(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
1 February 2007 (01.02.2007)

PCT

(10) International Publication Number
WO 2007/014282 A2

(51) International Patent Classification:
C12N 15/09 (2006.01)

(21) International Application Number:
PCT/US2006/029068

(22) International Filing Date: 26 July 2006 (26.07.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/702,477	26 July 2005 (26.07.2005)	US
60/710,905	24 August 2005 (24.08.2005)	US

(71) Applicant *(for all designated States except US)*: **WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH** [US/US]; Nine Cambridge Center, Cambridge, MA 02142-1479 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SABATINI, David, M.** [US/US]; 1105 Massachusetts Avenue, Apartment 6b, Cambridge, MA 02138 (US). **ALI, Siraj, M.** [US/US]; 44 Francis Street, Apartment 2, Boston, MA 02115 (US). **BAILEY, Steve, N.** [US/US]; 44 Franscis Street, Apartment 2, Boston, MA 02115 (US).

(74) Agent: **BELLIVEAU, Michael, J.**; Clark & Elbing LLP, 101 Federal Street, Boston, MA 02110 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2007/014282 A2

(54) Title: TRANSFECTION METHOD AND USES RELATED THERETO

(57) Abstract: The invention features a method of introducing nucleic acid molecules into cells by (a) affixing VSV-G pseudotyped viral vectors onto the surface of an object; and (b) contacting cells and the affixed viral vectors under appropriate conditions for entry of the viral vectors into the cells, whereby the nucleic acid molecules are introduced into the cells.

TRANSFECTION METHOD AND USES RELATED THERETO

Background of the Invention

Genome and expressed sequence tag (EST) projects are rapidly cataloging and cloning the genes of higher organisms, including humans. The emerging challenge is to uncover the functional roles of the genes and to quickly identify gene products with desired properties. The growing collection of gene sequences and cloned cDNAs demands the development of systematic and high-throughput approaches to characterizing the gene products. The uses of nucleic acid microarrays for transcriptional profiling and of yeast two-hybrid arrays for determining protein-protein interactions are recent examples of genomic approaches to the characterization of gene products.

Summary of the Invention

Here we demonstrate Lentiviral Cell-Based Microarray (LCBM) technology. We show robust and addressable infection of mammalian cells in a microarray based format with VSV-G pseudotyped lentiviruses encoding a variety of recombinant genes. We also demonstrate that LCBMs can infect primary dividing and non-dividing cells. Furthermore, this technology is compatible with RNA interference (RNAi), as lentiviruses expressing short hairpin RNAs (shRNAs) silence targeted genes *in vivo*. Previously printed and stored microarrays of virus are durable, as microarrays over a year old were competent to infect cells comparably to freshly printed microarrays. This technology provides a platform of broad utility for the investigation of gene function by loss-of-function and overexpression screens in mammalian cells.

In a first aspect, the invention features a method of introducing a target nucleic acid molecule into cells. This method includes the steps of: (a) affixing a plurality of VSV-G pseudotyped viral vectors onto a surface, the viral vectors including the target nucleic acid molecule; and (b) contacting eukaryotic cells (e.g., mammalian cells) with the affixed viral vectors) under appropriate conditions for entry of the VSV-G

pseudotyped viral vectors into the eukaryotic cells. Exemplary target nucleic acid molecules are those encoding polypeptides and those encoding double-stranded RNA molecules.

In a second aspect, the invention features a method of identifying a nucleic acid molecule capable of post-transcriptional gene silencing. This method includes the steps of: (a) affixing a plurality of VSV-G pseudotyped viral vectors onto a surface in discrete, defined locations, each of the viral vectors including a target nucleic acid molecule (e.g., encoding a double-stranded RNA molecule); (b) contacting eukaryotic cells (e.g., mammalian cells) with the affixed VSV-G pseudotyped viral vectors of step (a) under appropriate conditions for entry of the nucleic acid molecules into the eukaryotic cells; whereby the VSV-G pseudotyped viral vectors are introduced into the eukaryotic cells in the location in which the VSV-G pseudotyped viral vectors were affixed; and (c) determining the ability of the target nucleic acid molecules in the VSV-G pseudotyped viral vectors to post-transcriptionally silence expression of a gene in the cells. Post-transcriptional gene silencing at a discrete, defined location identifies the target nucleic acid molecules affixed at the location as being capable of post-transcriptional gene silencing.

In one embodiment, the affixed plurality of VSV-G pseudotyped viral vectors form an array of VSV-G pseudotyped viral vectors, and the cells into which theVSV-G pseudotyped viral vectors are introduced form an array of cells.

In a third aspect, the invention provides a method of making an array comprising steps of: (a) depositing a plurality of viral vectors onto a surface in discrete, defined locations, each of the viral vectors including a target nucleic acid molecule; and (b) maintaining the surface under conditions in which the viral vectors become affixed to the surface. The viral vectors may be deposited in a mixture that includes a liquid, and step (b) of the method may comprise allowing the liquid to evaporate whereby the viral vectors remain stably associated with the surface.

The array may include at least 96, 192, 300, 1,000, 5,000, or even 10,000 different discrete, defined locations of known sequence composition. In some embodiments the array includes between 300 and 10,000 features or between 300 and

100,000 features. In certain embodiments of the invention, each of the defined locations is 100-200 μm in diameter and is 200-500 μm apart from its nearest neighboring location. In certain embodiments of the invention each of the defined locations is 100- 500 μm in diameter and is 100-500 μm apart from its nearest neighboring location. In certain embodiments of the invention each of the defined locations is 100-300 μm in diameter and is 100-300 μm apart from its nearest neighboring location. In certain embodiments of the invention the center-to-center spacing of the discrete, defined locations is between 500 μm and 2 mm. The invention also features an array of VSV-G pseudotyped viral vectors. This array includes a surface having an array of at least 100 features per square centimeter, each feature including one or more VSV-G pseudotyped viral vectors covalently or non-covalently affixed to the surface. These VSV-G pseudotyped viral vectors are capable of being transfected into a eukaryotic cell under appropriate conditions. The array includes eukaryotic cells disposed on the surface and capable of being transfected by the one or more VSV-G pseudotyped viral vectors of at least one feature under appropriate conditions.

The invention also features a method of forming an array of transfected eukaryotic cells by (a) providing a surface having an array of features, wherein each feature comprises one or more VSV-G pseudotyped viral vectors covalently or non-covalently affixed to the surface; (b) contacting the surface with eukaryotic cells; and (c) transfecting the one or more VSV-G pseudotyped viral vectors into the eukaryotic cells to form the array of transfected eukaryotic cells.

The invention also features a method of forming an array of transfected eukaryotic cells by (a) providing a surface having an array of features, wherein each feature comprises one or more VSV-G pseudotyped viral vectors reversibly affixed to the surface; (b) contacting the surface with eukaryotic cells; and (c) transfecting the one or more VSV-G pseudotyped viral vectors into the eukaryotic cells to form the array of transfected eukaryotic cells.

Any VSV-G pseudotyped viral vector may be employed in the methods and compositions of the invention. Particularly desirable viral vectors are lentiviral

vectors and Moloney murine leukemia virus-derived retroviral vectors. In certain embodiments of the methods and compositions of the invention, the VSV-G pseudotyped viral vectors affixed to the surface are admixed with a carrier. The carrier may facilitate affixing of the viral vector to the surface and/or enhance uptake of the viral vector by cells that are subsequently overlaid on the array.

In the methods and compositions of the invention, the features of the array can have a density of at least 1,000, 10,000, or even 100,000 different features per square centimeter. The features have a density of at most 1,000,000 different features per square centimeter. In one embodiment, at least one feature comprises at least two different nucleic acid molecules included either in the same or in different viral vectors.

All patents and publications referred to herein are incorporated by reference. Also incorporated by reference are U.S. Provisional Application Nos. 60/702,477 and 60/710, 905. In the case of a conflict or inconsistency between the specification and an incorporated reference, the specification shall control.

Brief Description of the Drawings

Fig. 1 shows LCBM printed with multiple virus types. Fig. 1A shows a large array of 300 spots containing either GFP encoding lentivirus or shRNA-lamin encoding lentivirus. Blue is Hoechst (nuclei), red is an anti-lamin-cy3, and green is GFP. Fig. 1B shows quantification of the fluorescent intensity of spots in the array. GFP was only detected in spots with cells infected with GFP lentivirus and lamin knockdown (KD) was only detected in spots infected with shRNA-lamin lentivirus.

Fig. 2 shows demonstration of the sensitivity and flexibility of LCBMs. Fig. 2A shows an array of GFP-overexpressing viruses with different concentrations. Over a defined range, the amount of GFP expression correlates with the concentration of virus printed in each spot (n=5). Fig. 2B shows 10x10 grids of GFP-overexpressing virus seeded with HeLa, A549, 293T, and DU145 cells. Each array shows clusters of cells overexpressing GFP in the predicted pattern. The average GFP expression was similar in all four cell types. Fig. 2C shows primary human

fibroblasts and primary mouse dendritic cells were seeded on arrays similar to Fig. 2B. In these slow and non-dividing cell-types, local expression of GFP is seen on all printed spots of virus in the predicted pattern.

Fig. 3 shows the use of the LCBMs with multiple virus types and the detection of complex phenotypes. Fig. 3A shows an array of three viruses that use two different promoters (PGK and Ubiq-C) to drive the over expression of three genes (GFP, RFP, and Thy1.1). Blue is Hoechst (nuclei), red is an anti-thy1.1-cy3 and green if GFP (RFP not shown). The viruses were able to infect and induce gene-specific over-expression in cell clusters on the arrays. At higher magnification (100X), the expression of Thy1.1 co-localizes with the cell membrane. An identical array on a separate slide was stored for 15 months at –80°C and then assayed in an identical fashion. At both low and high magnification, the stored array is identical to the freshly printed array. Fig. 3B shows an LCBM of three viruses encoding GFP, shRNA for Lamin and shRNA for mTOR. Blue is Hoechst, red is an anti-pS6-cy3 and green is GFP. GFP expression is seen only on the cells growing on the spot printed with GFP-overexpressing virus. Compared to the cells on spots with GFP or shLamin lentivirus, the cells growing on the spot printed with shTOR lentivirus have reduced levels of phosphorylated S6, a downstream effector of mTOR. At higher magnification, the arrows point to two cells. The uninfected cell is bigger and has typical level of pS6, while the cell infected with the mTOR virus is smaller and has a reduced level of pS6.

Detailed Description of the Invention

The present invention features methods and reagents for making and using a gene expression system useful for the functional analysis of many gene products in parallel. Cells are cultured on a solid surface printed in defined locations with different VSV-G pseudotyped viral vectors that can be taken up by the cells. Rather than having to recover the transfected construct (viral vector) to ascertain its identity, the identity is determined by the position of the transfectant (infected cells) of interest

on the array. The subject assay can be particular useful where cell phenotype is the read-out used to identify a construct of interest.

In one embodiment, the invention features a microarray-based system useful to analyze the function in cells of many genes in parallel. Cells are cultured on a glass slide printed in defined locations with solutions containing different viral vectors. For example, VSV-G pseudotyped viral vectors (e.g., lentiviral vectors or Moloney murine leukemia virus-derived retrovirus vectors) can be used as described in more detail in Example 1. Cells located on the printed areas take up the vectors, creating spots of localized infected cells within a lawn of non-infected cells.

Infected cell microarrays can be of broad utility for the high-throughput expression cloning of genes, particularly in areas such as signal transduction and drug discovery. Another application of particular interest features use of infected cell microarrays for high throughput RNAi-based loss of function studies in eukaryotic cells, e.g., mammalian or avian cells. A further application of particular interest utilizes infected cell microarrays for high throughput overexpression studies in eukaryotic cells, e.g., mammalian or avian cells.

In certain embodiments, the method of the invention makes use of nucleic acids of known sequence and/or source, affixed to a surface, such as a slide or well bottom, and cells that are plated onto the DNA spots and maintained under conditions appropriate for entry of the nucleic acids into the cells. The size of the nucleic acid spots and the density of the spots affixed to the surface can be adjusted depending on the conditions used in the methods. For example, the nucleic acid spots can be from about 100 μm to about 200 μm in diameter and can be affixed from about 100 μm to about 500 μm apart on the surface. If desired, the nucleic acids are deposited uniformly. Thus, when the cells are plated on the nucleic acids, a lawn of infected cells is produced.

In certain embodiments of the method of the invention the viral vectors are viruses (e.g., viral particles) that include nucleic acids of known sequence and/or source. In one embodiment, the viruses are affixed to a surface such as a slide or well bottom to create an array of virus spots. A variety of suitable surfaces are described

below. The spots include the nucleic acid and the virus that contains it. Cells are plated onto the virus spots and maintained under conditions appropriate for entry of the viruses into the cells. For purposes of convenience, the process by which viruses are introduced into cells may be referred to herein as "transfection" or "infection", and the cells that have taken up viruses may be referred to as being "transfected" or "infected". However, it should be understood that the virus need not be capable of carrying out the complete infectious cycle, and it should be understood that "transfection" does not imply any particular mechanism of viral entry. Typically infection of cells by viral particles involves binding to a cell surface receptor, as described elsewhere herein. The virus typically comprises a genome encapsulated in a surrounding envelope or capsid and is thus distinct from a DNA plasmid or other vector consisting or consisting essentially of nucleic acid.

