

# Lessons from genetics: interpreting complex phenotypes in RNAi screens

Raphael Sacher<sup>1</sup>, Lilli Stergiou<sup>1</sup> and Lucas Pelkmans

Mammalian cell biology is witnessing a new era in which cellular processes are explained through dynamic networks of interacting cellular components. In this fast-paced field, where image-based RNAi screening is taking a central role, there is a strong need to improve ways to capture such interactions in space and time. Cell biologists traditionally depict these events by confining themselves to the level of a single cell, or to many population-averaged cells. Similarly, classical geneticists observe and interpret phenotypes in a single organism to delineate signaling processes, but have also described genetic phenomena in populations of organisms. The analogy in the two approaches inspired us to draw parallels with, and take lessons from concepts in classical genetics.

## Address

Institute of Molecular Systems Biology, ETH Zürich, Wolfgang-Pauli Street 16, 8093 Zürich, Switzerland

Corresponding author: Sacher, Raphael ([sacher@imsb.biol.ethz.ch](mailto:sacher@imsb.biol.ethz.ch)), Stergiou, Lilli ([stergiou@imsb.biol.ethz.ch](mailto:stergiou@imsb.biol.ethz.ch)) and Pelkmans, Lucas ([pelkmans@imsb.biol.ethz.ch](mailto:pelkmans@imsb.biol.ethz.ch))

<sup>1</sup>Authors contributed equally.

Current Opinion in Cell Biology 2008, 20:483–489

This review comes from a themed issue on  
Membranes and organelles  
Edited by Graça Raposo and Harald Stenmark

Available online 11th July 2008

0955-0674/\$ – see front matter

© 2008 Elsevier Ltd. All rights reserved.

DOI 10.1016/j.ceb.2008.06.002

## Introduction

Phenomena from genetics, such as quantitative traits, heterogeneity, pleiotropy, penetrance, and expressivity, are likely to be relevant for every cellular process, and will be visible in a population of isogenic tissue culture cells. For example, it is just as fundamental to implicate these phenomena in clathrin-mediated endocytosis, as it is for engulfment of apoptotic germ cell corpses in *Caenorhabditis elegans* [1–3]. By quantitatively studying large populations of cells whose components and regulatory circuits are known at the single cell level, we might better understand the mechanisms that, when disrupted, give rise to these genetic phenomena. This approach lies at the interface of genetics, molecular cell biology, and evolution, and will be a major focus of attention in the

developing discipline of systems cell biology. In this review, we will highlight some of the concepts of classical genetics and discuss them in the context of current challenges in image-based RNAi screening. We want to increase awareness among cell biologists by demonstrating these phenomena for clathrin-mediated endocytosis of transferrin in human cells.

## Endocytosis and RNAi

The process of endocytosis has a crucial role in the maintenance of cell physiology and homeostasis [4,5]. Endocytosis regulates a broad spectrum of signaling processes, ranging from cell migration and cell adhesion to anchorage-dependent cell growth, cell proliferation, and polarity of the cell [6–9]. The endocytic membrane system in higher eukaryotes also provides spatio-temporal control mechanisms of cellular signaling and, therefore, the machinery of endocytosis is itself under tight regulation. The complexity of this system has been difficult to fully dissect, and many basic aspects remain to be discovered.

Understanding this important aspect of cellular function will provide a better knowledge of the spatio-temporal coordination of signaling pathways. A powerful tool to functionally dissect membrane trafficking routes is to monitor successful infectious virus entry. Viruses have co-evolved into specific directions to take advantage of signaling cascades and either clathrin-mediated endocytosis, caveolae/raft-mediated endocytosis, macropinocytosis or non-clathrin, non-caveolae pathways to gain entry into host cells [10,11,12,13\*]. To study the more complex morphological and dynamic features of endocytic activity, fluorescent endocytic cargo and fluorescent protein-tagged markers or immunofluorescent stainings of endocytic pits, vesicles, and organelles are used [14,15].

These two principal types of assays provide a useful platform to explore the properties of the endocytic membrane system. Combined with RNA interference (RNAi), they allow us to systematically study the response of this highly complex cellular system to a large number of perturbations with single-gene specificity [16]. This makes it possible to address questions that go beyond the function of a single gene: we can simultaneously study a group of functionally related genes and gene products on the basis of similar loss-of-function phenotypes that arise upon silencing. Furthermore, we can assign a potential endocytic role to genes with unknown functions on

the basis of the similarity of their loss-of-function phenotype with previously identified endocytic phenotypes. By applying automated imaging techniques, we can extract and quantify phenotypic information with high accuracy, which is needed to generate functional networks of the identified genes.