The size of the virus spots and the density of the spots affixed to the surface can be adjusted depending on the conditions used in the methods. For example, the virus spots can be from about 100 μm to about 500 μm in diameter and can be affixed from about 100 μm to about 500 μm apart on the surface. Of course the spots could be spaced further apart, e.g., up to 2 or 3 mm apart in various embodiments of the invention.

Entry of the virus into the cells introduces the nucleic acid into the cells, wherein the nucleic acid is expressed or has an effect on or interacts with a cellular component or function. In the case of some viruses, the virus is not internalized but instead "injects" its genomic material into the target cell.

The nucleic acid of known sequence and/or source contained in the virus may, but need not be, incorporated into the viral genome. Once inside the cell, a copy of the viral genome or a portion thereof including the nucleic acid, may integrate into the genome of the cell and be inherited by progeny of the cell. In another embodiment the viral genome may be maintained as a replicable episome and inherited by progeny of the cell. In another embodiment the nucleic acid is transiently expressed in the cell following entry of the virus.

Detection of the effects of the introduced nucleic acids on the infected cells can be carried out by a variety of known techniques, such as immunofluorescence, in which a fluorescently labeled antibody that binds a protein of interest (e.g., a protein thought to be encoded by a reverse transfected DNA or a protein whose expression or function is altered through the action of the reverse transfected DNA) is used to determine if the protein is present in cells grown on the DNA spots.

As used herein, the term "nucleic acid" refers to polynucleotides such as deoxyribonucleic acid (DNA), and ribonucleic acid (RNA).

"Complementary DNA" or a "cDNA" as used herein includes recombinant genes synthesized by reverse transcription of mRNA and from which intervening sequences (introns) have been removed.

As used herein, the terms "heterologous nucleic acid" and "foreign nucleic acid" refer to a nucleic acid, e.g., DNA or RNA, that does not occur naturally as part of the cell or genome in which it is present or which is found in a location or locations in the genome that differs from that in which it occurs in nature. Heterologous DNA or RNA is not endogenous to the cell into which it is introduced, but has been obtained initially from another cell, from a virus, by chemical synthesis, etc. Examples of heterologous nucleic acid include, but are not limited to, DNA that encodes test polypeptides, receptors, reporter genes, transcriptional and translational regulatory sequences, selectable or traceable marker proteins, such as a protein that confers drug resistance. Examples of heterologous RNA include, but are not limited to, antisense RNA sequences, ribozymes, and double-stranded RNA (for inducing sequence-specific RNA interference).

The term "feature," as it is used in describing a transfection array (also referred to herein as an "infected cell array" or "viral vector array"), refers to an area of a substrate having a homogenous collection of a target sequence (or sequences in the case of certain co-transfection embodiments). One feature is different than another feature if the target sequences of the different features have different nucleotide sequences.

As used herein, a "desired phenotype" refers to a particular phenotype that the user of the subject method seeks to have selectively conferred on the host cell line upon expression of a target sequence.

As used herein, the terms "target nucleic acid" and "target sequence" refer to the component of a transfection array, e.g., the portion or portions of a nucleic acid being transfected into the host cells via the viral vector, which is of interest with respect to its ability to confer a change in the phenotype of the host cells. In general, though not always, the target nucleic acid will be that portion(s) of the nucleic acid of the transfection array that is varied from one portion of the array to the next. Thus in many embodiments of the invention the target nucleic acid will be that portion of the nucleic acid contained in the viral vector that is varied from one portion of the array to the next. The target nucleic acid can be a coding sequence for a protein, a "coding" sequence for an RNA molecule (e.g., which is transcribed into an anti-sense RNA sequence, a ribozyme or double-stranded RNA), or a regulatory sequence (e.g., as part of a reporter construct), to name but a few examples.

As used herein, the term "operatively linked" refers to the functional relationship of a nucleic acid with regulatory and effector nucleotide sequences, such as promoters, enhancers, transcriptional and translational start and stop sites, and other signal sequences. For example, operative linkage of DNA to a promoter refers to the physical and functional relationship between the DNA and the promoter such that the transcription of such DNA is initiated from the promoter by an RNA polymerase that specifically recognizes, binds to, and transcribes the DNA.

As used herein, the term "expression" refers to any number of steps comprising the process by which nucleic acids are transcribed into RNA, and (optionally) translated into peptides, polypeptides, or proteins. If the nucleic acid is derived from genomic DNA, expression may, if an appropriate eukaryotic host cell or organism is selected, include splicing of the RNA. The process may include a step in which RNA is reverse transcribed to produce DNA, which is then transcribed to produce RNA, optionally after integrating into a host cell genome.

As used herein, "recombinant cells" include any cells that have been modified by the introduction of heterologous nucleic acid. Control cells include cells that are substantially identical to the recombinant cells, but do not express one or more of the proteins encoded by the heterologous nucleic acid.

The terms "protein," "polypeptide," and "peptide" are used interchangeably herein.

The terms "recombinant protein," "heterologous protein," and "exogenous protein" are used interchangeably throughout the specification and refer to a polypeptide which is produced by recombinant techniques, wherein generally, a nucleic acid encoding the polypeptide is inserted into a suitable expression vector which is in turn used to transform a host cell to produce the heterologous protein.

As used herein, a "reporter gene construct" is a nucleic acid that includes a "reporter gene" operatively linked to at least one transcriptional regulatory sequence. Transcription of the reporter gene is controlled by these sequences to which they are linked. The activity of at least one or more of these control sequences is directly or indirectly regulated by the target receptor protein. Exemplary transcriptional control sequences are promoter sequences. A reporter gene is meant to include a promoter-reporter gene construct which is heterologously expressed in a cell.

In certain preferred embodiments, the viral vector array provides, in a single array, at least 10 different sequences, more preferably at least 100, 1000 or even 10,000 different, discrete sequences. Preferably, target sequences are arrayed in an addressable fashion, such as rows and columns where the substrate is a planar surface. If each feature size is about 100 microns on a side, each chip can have about 10,000 target sequence addresses (features) in a one centimeter square (cm^2) area. In certain preferred embodiments, the transfection array provides a density of at least 10^3 different features per square centimeter (10^3 sequences/cm^2), and more preferably at least 10^4 features/cm^2, 10^5 features/cm^2, or even at least 10^6 features/cm^2. Of course, lower densities are contemplated, such as at least 100 features/cm^2.

In certain embodiments, multiple different target sequences are present in a feature in order to promote co-infection of the host cells with at least two different

nucleic acids. Co-infection refers to the simultaneous introduction of two or more viral vectors into the same cell. If the viral vectors direct the expression of a gene product, such as a protein, RNA, or other gene product, the cell will then express both gene products at the same time.

Co-infections can be performed in cell microarrays if the solution spotted on the surface contains more than one VSV-G pseudotyped viral vector. The co-infection features can include, for example, 2-10 different nucleic acids per feature, 10-100 different nucleic acids per feature, or even more than 100 different nucleic acids per feature.

The capacity to co-infect cells in a cell microarray has many important uses. These include but are not limited to the ability to: infer the expression of a gene product by detecting the expression of a co-infected nucleic acid encoding a marker protein (e.g. GFP, luciferase, beta-galactosidase, or any protein to which a specific antibody is available), express all the components of a multi-subunit complex (e.g. the T-cell receptor) in the same cells, express all the components of a signal transduction pathway (e.g. MAP kinase pathway) in the same cells, and express all the components of a pathway that synthesizes a small molecule (e.g. polyketide synthetase). In addition, the capacity to co-infect allows the creation of microarrays with combinatorial combinations of co-expressed viral vectors. This capacity is particularly useful for implementing mammalian two-hybrid assays in which viral vectors encoding bait and prey proteins are co-infected into the same cells by spotting them in one feature of the microarray.

The capacity to co-infect is also useful when the goal is to promote differentiation of the infected cells along a certain tissue lineage. For example, combination of genes can be expressed in a stem or early progenitor cells that will force the differentiation of the cells into endothelial, liver, heart, pancreatic, lymphoid, islet, brain, lung, kidney or other cell types. In this fashion, arrays can be made with primary-like cells that can be used to examine interactions of protein or small molecules that are cell-type specific.

Furthermore, combinations of viral vectors can be printed in different patterns on the surface on which reverse transfection occurs. Patterns include, but are not limited to, bulls-eyes, squares, rectangles of varying heights and widths, and lines of single cell thickness. By printing, in particular patterns, combinations of nucleic acids that cause differentiation of cells into different tissue types, this technology can be used to obtain arrays with distinct cell types in distinct locations. This capacity can be useful when trying to create tissue-like structures on the array, such as blood capillaries and stromal structures, or when studying the response of one cell type to the protein secretions of another cell type. For example, a secreting cell type can be created in the center of a bulls-eye pattern and responder cell types of different tissues can be created on the edge of bulls-eye. The response of the responder cells to the secretions of the center cell can then be examined.

Arrays containing mixtures of viral vectors at each feature could be constructed by mixing viral vectors before printing, printing in serial, printing with masks, or printing with patterned printheads. For example, viral vectors could be mixed in a container before printing and printed as a homogenous mixture. Alternatively, viral vectors could be printed on top of one another or close to one another. In this method, the exact composition of the mixture containing each viral vectors could be modified to control the sequencing and timing of their entry into a cell, e.g. slower or faster release mixtures. Masks with different patterns of holes or print heads with different configurations could also be used to print combinations of viral vectors. For example, different enzymes involved in polyketide synthesis could be combined to generate different polyketides.

It will be appreciated that any suitable method could be used to deposit viral vectors on a surface, of which printing (contact printing, inket printing, etc.) is a convenient approach. Viral vectors could be deposited by hand, e.g., using a multipipette apparatus.

While certain embodiments employ arrays containing mixtures of viral vectors at each feature, it will be appreciated that non-array based methods are also encompassed by the invention.

The carrier for use in certain of the methods of the present invention can be any carrier that, when admixed with polynucleotides or viruses, facilitates their becoming affixed to a surface, and/or increases the transfection of the affixed polynucleotides or viruses into the plated cells. Suitable carriers include gelatin, fibronectin, acrylamide polycarboxylic acid, cellulosic polymer, polyvinylpyrrolidone, maleic anhydride polymer, polyamide, polyvinyl alcohol, polyethylene glycol, polyethylene oxide, carbohydrates (e.g., monosaccharides, disaccharides, polysaccharides), amino acids, peptides, proteins, lipids (e.g., fatty acids, triacylglycerides, waxes, phospholipids, sphingolipids, polar lipids), steroids, lipoproteins, vitamins, alcohols, aldehydes, ketones, carboxylic acids, alkanes, alkyl halides, organometallic compounds, ethers, alkenes, alkynes, nitriles, aromatic compounds, amines, isocyanates, carbamates, ureas, azides, diazo compounds, diazonium salts, thiols, sulfides, phosphines, phosphonium salts, carbanions. organosilicon compounds, aromatic halides, phenols, phenyl ethers, quinines, and heterocyclic compounds. In certain embodiments of the invention the carrier is a positively charged species, e.g., a polycation. The carrier, if used, is typically present in the mixture to be deposited on the surface at a concentration of from 0.0001% to 10% (w/v), but may be present in smaller or greater amounts, e.g., from 0.001% to 10% (w/v) or from 0.01% to 1% (w/v), or from 0.1% to 1% (w/v).

Any suitable surface which can be used to affix the nucleic acid containing mixture (i.e., the mixture containing the viral vector) to its surface can be used. For example, the surface can be glass, plastics (such as polytetrafluoroethylene, polyvinylidenedifluoride, polystyrene, polycarbonate, polypropylene), silicon, metal, (such as gold), membranes (such as nitrocellulose, methylcellulose, PTFE or cellulose), paper, biomaterials (such as protein, gelatin, agar), tissues (such as skin, endothelial tissue, bone, cartilage), minerals (such as hydroxylapatite, graphite). Additional compounds may be added to the base material of the surface to provide functionality. For example, scintillants can be added to a polystyrene substrate to allow Scintillation Proximity Assays to be performed. The substrate may be a porous solid support or non-porous solid support. The surface can have concave or convex

regions, patterns of hydrophobic or hydrophilic regions, diffraction gratings, channels or other features. The scale of these features can range from the meter to the nanometer scale. For example, the scale can be on the micron scale for microfluidics channels or other MEMS features or on the nanometer scale for nanotubes or buckyballs. The surface can be planar, planar with raised or sunken features, spherical (e.g. optically encoded beads), fibers (e.g. fiber optic bundles), tubular (both interior or exterior), a 3-dimensional network (such as interlinking rods, tubes, spheres) or other shapes. The surface can be part of an integrated system. For instance, the surface can be the bottom of a microtiter dish, a culture dish, a culture chamber. Other components, such as lenses, gratings, and electrodes, can be integrated with the surface. In general, the material of the substrate and geometry of the array will be selected based on criteria that it be useful for automation of array formation, culturing and/or detection of cellular phenotype.

In still other embodiments, the solid support is a microsphere (bead), especially a FACS sortable bead. Preferably, each bead is an individual feature, e.g., having a homogenous population of viral vectors and distinct from most other beads in the mixture, and one or more tags which can be used to identify any given bead and therefore the target sequence it displays. The identity of any given target sequence that can induce a FACS-detectable change in cells that adhere to the beads can be readily determined from the tag(s) associate with the bead. For example, the tag can be an electrophoric tagging molecules that are used as a binary code (Ohlmeyer et al. (1993) PNAS 90:10922-10926). Exemplary tags are haloaromatic alkyl ethers that are detectable as their trimethylsilyl ethers at less than femtomolar levels by electron capture gas chromatography (ECGC). Variations in the length of the alkyl chain, as well as the nature and position of the aromatic halide substituents, permit the synthesis of at least 40 such tags, which in principle can encode 2^{40} (e.g., upwards of 10^{12}) different molecules. A more versatile system has, however, been developed that permits encoding of essentially any combinatorial library. Here, the compound would be attached to the solid support via the photocleavable linker and the tag is attached through a catechol ether linker via carbene insertion into the bead matrix (Nestler

et al. (1994) J. Org. Chem. 59:4723-4724). This orthogonal attachment strategy permits the FACS sorting of the cell/bead entities and subsequent decoding by ECGC after oxidative detachment of the tag sets from isolated beads. In other embodiments, the beads can be tagged with two or more fluorescently active molecules, and the identity of the bead is defined by the ratio of the various fluorophores.

In still another embodiment, the VSV-G pseudotyped viral vector can be disposed on the end of a fiber optic system, such as a fiber optic bundle. Each fiber optic bundle contains thousands to millions of individual fibers depending on the diameter of the bundle. Changes in the phenotype of cells applied to the transfection array can be detected spectrometrically by conductance or transmittance of light over the spatially defined optic bundle. An optical fiber is a clad plastic or glass tube wherein the cladding is of a lower index of refraction than the core of the tube. When a plurality of such tubes are combined, a fiber optic bundle is produced. The choice of materials for the fiber optic will depend at least in part on the wavelengths at which the spectrometric analysis of the infected cells is to be accomplished.

In addition, the surface can be coated with, for example, a cationic moiety. The cationic moiety can be any positively charged species capable of electrostatically binding to negatively charged polynucleotides or viruses. Preferred cationic moieties for use in the carrier or for purposes of coating the surface are polycations, such as polylysine (e.g., poly-L-lysine), polyarginine, polyornithine, spermine, basic proteins such as histones (Chen et al. (1994) FEBS Letters 338:167-169), avidin, protamines (see e.g., Wagner et al. (1990) PNAS 87: 3410-3414), modified albumin (i.e., N-acylurea albumin) (see e.g., Huckett et al. (1990) Chemical Pharmacology 40: 253-263), and polyamidoamine cascade polymers (see e.g., Haensler et al. (1993) Bioconjugate Chem. 4:372-379). A preferred polycation is polylysine (e.g., ranging from 3,800 to 60,000 daltons). Alternatively, the surface itself can be positively charged (such as gamma amino propyl silane (GAPS) or other alkyl silanes, which can be used as a coating material). Without wishing to be bound by any theory, electrostatic interactions between a negatively charged viral envelope and a positively

charged surface may mediate or facilitate stable association between the virus and the surface.

The surface can also be coated with molecules for additional functions. For instance, these molecules can be capture reagents such as antibodies, biotin, avidin, Ni-NTA to bind epitopes, avidin, biotinylted molecules, or 6-His tagged molecules. Alternatively, the molecules can be culture reagents such as extracellular matrix, fetal calf serum, collagen. In certain embodiments of the invention the surface is coated with an affinity ligand that binds to the viral vector. For example, the surface may be coated with a lectin that binds to sugar molecules present on the cell surface of the virus. Any binding moiety that interacts with a component present at the surface of the viral envelope or capsid could be used as an affinity ligand to promote stable association of the virus with the surface.

In certain embodiments, once the arrays of infected cells have formed (i.e., the viral vectors in the spots have entered cells), the arrays can be transferred onto a variety of surfaces. Surfaces can be flexible or non-flexible and porous or non-porous. The surfaces can be flat or patterned with concave or convex regions, patterns of hydrophobic or hydrophilic regions, diffraction gratings, channels or other features. The scale of these features can range from the meter to the nanometer scale. Examples of surfaces include, but are not limited to, glass, plastics (such as polytetrafluoroethylene, polyvinylidenedifluoride, polystyrene, polycarbonate, polypropylene), silicon, metal, (such as gold), membranes (such as nitrocellulose, methylcellulose, PTFE or cellulose, polyvinylidene fluoride (PVDF)), paper, biomaterials (such as protein, gelatin, agar), tissues (such as skin, endothelial tissue, bone, cartilage), minerals (such as hydroxylapatite, graphite). Furthermore, many of these surfaces can be derivatized to provide additional functionalities. For example, scintillants can be added to a polystyrene substrate to allow Scintillation Proximity Assays to be performed. In another example, nitrocellulose membranes can be covalently modified with metal chelators that immobilize metals, such as nickel or cobalt, and allow the selective binding of proteins carrying a specific amino acid sequence, such as a hexa-histidine tag (6X His).

Transfers can be performed so that 1) the entire cellular material on the array is transferred or 2) so that only the infected cells are transferred. The transfer of the array to another surface is accomplished by directly contacting the array to the other surface and allowing the material to move to the new surface under the influence of a force, such as capillary forces (commonly referred to as "blotting"), electric or magnetic fields, vacuum suction forces, or other forces. The material binds to the new surface through an interaction mediated by hydrophobic, hydrophillic, Van der Waals, ionic or other forces, or through specific receptor-ligand interactions (e.g. antibody-epitope interactions) or by becoming entangled in the molecular structure of the other surface.

The ability to transfer cellular material from the arrays to another surface has many important uses. These include, but are not limited to, the capacity to detect cellular phenotypes or protein properties using techniques normally performed on specific surfaces and the capacity to in parallel purify the recombinant gene products expressed in the array. Examples of techniques normally performed on specific surfaces include western blotting, far-western blotting, southwestern blotting, surface plasmon resonance (SPR), mass spectroscopy, and others. These techniques normally require the immobilization of native or denatured proteins on nitrocellulose, nylon, paper, polyvinylidene fluoride (PVDF), or gold or other metal surfaces or membranes. Southwestern blotting is used to detect the interaction of a nucleic acid (such as DNA or RNA) with a protein. After transfer to an appropriate membrane, arrays of cells expressing a collection of DNA binding proteins, such as transcription factors, could be used to identify binding proteins for genomic DNA sequence elements.

The transfer of arrays to other surfaces is also useful for the in parallel purification of the recombinant proteins expressed on the microarray. In one embodiment of this approach, all the recombinant proteins expressed on the microarray contain an amino acid sequence that is a ligand for a specific protein or chemical reagent (e.g. an epitope recognized by a polyclonal or monoclonal antibody or a hexa-histidine tag recognized by a nickel affinity matrix). Arrays expressing these proteins are then transferred by direct contact to a surface that has been

derivatized with the reagent that binds the ligand (e.g. a nitrocellulose membrane to which an anti-epitope monoclonal antibody is bound or a nitrocellulose membrane derivatized with a metal chelator that allows the binding of nickel to its surface). After the material has bound to the new surface, the surface is washed with an appropriate buffer that does not disrupt the specific interaction but eliminates non-specific interactions with the surface. Non-specific interactions include but are not limited to the interactions of any cellular components that do not contain the specific ligand recognized by the surface to which the microarray has been transferred. The array of recombinant proteins can then used to detect the interaction of other proteins or small molecules with the array. The binding of proteins or small molecules with the array can be detected with autoradiography, fluorescence, mass spectroscopy, immunofluorescence, or calorimetry.

Viral Vectors

Any of a wide variety of viruses may be used in viral vector arrays and methods of the present invention. The use of lentiviruses in the methods and arrays of the invention is exemplified herein, but each aspect of the invention encompasses use of other viruses. Lentiviral vectors are attractive in part because, like other retroviral vectors, following their entry into a cell a DNA copy of the viral RNA genome is synthesized and integrates into the genome of the cell, which can allow for stable expression of a target nucleic acid included in the viral vector. Any retroviral vector may be used in the present invention. Lentiviruses are retroviruses of particular interest, in part because of their ability to transduce non-dividing cells. Exemplary lentiviruses may include sequences derived from any of a wide variety of lentiviruses including, but not limited to, primate lentivirus group viruses such as human immunodeficiency viruses HIV-1 and HIV-2 or simian immunodeficiency virus (SIV); feline lentivirus group viruses such as feline immunodeficiency virus (FIV); ovine/caprine immunodeficieny group viruses such as caprine arthritis encephalitis virus (CAEV); bovine immunodeficiency-like virus (BIV); equine lentivirus group viruses such as equine infectious anemia virus; and visna/maedi virus. It will be

appreciated that each of these viruses exists in multiple variants or strains. Exemplary retroviruses include Moloney murine leukemia virus (MoMuLV or MMLV), Harvey murine sarcoma virus (HaMuSV or HSV), murine mammary tumor virus (MuMTV or MMTV), gibbon ape leukemia virus (GaLV or GALV), Rous sarcoma virus (RSV), avian sarcoma and leucosis virus, and spleen necrosis virus (SNV).

Any of a wide variety of viruses capable of infecting eukaryotic cells, e.g., mammalian or avian cells, could be used in the present invention. The invention contemplates use of any virus known in the art to be useful for expressing heterologous nucleic acids in eukaryotic cells. In certain embodiments of the invention the virus is an enveloped virus. Examples include adenovirus, adeno-associated virus, herpes viruses such as herpes simplex virus or Epstein-Barr virus, baculovirus, measles virus, hepatitis virus, etc. It will typically be desirable to select a virus with a tropism that allows it to infect a cell type of interest that will be used to form the cell based array. For example, if hepatocytes are of interest, a hepatitis virus may be a suitable choice since these viruses are able to infect hepatocytes with high efficiency. Influenza virus is capable of infecting respiratory epithelial cells. In general, one would select a virus that, either naturally or as a result of pseudotyping (see below), is able to infect a cell type of interest, e.g., displays an envelope or capsid protein for which the cell type of interest has a receptor. The receptor may or may not be known. In some embodiments of the invention a virus capable of infecting both human and rodent (e.g., mouse) cells is used.

In certain embodiments of the invention the viral vector is pseudotyped. In general, a pseudotyped virus is one in which at least some components of the outer shell (e.g., the envelope glycoproteins of an enveloped virus or the capsid proteins of a nonenveloped virus) originates from a virus that differs from the source of the genome and the genome replication apparatus. For example, a pseudotyped retrovirus would differ from its non-pseudotyped counterpart in that its envelope would incorporate a non-retroviral envelope protein, or an envelope protein from a different retrovirus, instead of, or in some cases in addition to, the native retroviral envelope

protein. The two viruses may differ considerably (e.g., retrovirus and rhabdovirus), or they may be closely related (e.g., two different retroviruses or different serotypes of a virus). Pseudotyping makes it possible to alter the range of cell types and/or species that the virus can infect. Pseudotyping a viral vector can provide it with an expanded set of target cells or can restrict it to specific cells that are of experimental or therapeutic interest. A pseudotyped vector can have an altered stability and/or interaction with the host cell or organism. One example is pseudotyped viral vectors that can be produced and/or concentrated to higher transduction titers than the viral vector with its native outer shell. The present invention contemplates use of envelope proteins that confer any of these or other desired characteristics on the virus.

Vesicular stomatitis virus (VSV) is a rhabodvirus able to infect a broad range of cell types, likely as a result of interactions of its envelope glycoprotein (VSV-G) with cell surface molecules present on these cells. Thus VSV-G pseudotyped viral vectors are of particular interest in the present invention. It will be appreciated that VSV-G from any VSV serotype (e.g., New Jersey or Indiana) could be used. Furthermore, one of skill in the art will appreciate that VSV-G pseudotyped viral vectors may employ naturally occurring variants or engineered forms of VSV-G so long as the ability of the protein to mediate entry into cells is preserved.

In one embodiment, an inducible promoter system is used so that VSV-G expression can be regulated, e.g., turned off, when it is not required (e.g., after infection). For example, the tetracycline-regulatable gene expression system (Gossen & Bujard, *Proc. Natl. Acad. Sci.* 89:5547-5551, 1992) and more recently introduced variants thereof can be employed to provide for inducible or repressible expression of VSV-G. To this end, the VSV-G coding sequence may be cloned downstream from a promoter controlled by *tet* operator sequences. Other inducible/repressible systems are known in the art.

A wide range of other viral envelope glycoproteins or capsid proteins could be used in the pseudotyped viral vectors. Examples include viral envelope proteins from any of the afore-mentioned lentiviruses or retroviruses. Envelope glycoproteins from other rhabdoviruses such as rabies virus or rabies related viruses such as Mokola or

Ebola virus, alphaviruses such as Ross River virus, arenaviruses such as lymphocytic choriomeningitis virus, hepatitis B or C virus, or influenza virus, could be used. It will be appreciated that there are multiple naturally occurring strains or serotypes of some of these viruses, as well as engineered variants, and envelope proteins from any of these could be used in the present invention.

In some embodiments a viral vector is pseudotyped with a protein that includes a portion (binding moiety) selected to bind to a receptor expressed by a target cell (e.g., present at the cell surface), or selected to bind to a ligand that is coated onto the surface on which the array is to be formed. Interaction between the virus and the ligand promotes stable association of the virus and the surface. By inserting a nucleic acid sequence of interest (optionally including a regulatory region) into the viral vector, along with another gene which encodes the ligand for a receptor on a specific cell type the vector can be made specific for particular cells. Targeting may be accomplished by using an antigen-binding portion of an antibody or a recombinant antibody-type molecule, such as a single chain antibody, to target the viral vector.

Methods of generating recombinant viruses containing a target nucleic acid sequence are known in the art. See, e.g., U.S. Pat. Nos. 6,013,516; 6,682,907; and U.S. Pub. No. 20050251872 for discussion of systems and methods for preparing recombinant retroviral, e.g., lentiviral particles. One of skill in the art will readily be able to prepare recombinant viruses of the various types mentioned herein.

A pseudotyped viral vector may be produced using standard recombinant DNA methods to replace a portion of the viral genome with a sequence that encodes the desired protein that is to be incorporated into the viral envelope or capsid. The introduced sequence may, but need not, replace all or part of the native envelope or capsid protein gene. The introduced sequence may be positioned so that it is operatively linked to expression control sequences such as a promoter already present in the viral genome or may include such expression control sequences. High titer preparations of pseudotyped viral vectors may be produced using standard methods. In one embodiment an optimized method for large-scale production of high titer lentivirus vector pseudotypes is employed (Sena-Esteves, M., et al., *J Virol Methods.*

22(2):131-9, 2004), which can be used to prepare large amounts of virus having titers above 10^9 infectious units (IFU)/ml for a wide range of pseudotyped lentiviral vectors. In some embodiments the invention contemplates use of retroviral, e.g., lentiviral, solutions having titers of between 10^9 and 10^{10} IFU/ml or even higher, e.g., between 10^{10} and 10^{11} IFU/ml. Of course lower tites, e.g., between 10^8 and 10^9 IFU/ml may be used. The viral vectors may be stored in a solution that allows for their stable preservation at low temperatures. The same solution may be used for depositing the viral vector on the surface. Alternately, one or more additional components may be added to the solution prior to depositing on the surface.

A variety of components may be included in the viral vector solution. The solution may contain a buffer, e.g., HEPES or another physiologically acceptable buffer, etc. The solution may include a suitable amount of a stabilizing agent or preservative. The stabilizing agent may contribute to maintaining the virus in an infectious state during the process of freezing or freeze drying and/or while maintained in a low temperature state either prior to deposition on the surface or thereafter. Suitable agents and concentrations thereof are known in the art and include carbohydrates, e.g., disaccharides such as trehalose, etc. In certain embodiments of the invention an infection-enhancing agent is included in the solution. Suitable agents include those known in the art to enhance viral infection. Examples include positively charged compounds, e.g., polycations such as protamine sulfate. Without wishing to be bound by any theory, such compounds may facilitate interaction between the negatively charged viral envelope and negatively charged cell surface. Other suitable components include DEAE-dextran, dextran sulfate, polybrene, heparan sulfate, etc. Any of the various components may be present in the mixture to be deposited on the surface at a concentration of from 0.0001% to 10% (w/v), but may be present in smaller or greater amounts, e.g., from 0.001% to 10% (w/v) or from 0.01% to 1% (w/v), or from 0.1% to 1% (w/v).

The viral vector solutions may be processed to facilitate their printing onto a surface. In one embodiment, multiple different viral vector solutions are placed into a multiwell plate, e.g., a 384 well plate, and centrifuged. The volume of solution

printed may vary depending, e.g., on the size of the print head used and the desired size of the features. Conveniently, between 1 nl and 10 nl of solution, e.g., between 2 and 5 nl of solution, may be deposited on the surface. The numbef of IFU/spot may vary. For example, the number of IFU/spot may be 100-100,000; 500-20,000; or 1000-10,000, etc. One aspect of the present invention is a method of preparing a viral vector array comprising steps of: (a) preparing a plurality of viral vectors containing different target nucleic acid sequences; (b) distributing the viral vectors into a multiwell plate having a defined, discrete pattern of wells; and (c) depositing the viral vectors onto a surface. If described the viral vectors are deposited in a highly parallel manner such as by using a multipin printing device or multijet inket printer. Optionally the method includes a step of centrifuging the multiwell plate. "Multiwell plate" refers to any container that contains a plurality of wells or vessels suitable for holding solutions in discrete, defined locations, preferably in a planar grid of mutually perpendicular rows and columns. The viral vector solutions may be allowed to dry on the surface for variable periods of time to allow evaporation of the liquid, whereby the viruses are affixed to the surface.

In some embodiments, the viral solution contains a salt, e.g., KCl, NaCl, etc. Other salts, e.g., those containing a monovalent or divalent cation such as Mg^{++}, Ca^{++}, etc., could also be used. Without wishing to be bound by any theory, the salt may help to constrain spreading of the spots of solution deposited on the slide. The concentration of the salt may range, e.g., from 100 mM to 2 M, e.g., from 500 mM to 1.5 M, or from 1.0 M to 1.4 M.

Cells

Suitable host cells for generating the subject assay include animal cells, especially mammalian cells, and even more preferably are primate cells such as human cells. Other preferred species of mammalian cells include canine, feline, bovine, porcine, mouse and rat. For instance, such cells can be hematopoietic cells, neuronal cells, pancreatic cells, hepatic cells, chondrocytes, osteocytes, or myocytes. The cells can be fully differentiated cells or progenitor/stem cells. They may be

primary cells or cell lines. They may be dividing or non-dividing cells. Moreover, the cells can be derived from normal or diseased tissue, from differentiated or undifferentiated cells, from embryonic or adult tissue.

The cells may be dispersed in culture, or can be tissues samples containing multiple cells that retain some of the microarchitecture of the organ.

In certain embodiments, the method of the invention is used to infect a cell that can be co-cultured with a target cell. A biologically active protein secreted by the cells expressing a recombinant nucleic acid will diffuse to neighboring target cells and induce a particular biological response, such as proliferation or differentiation, or activation of a signal transduction pathway which is directly detected by other phenotypic criteria. Likewise, antagonists of a given factor can be selected in similar fashion by the ability of the cell producing a functional antagonist to protect neighboring cells from the effect of exogenous factor added to the culture media. The host and target cells can be in direct contact, or separated, e.g., by a cell culture insert (e.g. Collaborative Biomedical Products, Catalog #40446).

The choice of appropriate host cell will also be influenced by the choice of detection signal. For instance, reporter constructs can provide a selectable or screenable trait upon gain-of-function or loss-of-function induced by a target nucleic acid. The reporter gene may be an unmodified gene already in the host cell pathway, or it may be a heterologous gene (e.g., a reporter gene construct). In other embodiments, second messenger generation can be measured directly in a detection step, such as mobilization of intracellular calcium or phospholipid metabolism, in which case the host cell should have an appropriate starting phenotype for activation of such pathways.

The host cells are plated (placed) onto the surface bearing the viral vectors, e.g., VSV-G pseudotyped viral vectors, in sufficient density and under appropriate conditions for introduction/entry of the nucleic acid into the cells. Preferably, the host cells (in an appropriate medium) are plated on the array at high density (e.g., on the order of 0.5-1 x $10^5/cm^2$), in order to increase the likelihood that infection will occur. For example, the density of cells can be from about 0.3 x $10^5/cm^2$ to about 3 x

$10^5/cm^2$, and in specific embodiments, is from about 0.5 x $10^5/cm^2$ to about 2 x $10^5/cm^2$ and from about 0.5 x $10^5/cm^2$ to about 1 x $10^5/cm^2$. In exemplary embodiments the number of cells plated onto a viral vector array on a glass slide is between 1 million and 100 million, e.g., between 10 million and 50 million. Of course the number of cells would be correspondingly larger in the case of larger viral vector arrays. The size of the spots of infected cells generated by the viral vectors of the array, and their center-to-center distance, can vary depending, e.g., on the size of the viral vector spots and on the size of the cells.

In certain embodiments, the host cells can be engineered to express other recombinant genes. For instance, the host cells can be engineered with a reporter gene construct, and the ability of nucleic acids in the VSV-G pseudotyped viral vectors to alter the level of expression of the reporter gene can be assessed. For example, the transfection array can be assessed for members that encode transcriptional activators or transcriptional repressors of the reporter gene, and may include native and non-native sequences. For instance, the host cell can express a reporter gene construct including a promoter sequence for which a protein that binds that sequence is sought. The VSV-G pseudotyped viral vectors can encode a library of potential DNA binding domains fused to a polymerase activation domain. Members of the library are selected by their ability to induce expression of the reporter gene. Conversely, the DNA binding specificity of a DNA binding protein can be determined by arraying a library of reporter gene constructs that are variegated with respect to the sequence of a transcriptional regulatory element. The cell also expresses the DNA binding protein, e.g., which naturally or by engineering includes a transcriptional activation domain. Those members of the reporter gene construct library that include appropriate regulatory sequences are expressed, and the position of those constructs in the array used to determine the consensus sequence for the DNA binding protein.

In other instances, the host cells can be engineered so as to have a loss-of-function or gain-of-function phenotype, and the ability of the ability of nucleic acids in the VSV-G pseudotyped viral vectors to counteract such a phenotype is assessed.

In still other instances, the host cells are engineered to express a recombinant cell surface receptor, and the VSV-G pseudotyped viral vectors encode a variegated library of gene products or peptides, and the ability of one or more members of that library to induce or inhibit signal transduction by the receptor is assessed. For instance, the VSV-G pseudotyped viral vectors can provide a library of secreted peptides, and the ability of a given peptide to induce signal transduction is detected by the conversion of the cell to an autocrine phenotype.

Detection

A variety of methods can be used to detect the consequence of uptake and expression of VSV-G pseudotyped viral vectors. In a general sense, the assay provides the means for determining if the target sequence is able to confer a change in the phenotype of the cell relative to the same cell but which lacks the target sequence. Such changes can be detected on a gross cellular level, such as by changes in cell morphology (membrane ruffling, rate of mitosis, rate of cell death, mechanism of cell death, dye uptake, and the like). In other embodiments, the changes to the cell's phenotype, if any, are detected by more focused means, such as the detection of the level of a particular protein (such as a selectable or detectable marker), or level of mRNA or second messenger, to name but a few. Changes in the cell's phenotype can be determined by assaying reporter genes (encoding, e.g., beta-galactosidase, green fluorescent protein, beta-lactamase, luciferase, or chloramphenicol acetyl transferase), assaying enzymes, using immunoassays, staining with dyes (e.g. DAPI, calcofluor), assaying electrical changes, characterizing changes in cell shape, examining changes in protein conformation, and counting cell number. Other changes of interest could be detected by methods such as chemical assays, light microscopy, scanning electron microscopy, transmission electron microscopy, atomic force microscopy, confocal microscopy, image reconstruction microscopy, scanners, autoradiography, light scattering, light absorbance, NMR, PET, patch clamping, calorimetry, mass spectrometry, surface plasmon resonance, time resolved fluorescence. Data could be collected at single or multiple time points and analyzed by the appropriate software.

In one example, immunofluorescence can be used to detect a protein. Alternatively, expression of proteins that alter the phosphorylation state or subcellular localization of another protein, proteins that bind with other proteins or with nucleic acids or proteins with enzymatic activity can be detected.

When screening for bioactivity of test compounds, intracellular second messenger generation can be measured directly. Such embodiments are useful where, for example, the arrayed library is being screened for target sequences that activate or inactivate a particular signaling pathway. A variety of intracellular effectors have been identified as being receptor- or ion channel-regulated, including adenylyl cyclase, cyclic GMP, phosphodiesterases, phosphoinositidases, phosphoinositol kinases, and phospholipases, as well as a variety of ions.

In still another embodiment, a heterologous reporter gene construct can be used to provide the function of an indicator gene. Reporter gene constructs are prepared by operatively linking a reporter gene with at least one transcriptional regulatory element. If only one transcriptional regulatory element is included it must be a regulatable promoter. At least one the selected transcriptional regulatory elements must be indirectly or directly regulated by the activity of the selected cell-surface receptor whereby activity of the receptor can be monitored via transcription of the reporter genes.

Transcriptional control elements for use in the reporter gene constructs, or for modifying the genomic locus of an indicator gene include, but are not limited to, promoters, enhancers, and repressor and activator binding sites. Suitable transcriptional regulatory elements may be derived from the transcriptional regulatory regions of genes whose expression is linked to the desired phenotype sought from the arrayed library.

In the case of receptors which modulate cyclic AMP, a transcriptional based readout can be constructed using the cyclic AMP response element binding protein, CREB, which is a transcription factor whose activity is regulated by phosphorylation at a particular serine (S133). When this serine residue is phosphorylated, CREB binds to a recognition sequence known as a CRE (cAMP Responsive Element) found to the

5' of promoters known to be responsive to elevated cAMP levels. Upon binding of phosphorylated CREB to a CRE, transcription from this promoter is increased.

Phosphorylation of CREB is seen in response to both increased cAMP levels and increased intracellular Ca levels. Increased cAMP levels result in activation of PKA, which in turn phosphorylates CREB and leads to binding to CRE and transcriptional activation. Increased intracellular calcium levels results in activation of calcium/calmodulin responsive kinase II (CaM kinase II). Phosphorylation of CREB by CaM kinase II is effectively the same as phosphorylation of CREB by PKA, and results in transcriptional activation of CRE containing promoters. Therefore, a transcriptionally-based readout can be constructed in cells containing a reporter gene whose expression is driven by a basal promoter containing one or more CRE. Changes in the intracellular concentration of Ca++ (a result of alterations in the activity of the receptor upon engagement with a ligand) will result in changes in the level of expression of the reporter gene if: a) CREB is also co-expressed in the cell, and b) either an endogenous or heterologous CaM kinase phosphorylates CREB in response to increases in calcium or if an exogenously expressed CaM kinase II is present in the same cell. In other words, stimulation of PLC activity may result in phosphorylation of CREB and increased transcription from the CRE-construct, while inhibition of PLC activity may result in decreased transcription from the CRE-responsive construct.

In preferred embodiments, the reporter gene is a gene whose expression causes a phenotypic change that is screenable or selectable. If the change is selectable, the phenotypic change creates a difference in the growth or survival rate between cells that express the reporter gene and those that do not. If the change is screenable, the phenotype change creates a difference in some detectable characteristic of the cells, by which the cells that express the marker may be distinguished from those that do not. Selection is preferable to screening in that it can provide a means for amplifying from the cell culture those cells that express a test polypeptide that is a receptor effector.

The reporter gene is coupled to the receptor signaling pathway so that expression of the reporter gene is dependent on activation of the receptor. This coupling may be achieved by operably linking the marker gene to a receptor-responsive promoter. The term "receptor-responsive promoter" indicates a promoter that is regulated by some product of the target receptor's signal transduction pathway.

Alternatively, the promoter may be one that is repressed by the receptor pathway, thereby preventing expression of a product that is deleterious to the cell. With a receptor repressed promoter, one screens for agonists by linking the promoter to a deleterious gene, and for antagonists, by linking it to a beneficial gene. Repression may be achieved by operably linking a receptor-induced promoter to a gene encoding mRNA which is antisense to at least a portion of the mRNA encoded by the marker gene (whether in the coding or flanking regions), so as to inhibit translation of that mRNA. Repression may also be obtained by linking a receptor-induced promoter to a gene encoding a DNA binding repressor protein, and incorporating a suitable operator site into the promoter or other suitable region of the marker gene.

Exemplary Uses

This section describes a variety of uses for the methods and arrays of the present invention. It is noted that the uses are divided into different categories for purposes of convenience and that such categories do not imply any limitation on the purposes for which the invention may be used.

Target identification and analysis

The binding partners for molecules such as drugs, hormones, interleukins, or secreted proteins can be identified by incubating the compounds of interest with an array that overexpresses potential targets within each array feature or combinations of potential targets within each cell of an array feature. Binding could be detected by methods such as SPR, SPA, TRF, or autoradiography. In addition, the binding partners for cells could be identified by incubating the cell of interest with arrays or

color-encoded beads. For instance, migratory or free-floating test cells could be incubated with an array, allowed to migrate or bind, and then the binding or migration detected by standard methods, e.g. expressing GFP or other markers in the test cells. Alternatively, the test cells could be mixed with a collection of color-encoded beads, each expressing a distinct DNA construct with a unique color code, e.g. a unique ratio of red to green dyes. Binding could then be detected by fluorescence activated cell sorting or other methods.

The array could also be used to identify the targets of an organism's immune response to cancer, an infectious or autoimmune disease, exposure to chemicals, or environmental changes. An array expressing target proteins could be incubated with sera from the organism. Binding of antibodies could be detected by labeling the sera or using the appropriate secondary antibody. The identified targets of the immune response could be used to design vaccines against tumors or infectious diseases, immunosuppressive drugs, anti-infective drugs or others.

In other embodiments, the present invention facilitates drug target discovery by permitting the identification of an endogenous gene the inhibition or activation of which may be of therapeutic value. The strategy relies, in part, on the ability of small gene fragments to encode dominant-acting synthetic genetic elements (SGEs), e.g., molecules that interfere with the function of genes from which they are derived (antagonists) or that are dominant constitutively active fragments (agonists) of such genes. SGEs that can be identified by the subject method include, but are not limited to, polypeptides, inhibitory antisense RNA molecules, ribozymes, nucleic acid decoys, and small peptides. For instance, a gene whose activity is inactivated by an identified SGE can itself be used as a target for drug development, e.g., to identify other agents, such as small molecules and natural extracts, which can also inhibit the function of the endogenous gene. Thus, another aspect of the present invention provides drug screening assays for detecting agonists or antagonists, as appropriate, of a gene (or gene product thereof) that corresponds to a selected SGE. Likewise, the identification of an SGE that can inhibit a particular pathological phenotype will

indicate diagnostic assays that can assess loss-of-function or gain-of-function mutations, as appropriate, to the corresponding endogenous gene.

In other embodiments, the use of transcription arrays which give rise to dsRNA in the host cell can be used to assess the loss-of-function of a particular gene. "RNA interference," "post-transcriptional gene silencing," "quelling"— these different names describe similar effects that result from the overexpression or misexpression of transgenes, from the deliberate introduction of double-stranded RNA (e.g., short interfering RNA) into cells, or from the engineered expression of dsRNA such as shRNAs in cells. In certain organisms, the injection of double-stranded RNA into a cell can act systemically to cause the post-transcriptional depletion of the homologous endogenous RNA. RNA interference, commonly referred to as RNAi, offers a way of specifically and potently inactivating a cloned gene, and is proving a powerful tool for investigating gene function. RNAi operates by a number of different mechanisms to silence gene expression, including cleavage of mRNA transcribed from the gene, translational repression of the mRNA, etc. As is known in the art, the precise mechanism by which any particular dsRNA operates to effect post-transcriptional gene silencing may depend on certain structural features of the dsRNA

To illustrate, the subject method contemplates (a) constructing a cDNA or genomic transfection array including cDNA or genomic DNA in an orientation relative to a promoter(s) capable of initiating transcription of the cDNA or genomic DNA to double stranded RNA; (b) introducing the transfection array into cells by the subject method; (c) identifying and isolating cells in which a member of the transfection array confers a particular phenotype; and (d) identifying the gene sequence from the library which gave rise to the dsRNA construct responsible for conferring the phenotype.

In mammalian cells, short dsRNA molecules known as short interfering RNAs (siRNAs) are widely used to silence gene expression via RNAi. These molecules are typically introduced into cells using a variety of methods such as lipid based transfection. As an alternative to introducing siRNAs into cells from the exterior,

viruses carrying expression cassettes that encode short hairpin RNAs (shRNAs) can be used to generate gene-specific siRNAs within cells. This approach can achieve stable and highly effective gene suppression in a variety of mammalian cell types (see, e.g., Brummelkamp et al., *Science*, 296(5567):550-3, 2002, Rubinson et al., *Nat. Genet.*, 33: 401-406, 2003. Paddison et al., *Methods Mol Biol.*, 265:85-100, 2004. See also U.S.S.N. 10/324,184 and U.S.S.N. 10/350,798). Suitable expression cassettes may contain an RNA polymerase III promoter such as the U6 or H1 promoter operatively linked to a template from which an RNA containing sense and antisense sequences corresponding to a target gene can be transcribed and typically followed by a terminator. Alternately cassettes containing RNA polymerase I or II promoters can be used. The complementary sense and antisense regions are separated by a region that forms a loop when the sense and antisense portions hybridize to each other to form a hairpin structure. One of ordinary skill in the art will appreciate that shRNAs typically contain a loop (typically between 6-15 nucleotides long) and a double-stranded (duplex) stem, which is typically between 15-50 nucleotides in length, more typically between 15-30 nucleotides in length, e.g., 19-23 nucleotides in length. One of skill in the art will also appreciate that the duplex stem may contain one or more mismatches or unpaired nucleotides.

In another embodiment, the viral vector encodes an RNA molecule that is a precursor of a microRNA-like RNA. MicroRNAs (miRNAs) are a form of endogenous single-stranded RNA which is typically 20-25 nucleotides long and regulates the expression of target gene(s) to which it is partially complementary. miRNAs are derived from larger precursors that include the miRNA sequence (which corresponds to the sequence of a target gene) and an approximate reverse complement. The RNA forms a double-stranded structure (which typically includes several mismatches or unpaired nucleotide bulges). This precursor structure is processed within the cell to generate an active silencing species. See, e.g., Bartel, D., *Cell*, 116(2):281-97, 2004. As used herein, a microRNA-like RNA is a molecule that is identical to or closely resembles an endogenous miRNA in sequence and overall structure. However, rather than being naturally expressed by the cell the miRNA-like

molecule is introduced by transfection, or its intracellularly expression is achieved by introducing a viral vector (or other vector) that encodes the RNA or a precursor of the RNA into the cell, e.g., using a viral vector array of the present invention.

One of skill in the art will appreciate that any of a variety of short double-stranded RNA molecules (e.g., shRNAs, microRNA-like RNA precursors), may undergo post-transcriptional processing in the cell to generate an active silencing species. As used herein, a dsRNA synthesized in a cell and corresponding to a target gene is said to silence the target gene notwithstanding the fact that it may need to undergo additional processing in the cell to form the active silencing species.

The present invention contemplates a method of post-transcriptionally silencing a gene in cells by using an array of the invention comprising a plurality of viral vectors to introduce into the cells a viral vector that encodes a dsRNA that post-transcriptionally silences expression of a the gene. In one embodiment the method comprises (a) contacting eukaryotic cells with a plurality of viral vectors affixed to a surface in a discrete, defined location, each of said viral vectors comprising a target nucleic acid molecule that encodes a dsRNA molecule capable of post-transcriptionally silencing the gene, under appropriate conditions for entry of the viral vectors into said eukaryotic cells; whereby said viral vectors are introduced into said eukaryotic cells in the location in which said viral vectors were affixed; and (b) maintaining the eukaryotic cells under conditions in which said target nucleic acid molecule is expressed to produce a dsRNA molecule, whereby said gene is post-transcriptionally silenced. The invention contemplates using the viral vector arrays to silence large numbers of genes in parallel (e.g., thousands of different genes). To this end, the method employs a viral vector array comprising a plurality of features, wherein the viral vectors of each feature encode a dsRNA molecule capable of silencing a gene. The dsRNA molecules of each feature may silence a different gene, or some of the dsRNA molecules in different features may silence the same gene. Collectively, the viral vectors of the array encode dsRNA molecules capable of silencing a plurality of different genes, so that when cells are plated on the array, the

resulting transfected cell array contains features in which different genes are silenced in the cells of each feature.

The subject method contemplates (a) constructing or providing a transfection array including viral vectors that contain a nucleic acid in an orientation relative to a promoter(s) capable of initiating transcription of the nucleic acid to double stranded RNA (e.g., shRNA, precursors of miRNA-like RNA, etc.); (b) introducing the transfection array into cells by the subject method; (c) identifying and isolating cells in which a member of the transfection array confers a particular phenotype; and (d) identifying the gene sequence which gave rise to the dsRNA construct responsible for conferring the phenotype. The dsRNAs may be members of a library of nucleic acid molecules. A variety of libraries of interest are envisioned. The library may include dsRNAs capable of silencing a substantial fraction (e.g., at least 10%, 20%, 30%-100%) of the genes in a mammalian genome). Specific isoforms, alleles, polymorphic variants, or splice variants may be selectively silenced. In another embodiment the library may include dsRNAs capable of silencing genes that encode proteins of a particular functional category or that fall into a particular family based on sequence or possession of particular sequence motifs. For example, the library may include dsRNAs that silence kinases, phosphatases, DNA repair proteins, deubiquitinating enzymes, or any functional category of interest. The library may include dsRNAs that silence all or substantially all (e.g., at least 90-95%) of the known or predicted genes of a particular functional class.

The subject method further contemplates (a) constructing or providing a transfection array including viral vectors that contain a nucleic acid in an orientation relative to a promoter(s) capable of initiating transcription of the nucleic acid to double stranded RNA (e.g., shRNA, precursors of miRNA, etc.); (b) introducing the transfection array into cells by the subject method; (c) analyzing the phenotypes resulting from silencing each of a plurality of different genes. Optionally the method includes identifying and isolating cells in which a member of the transfection array confers a particular phenotype; and identifying the nucleic acid responsible for conferring the phenotype, thereby identifying the gene that was silenced. In some

embodiments of the invention, the method includes contacting the cells with a compound, e.g., a test agent, that has a particular effect on cells and identifying a nucleic acid that prevents or inhibits the effect. In one embodiment, the nucleic acid encodes a dsRNA that silences expression of a gene, and the methods identifies the gene (or its encoded protein) as being a target of the compound. The invention thus finds use in drug discovery use to identify targets of known drugs or compounds under investigation. Once having identified a target, conventional compound screening methods may be employed to identify additional compounds that interact with (e.g., activate or inhibit) the target. Alternately, identifying unknown targets of known drugs or compounds under investigation may provide insight into causes of side effects.

Target validation

The expression pattern of potential genes of interest could be tested by constructing an array where each spot contains a construct fusing regulatory sequences from the genes of interest with a reporter gene. The regulatory sequences could be involved in transcription, RNA processing or translation. The reporter gene could encode, for example, GFP, beta galactosidase, luciferase, or beta lactamase. The expression of the genes of interest could be tested by incubating the array with different combinations of conditions and cell lines and then assaying for the activity of the reporter gene. Genes with the appropriate expression patterns could then be studied further as potential drug targets.

The function of the gene of interest could be tested by making arrays where DNA constructs modify the function of the gene of interest and assaying the phenotype. These modifications could be derived from methods such as overexpression, knockout constructs, dominant negative mutants, antisense RNA, ribozyme RNA or others. The resulting phenotypic change could be assayed under different environmental conditions, genetic backgrounds and cell types. For instance, genes which activate or inhibit a pathway could be identified by examining the phenotype of cells on an array where each feature overexpresses or underexpresses a

gene of interest. Genes with the appropriate phenotypes could then be studied further as potential drug targets.

The function of a gene of interest could also be inferred by identifying the binding partners for a protein of interest. For instance, an array expressing proteins of interest could be tested for DNA binding, RNA binding, protein binding, nucleotide binding or other functions by incubating the array with the appropriately labeled molecule and/or detection system. Different classes of proteins, e.g., DNA-binding proteins, could be identified and the sequences examined for the discovery of novel binding motifs. Alternatively, a two hybrid or three hybrid system could be used to identify potential protein, RNA, or other classes of binding partners in vivo. For instance, the gene of interest could be cloned into the appropriate "bait" vector and stably transfected in a cell line with the appropriate reporter construct. The interaction of the gene of interest with other potential partners could be tested by using this cell line in an array of constructs where test proteins are cloned into the appropriate "test" vector. Alternatively, an array of affinity tagged constructs (e.g., 6xHis, epitopes, avidin) could be transferred to an affinity membrane, e.g., (Ni-NTA, anti-epitope antibody, biotinylated). Associated proteins could be detected and identified by mass spec or other methods. Proteins with the appropriate binding partners could then be further investigated as potential targets.

The function of a gene of interest could also be inferred by identifying its post-translational modifications. An array expressing proteins of interest could be tested for phosphorylation, sulfation, ubiquitination, glycosylation or other post-translational modifications by incubation with the appropriate labeling or detection reagent such as radiolabeled precursors, anti-phosphoamino acid antibodies, anti-ubiquitin, lectins or other specific detection reagents. Alternatively, post-translational modifications could be detected by transferring the array to an affinity membrane and then using mass spectrometry.

Subcellular localization of a protein could be investigated by making an array where each feature contains a DNA construct with the protein of interest fused to an epitope tag, GFP or other marker. After transfection and cell growth,

immunofluorescence could be performed with a microscope, high resolution scanner or other detection method to determine whether the proteins of interest localized to the nucleus, cytoplasm, membrane, extracellular or other compartments. Proteins with the appropriate subcellular localization could then be further investigated as potential targets.

Screening

Large molecule therapeutics (such as proteins, nucleic acids, sugars) could be identified by making an array of the appropriate constructs and screening for the desired phenotype. For instance, a screen for secreted proteins could involve an array where cells expressing secreted proteins are mixed with tester cells with the potential for an assayable response to the secreted proteins. After transfection and growth, the response of the tester cells could be measured to identify features producing secreted proteins with the desired effect.

Multiplexed screening could be performed by making arrays on the bottom of each well of a microtiter dish. The binding of molecules to an array of 100 or more potential targets in the bottom of each well may be assessed. These targets could be pharmacogenomic variants, families of proteins, or other collections of proteins. The binding could then be assayed by a scanner, plate reader or other instrument, (e.g., Cellomics ARRAYSCAN II).

Lead optimization

Potential drug candidates may be evaluated for selectivity by incubating the candidate with the appropriate array of potential targets. The arrays could be the entire set of genes in the genome(s) of interest or focused subsets, e.g. GPCRs, ion channels, enzymes, nuclear hormone receptors. The relative binding of the drug candidate to the known target and other potential targets could be determined. Candidates with a high degree of non-selective binding could be abandoned or modified to reduce non-selective binding before additional testing such as ADME ortoxicology other tests. Potential drug candidates could be evaluated for toxicity by

incubating the candidate with the appropriate array of targets, such as cytochrome P-450s, including pharmacogenomic variants or other variations.

Selectivity tests could also be performed on the metabolites of a drug candidate. For instance, a radiolabeled drug could be reacted with the appropriate biotransformation agent, such as a liver extract, tissue culture system, or living organism such as a rodent or dog. The radiolabeled metabolites could then be extracted and purified and tested for binding with the array. Metabolites with binding activity could then be characterized further by standard methods.

Optimization of plasmids

In still another embodiment, the subject method can be used to optimize an expression system for a particular cell type. Briefly, the transfection array can be a collection of various permutations of a vector system. For instance, the vector library can test various combinations and permutations of promoter and enhance sequences, replication origins, and other components that could effect the level of expression of a protein or the stability of the cell line for the plasmid.

Kits

Any of the viral vector arrays of the present invention may be provided in the form of a kit. The kit may include any of a number of components in addition to one or more viral vector arrays. Such components may include cells, media, detection reagents (e.g., for detecting expression of a nucleic acid introduced into cells using the viral vector array), and instructions for use.

Exemplary Embodiment

The present invention is illustrated by the following example, which is not intended to be limiting in any way.

Example 1: Development of LCBM

We created LCBMs by individually infecting small distinct cell clusters with one or more different lentiviruses, all surrounded by a confluent monolayer of uninfected cells on a glass slide. To do this, we printed nanoliters of concentrated lentiviral solutions on a microscope slide. The printed solution contained virus prepared using a novel technique adapted for using lentivirus in a microarray format. After drying, slides containing arrays of lentiviral spots were placed in a petri dish and covered with a confluent layer of adherent mammalian cells, creating an LCBM. After a specified incubation, typically 72 or 96 hours, the LCBM was fixed and processed by conventional cell based assay methods, such as immunofluorescence. Finally, the clusters of cells growing on each spot were photographed using automated microscopy and analyzed for a phenotype of interest.

To demonstrate dense and addressable infections with LCBM technology, we printed a large grid of 300 spots. Each spot was 200-300 μm in diameter and contained either lentivirus that expressed GFP or a shRNA targeting human Lamin A/C. Subsequently, we seeded the array with $2x10^6$ HeLa cells and allowed the LCBM to incubate for 72 hours. After this incubation, the arrays were fixed and then stained with an anti-LaminA/C-antibody and Hoechst 33342. Against an unaffected lawn of cells, defined cell clusters either express GFP or are deficient in Lamin A/C. The pattern of infected spots was consistent and confined to the pattern of lentivirus printed on the array demonstrating local infection of the targeted cells (Fig. 1A). Under high magnification, each cluster consisted of about 150 cells infected by the virus printed in that spot. At this feature density, it is possible to print 5000 different viruses on a single slide, allowing for 5000 distinct infections.

We printed the two types of virus in an alternating pattern to assess whether cells in one feature will be infected by lentivirus from proximal features. Cells growing on features with lentivirus encoding GFP express high levels of green fluoresence, unlike uninfected or shLamin A/C expressing cells. The cells on these features expressed Lamin A/C at levels similar to those of uninfected cells (Fig. 1B). Similarly, the cells on features with shLamin A/C lentivirus showed a 5-fold

reduction in Lamin A/C expression when compared to uninfected cells or cells infected with GFP-overexpressing virus. In addition, the number of knocked down cells per spot was increased 6-fold on features with shLamin A/C lentivirus when compared to uninfected or GFP infected cells. These results indicate that features containing printed lentivirus infect small clusters of cells in a local, viral specific manner.

Having demonstrated the viability of LCBM, we tested the effect of printing different amounts of lentivirus in each spot. The typical titer of a printed viral solution is $1x10^9$ IFU/ml. In each spot we deposited ~3.9 nl of solution or 3900 IFU/spot. Each feature had ~150 cells and roughly 26 IFU/cell. We printed two-fold, serial dilutions of a GFP-overexpressing lentiviral solution (initial titer of $2x10^9$ IFU/ml) on multiple arrays and quantified the subsequent level of GFP expression in the cells growing on those features. As seen with conventional lentiviral infection, the amount of lentivirus deposited in each spot was proportional to the amount of infection seen in cells growing on the corresponding feature (Fig. 2A).

VSV-G pseudotyped viruses are used to readily infect many different mammalian cell types. We printed four identical, 100-spot arrays of GFP-overexpressing lentivirus and covered each with a different cell-type (Fig. 2B). Although the average GFP expression was highest in A549 and DU145 cells, all four cell-types tested had levels of GFP expression that were well above the background of uninfected cells. This result demonstrates that LCBMs can be used with commonly used adherent, cultured cell lines, both cancerous and non-cancerous.

To demonstrate the compatibility of LCBMs with primary mammalian cells, we infected both human BJ fibroblasts and mouse dendritic cells with lentivirus encoding GFP (Fig. 2C). The arrays produced grids of GFP-expressing cell clusters in the printed pattern, showing that LCBMs are compatible with non-dividing and dividing primary cells.

We have demonstrated that LCBMs are compatible with overexpression and shRNA-mediated knockdown by lentivirus. To ensure that functionality was not specific to the lentiviral backbones used and the gene overexpressed, we obtained

several different lentiviruses. We printed small grids of virus that upon infecting cells should overexpress GFP from the PGK promoter, RFP from the Ubiq-C promoter, or thy1.1 from the Ubiq-C promoter (Fig. 3A). The backbones for these overexpression viruses were LKO.1, Lentilox 3.7, and LKO.2 respectively. All three viruses infected localized cell clusters and produced overexpression of only the expected protein, suggesting the LCBMs are compatible with a wide variety of promoters for gene overexpression and distinct viral backbones. At high magnification, both the RFP and GFP are expressed throughout infected cells. In contrast, the Thy1.1 is co-localized with the cell membrane, consistent with its known subcellular localization. This indicates that infected cells can accurately express biological relevant proteins in the microarray format. Our results suggest that LCBMs preserve the desired features of VSV-G lentiviruses such as high efficiency infection of mammalian cells, ability to infect primary non-dividing cells, and ability to deliver shRNAs.

Stability and durability of arrays for reverse transfection is desirable. To determine whether the lentilviral arrays were stable and durable, we printed multiple copies of microarrays, and then stored the replicate arrays at -80°C for 15 months. After storage, the slides were thawed, seeded, and analyzed in the same way as the freshly assayed slide. The stored arrays result in infection and expression that closely resembles that of the freshly printed arrays (Fig. 3A). The similarity of expression suggests that the viral microarrays can be stored for extended periods of time before infection.

Finally, we demonstrated that infection of a small cluster of cells in the array format could produce a complex phenotype. To accomplish this, we printed a lentivirus expressing an shRNA targeted against human mTOR, a central regulator of growth (cell size) and proliferation (cell number) in mammalian cells. We assayed mTOR activity through the phosphorylation of its downstream effector, the ribosomal protein S6 (Fig. 3B). The cells growing on spots containing shTOR lentivirus had a dramatic reduction in S6 phosphorylation, while the level of S6 phosphorylation in cells growing on the shLamin and GFP-overexpressing spots was unaffected. Furthermore, spots infected with shTOR lentivirus had fewer and smaller cells when

compared to controls. These phenotypes are concordant with mTOR knockdown by lentiviral shRNA performed by conventional means.

Materials and Methods

Reagent preparation

The GFP-expressing virus used herein is PRRL PGK GFP (Stewart et al. (2003) RNA 9:493-501). The shRNA lamin virus used was LKO.1 puro shRNA Lamin A/C (F oligo: CCGGCTGGACTTCCAGAAGAACATCCTCGAGTTTTTGG ATGTTCTTCTGGAAGTCCAG, B oligo: AATTCAAAAACTGGACTTCCAGAA GAACATCCTCGAGGATGTTCTTCTGGAAGTCCAG). The Thy1.1 overexpressing virus used was LKO.3 Thy1. The shRNA mTOR virus used was LKO.1 puro shRNA mTOR (F oligo: CCGGTCAGCGTCCCTACCTTCTTCTCTT CCTgTCAAGAAGAAGGTAGGGACGCTGATTTTTG, B oligo: AATTCAAAAA TCAGCGTCCCTACCTTCTTCTTGAcAGGAAGAGAAGAAGGTAGGGACGCT GAA (Kawasaki et al. (2003) Nucleic Acids Res. 31:700-707). To make virus, $6x10^6$ 293T cells were seeded in a 15 cm petri dish and incubated for 48 hrs at 37°C and 5% CO_2. Cells were then transfected with Fugene 6 (9 μg backbone, 6 μg Delta VPR 8.9 and 3 μg VSVG (Naldini et al. (1996) Science 272: 263-267) and incubated for an additional 24 hrs. After 24 hrs, the virus-containing supernatant was removed and filtered through a 0.45 μm cellulose acetate filter. To concentrate the virus, the filtered supernatant was ultracentrifuged at 23000 rpm for 1.5 hrs. To get more virus, fresh media was added to the 15 cm dish of 293T cells and was harvested after an additional 24 hr incubation. After centrifugation, pellets containing virus were re-suspended in a solution containing 0.4 M HEPES (pH to 7.4 with KOH), 12.5mg/ml trehalose (Calbiochem; Cat#625625), 6 μg/ml protamine sulfate (Sigma; Cat# P4020) and 1.23 M KCL (Sigma, Cat# 3911). Re-suspended virus was frozen at -80°C until use.

Microarray printing

A microarraying robot (Pixsys 5500; Genomic Solutions) with SMP10 pins (Arrayit SMP10; Telechem Inc.) was used to deposit viral solutions on gamma amino propyl silane-coated glass slides (Corning; UltraGAPS; cat#40015). UltraGAPS slides were used because the surface allowed for the printing of small, well-formed spots and was compatible with cell culture. Prior to printing the microarray, 20 μL of viral solution was loaded into a round-bottom, polypropylene 384-well microtiter plate and centrifuged at 1000 rpm for 30 seconds. All printing was performed at 23°C with 55% humidity. Virus was printed on UltraGAPS slides using SMP10 pins with a pins-down time of 350ms. Before each spot was printed on an array, 12 pre-printing spots were made on a pre-printing slide. The titer of a typical virus solution concentrated 600x is roughly $1x10^9$ IFU/ml. An SMP10 spot deposits ~ 3.9nl of solution, making a spot with a diameter between 200 and 300 μm. Based on the titer, we estimate there are 3900 IFU/spot. On a typical array seeded with HeLa cells, one spot has around 150 cells and 26 IFU/cell. After printing, slides were either used in an experiment or stored at -80°C. For the experiments in Figs. 1, 2A, 2B and 3, the spot center-to-center distance was 500 μm. In Fig. 2C, the spot center-to-center distance was 1.5 mm for the arrays seeded with BJ cells and 500 μm for arrays seeded with mouse dendritic cells.

Viral microarray assay

If previously printed and stored, slides containing arrays were desiccated at RT for 1 hr prior to use. All slides were then 'blocked' in DME (Dulbecco's Modified Eagle's Medium) /10% inactivated fetal calf serum (IFCS) for 30 min in a 15 cm petri dish at room temperature. During the blocking, the slides were propped such that the printed area on the slide was face down and not in contact with anything other than media. After blocking, slides were placed array side up in a 100 x 100 x 10 mm square tissue culture dish. Actively growing mammalian cells in 25 mL of culture medium (A549 cells: DMEM with 10% fetal bovine serum (FBS), 50 units mL^{-1} penicillin and 50 μg mL^{-1} streptomycin; HeLa, 293T, DU145, BJ cells: DMEM

with 10% IFCS, 50 units mL^{-1} penicillin and 50 $\mu g\ ml^{-1}$ streptomycin) were seeded on arrays. Finally, arrays with cells were incubated at 37°C in 5% CO_2 to grow. When performing longer assays, fewer cells were seeded to achieve similar final cell densities. For Fig. 2C, we seeded arrays with two million BJ cells for 72 hrs or $1.6x10^6$ dendritic cells for 92 hrs. For all other figures, we seeded 2 $x10^6$ cells for 72 hrs. After growth for the specified amount of time, slides were removed from culture medium and fixed using 3.7% formaldehyde and 4% sucrose in PBS for 20 minutes at room temperature.

Immunofluorescence

After fixation, cells on arrays in Figs. 1 and 3 were probed with primary and secondary antibodies as previously described (Sarbassov dos et al. (2004). Curr. Biol. 14: 1296-1302). The cells on arrays in Fig. 1 were permeabilized with 0.2% Triton-X for 20 minutes, stained with an anti-Lamin A/C (636) primary mouse moloclonal antibody (Santa Cruz Biotechnology, Inc.; sc-7292; 1:500), anti-mouse-cy3 secondary antibody (Jackson Laboratories; 1:1000), and Hoechst 33342 (Molecular Probes; H3570; 1:10000). Cells on arrays in Fig. 2 were stained with Hoechst (1:10000). Cells on arrays in Fig. 3A were stained with an anti-thy1-biotin antibody (PharMingen;1:500), streptavidin-cy3 secondary (PharMingen; 1:1000), and Hoechst 33342 (1:10000). Cells on arrays in Fig. 3B were stained with anti-p-S6 antibody (ser240/4; Cell Signaling; 1:500), anti-rabbit cy3 (Jackson Laboratories; 1:1000), and Hoechst (1:10000).

Image capture and analysis

We captured images of all slides, both light and fluorescent microscopy using an Axiovert 200 (Carl Ziess, Inc.) and custom software designed on the KS400 software platform. Photographs of large arrays were captured at 50x magnification, images of single spots were captured at 100x magnification, and the image of the spot containing the shRNA-mTOR virus was captured at 400x magnification. In all graphs, the error bars are the standard deviation.

For the GFP levels in Fig. 1, we measured the mean grayscale values for each spot (a circular ROI with a radius of 100 μm) and the 'background' grayscale values of uninfected cells on the slide. We then subtracted the background from grayscale values for each spot and averaged the mean grayscale values for all spots with the same type of virus. To evaluate the levels of lamin expression, we measured the mean grayscale intensity of lamin in each cell on each spot. Cells on spots with mean grayscale values below 80% of the grayscale values of uninfected cells were considered knocked down (KD) for lamin. The number of cells knocked down is expressed as a percentage of the total cells on each spot. In addition, we report the average grayscale value of the KD cells on spots with shLamin virus in relation to cells on spots containing GFP-overexpressing virus and uninfected cell on other parts of the slide. For Fig. 2, we measured the mean grayscale values of GFP as described above.

While this invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the scope of the invention encompassed by the appended claims.

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. The scope of the present invention is not intended to be limited to the above Description, but rather is as set forth in the appended claims. In the claims articles such as "a,", "an" and "the" may mean one or more than one unless indicated to the contrary or otherwise evident from the context. Claims or descriptions that include "or" between one or more members of a group are considered satisfied if one, more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process unless indicated to the contrary or otherwise evident from the context. The invention includes embodiments in which exactly one member of the group is present in, employed in, or otherwise relevant to a given product or process. The invention also includes embodiments in which more than one, or all of the group members are present in,

employed in, or otherwise relevant to a given product or process. Furthermore, it is to be understood that the invention encompasses variations, combinations, and permutations in which one or more elements, components, descriptive terms, etc., from one or more of the listed claims is introduced into another claim. In particular, any claim that is dependent on another claim can be modified to include one or more elements, components, etc., present in any other claim that is dependent on the same base claim. Furthermore, where the claims recite a composition, it is to be understood that methods of making the composition according to any of the methods disclosed herein or other suitable methods, and methods of using the composition for any of the purposes disclosed herein or other suitable purposes are provided, unless otherwise indicated or unless it would be evident to one of ordinary skill in the art that a contradiction or inconsistency would arise. Likewise, where the claims recite a method, it is to be understood that the method may employ any suitable composition described herein.

It should it be understood that, in general, where embodiments or aspects of the invention is/are referred to as comprising particular elements, components, etc., the invention provides embodiments or aspect that consist, or consist essentially of, such elements, components, etc. For purposes of simplicity these have not been specifically set forth herein.

Where ranges are given, endpoints are included. Furthermore, it is to be understood that unless otherwise indicated or otherwise evident from the context and understanding of one of ordinary skill in the art, values that are expressed as ranges can assume any specific value or subrange within the stated ranges in different embodiments of the invention, to the tenth of the unit of the lower limit of the range, unless the context clearly dictates otherwise.

What is claimed is:

Claims

1. A method of introducing a target nucleic acid molecule into cells, said method comprising the steps of:

(a) affixing a plurality of VSV-G pseudotyped viral vectors onto a surface, said VSV-G pseudotyped viral vectors comprising said target nucleic acid molecule; and

(b) contacting eukaryotic cells with the affixed VSV-G pseudotyped viral vectors of step (a) under appropriate conditions for entry of the VSV-G pseudotyped viral vectors into said eukaryotic cells, whereby said target nucleic acid molecule is introduced into said cells.

2. The method of claim 1, wherein said viral vectors are lentiviral vectors or Moloney murine leukemia virus-derived retroviral vectors.

3. The method of claim 1, wherein said target nucleic acid molecule is expressed in said cells.

4. The method of claim 1, wherein said target nucleic acid molecule comprises nucleic acid molecules encoding double-stranded RNA molecules.

5. The method of claim 1, wherein said cells are mammalian cells.

6. A method of identifying a nucleic acid molecule capable of post-transcriptional gene silencing, said method comprising the steps of:

(a) affixing a plurality of VSV-G pseudotyped viral vectors onto a surface in discrete, defined locations, each of said viral vectors comprising a target nucleic acid molecule;

(b) contacting eukaryotic cells with the affixed VSV-G pseudotyped viral vectors of step (a) under appropriate conditions for entry of the nucleic acid molecules into said eukaryotic cells; whereby said VSV-G pseudotyped viral vectors are

introduced into said eukaryotic cells in the location in which said VSV-G pseudotyped viral vectors were affixed; and

(c) determining the ability of said target nucleic acid molecules in said VSV-G pseudotyped viral vectors to post-transcriptionally silence expression of a gene in said cells,

wherein post-transcriptional gene silencing at a discrete, defined location identifies the target nucleic acid molecules affixed at said location as being capable of post-transcriptional gene silencing.

7. The method of claim 6, wherein said target nucleic acid molecules of step (a) encode double-stranded RNA molecules.

8. The method of claim 6, wherein said cells are mammalian cells.

9. The method of claim 6, wherein said affixed plurality of VSV-G pseudotyped viral vectors form an array of VSV-G pseudotyped viral vectors and wherein said cells into which said VSV-G pseudotyped viral vectors are introduced form an array of cells comprising said VSV-G pseudotyped viral vectors.

10. The method of claim 9, wherein the array comprises at least 300 different discrete, defined locations of known sequence composition.

11. The method of claim 10, wherein the array comprises at least 5000 different discrete, defined locations of known sequence composition.

12. The method of claim 11, wherein the array comprises up to 10,000 to 15,000 different discrete, defined locations of known sequence composition.

13. The method of claim 6, wherein each of the defined locations is 100-300 μm in diameter.

14. The method of claim 6, wherein each of the defined locations is 200-500 μm apart from its nearest neighboring location.

15. An array of VSV-G pseudotyped viral vectors comprising a surface having an array of at least 100 features per square centimeter, wherein each feature comprises one or more VSV-G pseudotyped viral vectors non-covalently affixed to the surface, wherein the one or more VSV-G pseudotyped viral vectors are capable of being transfected into a eukaryotic cell under appropriate conditions.

16. The array of claim 15, further comprising eukaryotic cells disposed on the surface and capable of being transfected by the one or more VSV-G pseudotyped viral vectors of at least one feature under appropriate conditions.

17. The array of claim 15, wherein at least one feature comprises at least two different nucleic acid molecules.

18. The array of claim 15, wherein the features have a density of at least 1000 different features per square centimeter.

19. The array of claim 18 wherein the features have a density of at least 10,000 different features per square centimeter.

20. The array of claim 19, wherein the features have a density of at least 100,000 different features per square centimeter.

21. The array of any of claims 15-20, wherein the features have a density of at most 1,000,000 different features per square centimeter.

22. The array of claim 15, wherein, in at least one feature, the one or more defined nucleic acid molecules encode a double-stranded RNA molecule.

23. A method of forming an array of transfected eukaryotic cells comprising:

(a) providing a surface having an array of features, wherein each feature comprises one or more VSV-G pseudotyped viral vectors non-covalently affixed to the surface;

(b) contacting the surface with eukaryotic cells; and

(c) transfecting the one or more VSV-G pseudotyped viral vectors into the eukaryotic cells to form the array of transfected eukaryotic cells.

24. The method of claim 23, wherein, in at least one feature, the one or more VSV-G pseudotyped viral vectors affixed to the surface are admixed with a carrier.

25. The method of claim 23, wherein one or more VSV-G pseudotyped viral vectors of at least one feature are expressed in the transfected eukaryotic cells.

26. The method of claim 23, wherein one or more VSV-G pseudotyped viral vectors of least one feature encode polypeptides that are expressed in the transfected eukaryotic cells.

27. The method of claim 23, wherein at least one of the features comprises at least two different VSV-G pseudotyped viral vectors.

28. The method of claim 23, wherein, in the array of step (a), the features have a density of at least 1000 different features per square centimeter.

29. The method of claim 28, wherein the features have a density of at least 10,000 different features per square centimeter.

30. The method of claim 29, wherein the features have a density of at least 100,000 different features per square centimeter.

31. The method of any of claims 23-30, wherein, in the array of step (a), the features have a density of at most 1,000,000 different features per square centimeter.

32. The method of claim 23, wherein, in at least one feature, one or more VSV-G pseudotyped viral vectors encode a double-stranded RNA molecule.

33. An array of VSV-G pseudotyped viral vectors comprising a surface having an array of at least 100 features per square centimeter, wherein each feature comprises one or more VSV-G pseudotyped viral vectors reversibly affixed to the surface, wherein the one or more VSV-G pseudotyped viral vectors are capable of being transfected into a eukaryotic cell under appropriate conditions.

34. The method of claim 33, wherein the features have a density of at least 1000 different features per square centimeter.

35. The method of claim 34, wherein the features have a density of at least 10,000 different features per square centimeter.

36. The method of claim 35, wherein the features have a density of at least 100,000 different features per square centimeter.

37. The method of any of claims 33-36, wherein the features have a density of at most 1,000,000 different features per square centimeter.

38. A method of forming an array of transfected eukaryotic cells comprising

(a) providing a surface having an array of features, wherein each feature comprises one or more VSV-G pseudotyped viral vectors reversibly affixed to the surface;

(b) contacting the surface with eukaryotic cells; and

(c) transfecting the one or more VSV-G pseudotyped viral vectors into the eukaryotic cells to form the array of transfected eukaryotic cells.

39. A method of post-transcriptionally silencing a gene, said method comprising the steps of:

(a) contacting eukaryotic cells with a plurality of viral vectors affixed to a surface in a discrete, defined location, each of said viral vectors comprising a target nucleic acid molecule that encodes a dsRNA molecule capable of post-transcriptionally silencing the gene, under appropriate conditions for entry of the viral vectors into said eukaryotic cells; whereby said viral vectors are introduced into said eukaryotic cells in the location in which said viral vectors were affixed; and

(b) maintaining the eukaryotic cells under conditions in which said target nucleic acid molecule is expressed to produce a dsRNA molecule, whereby said gene is post-transcriptionally silenced.

40. The method of claim 39, wherein said viral vectors are VSV-G pseudotyped viral vectors.

41. The method of claim 39, wherein said viral vectors are lentiviral or Moloney murine leukemia virus-derived retroviral vectors.

Figure 1

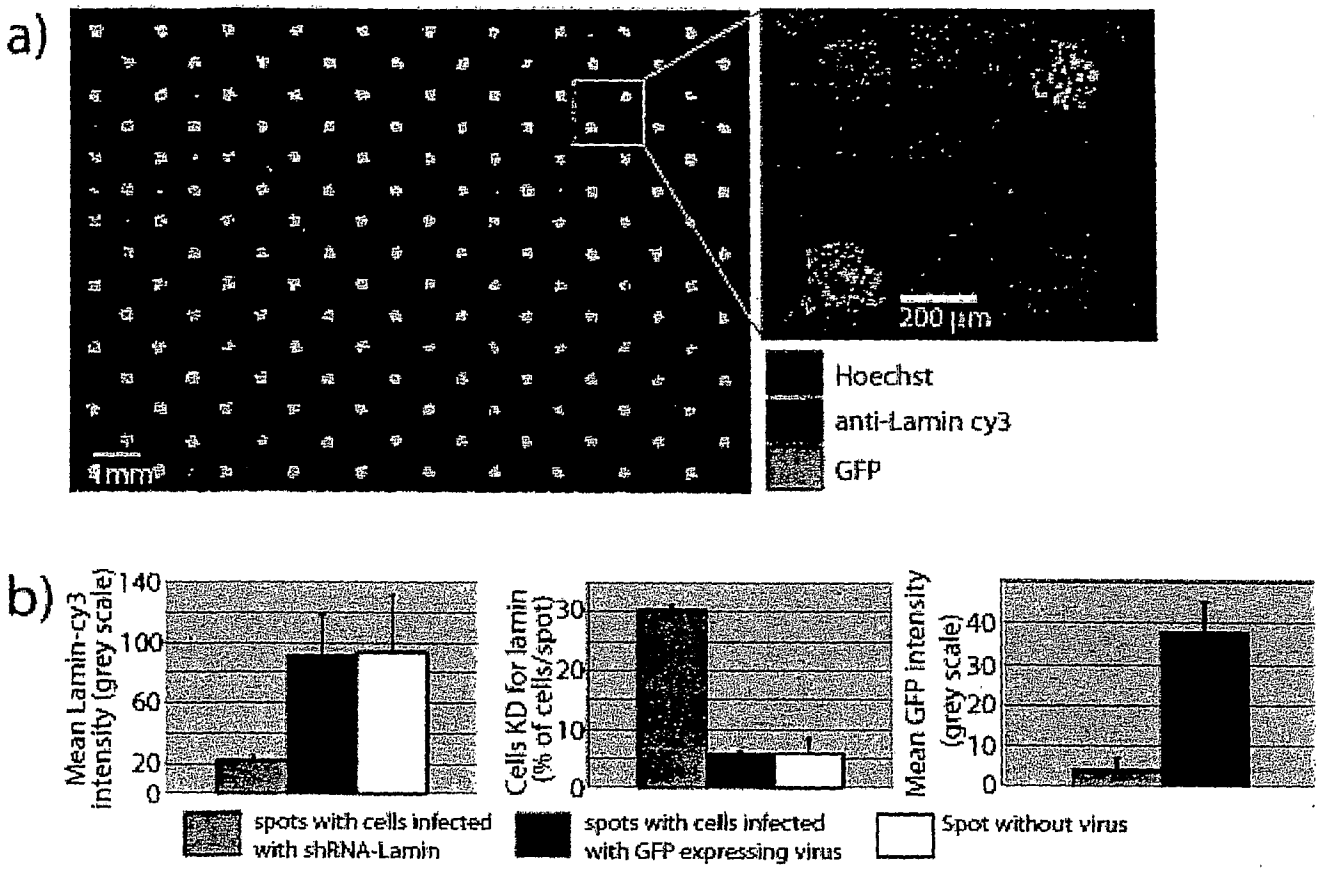


Figure 1: a) A large array of 300 spots containing either GFP expressing lentivirus or shRNA lamin lentivirus. Blue is Hoechst (nuclei), red is a anti-lamin-cy3, and green is GFP. b) Quantitation of fluorescent intensities on spots in the array. GFP was only seen on spots with cells infected with GFP lentivirus and lamin knockdown (KD) was only seen on spots infected with shRNA-lamin lentivirus.

Figure 2

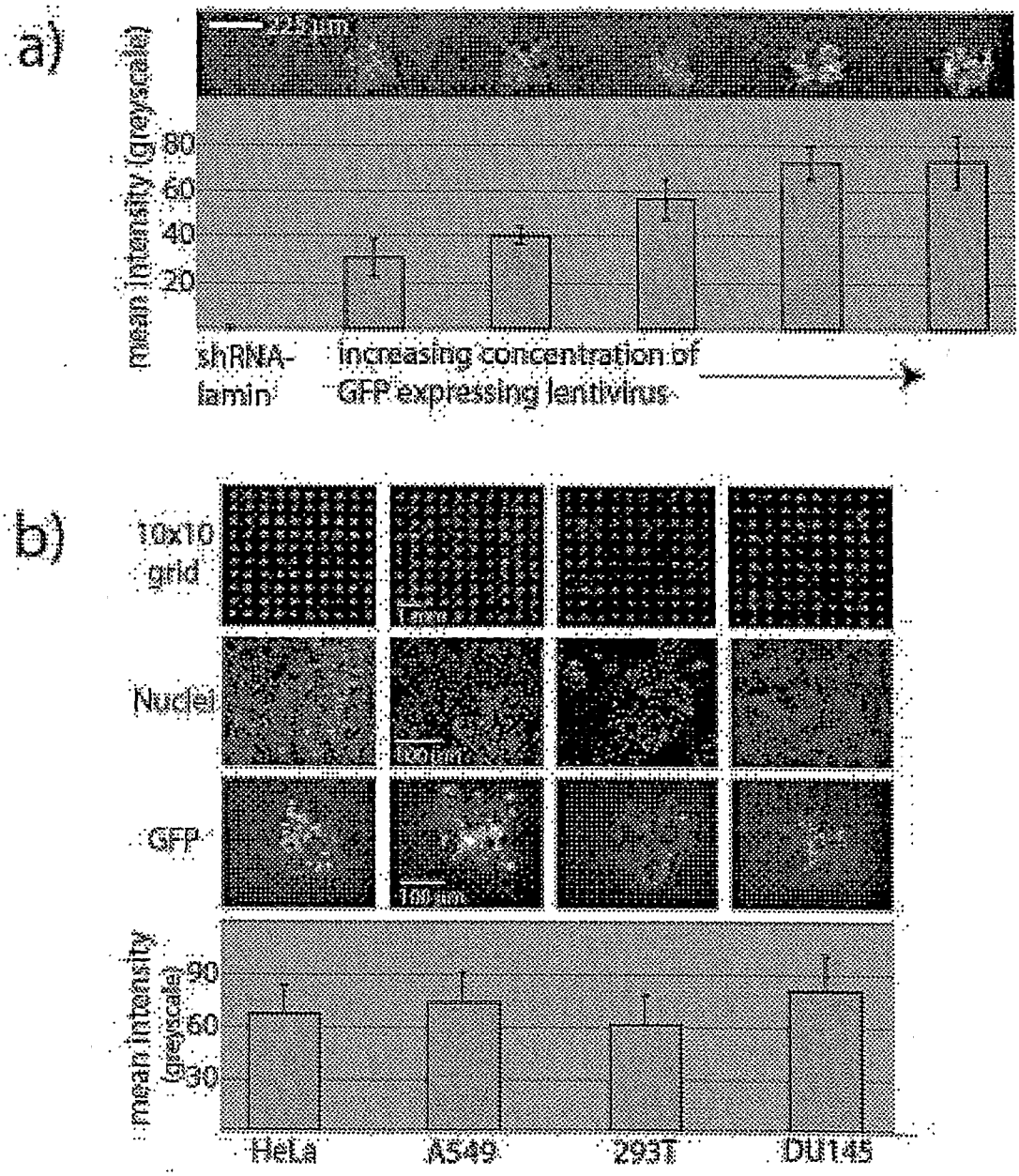
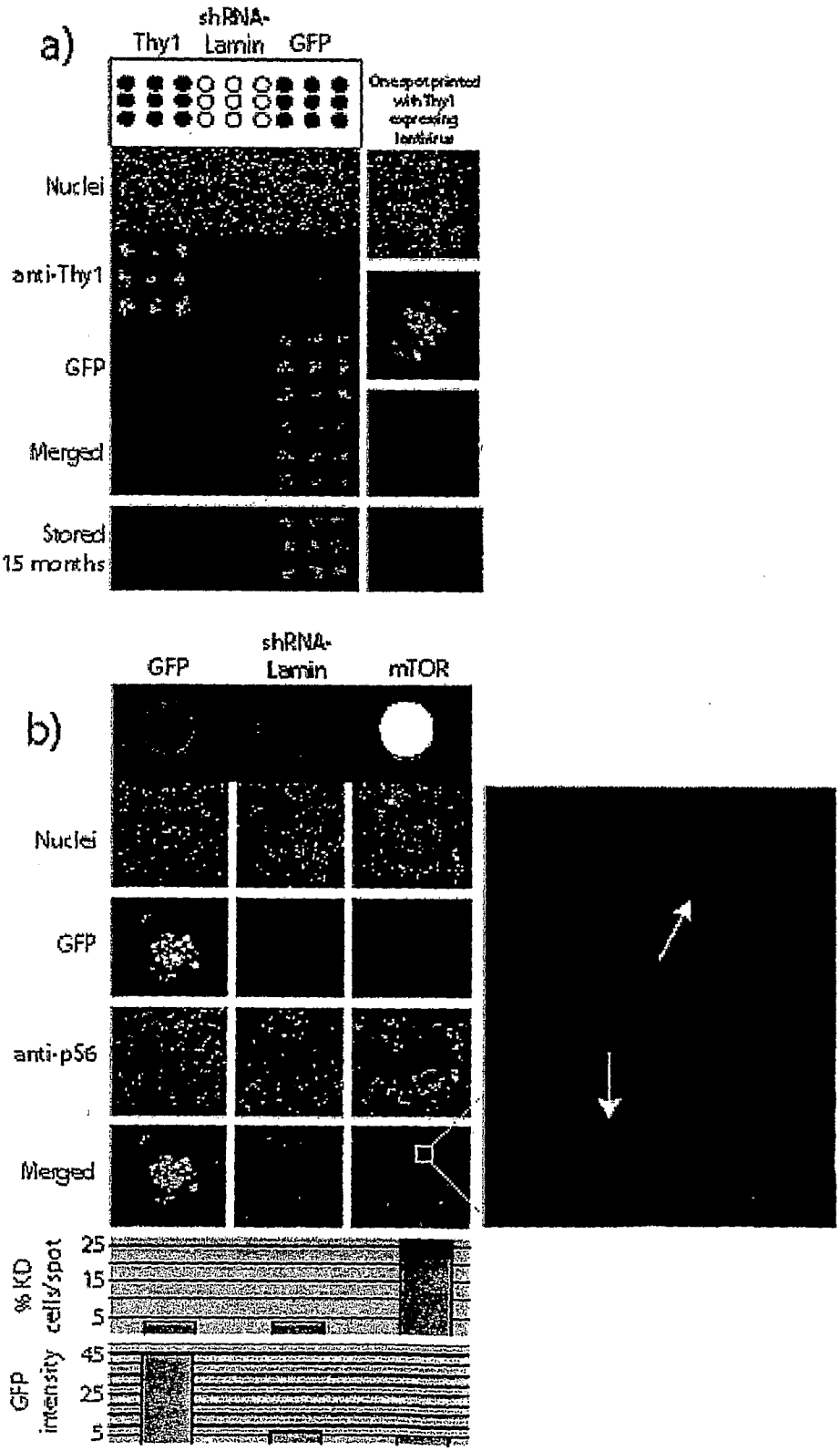

a)
mean intensity (greyscale)
80
60
40
20
shRNA-lamin
increasing concentration of GFP expressing lentivirus
b)
10x10 grid
Nuclei
GFP
mean intensity (greyscale)
90
60
30
HeLa
A549
293T
DU145
c)
BJ (primary fibroblasts)
Primary Mouse Dendritic cells
GFP levels in arrays printed with GFP virus
Nuclei
GFP
Nuclei
GFP
One spot

Figure 3


a)
shRNA-
Thy1
Lamin
GFP
One spot printed with Thy1 expressing lentivirus
Nuclei
anti-Thy1
GFP
Merged
Stored 15 months
b)
shRNA-
GFP
Lamin
mTOR
Nuclei
GFP
anti-p56
Merged
% KD
cells/spot
25
15
5
GFP intensity
45
25
5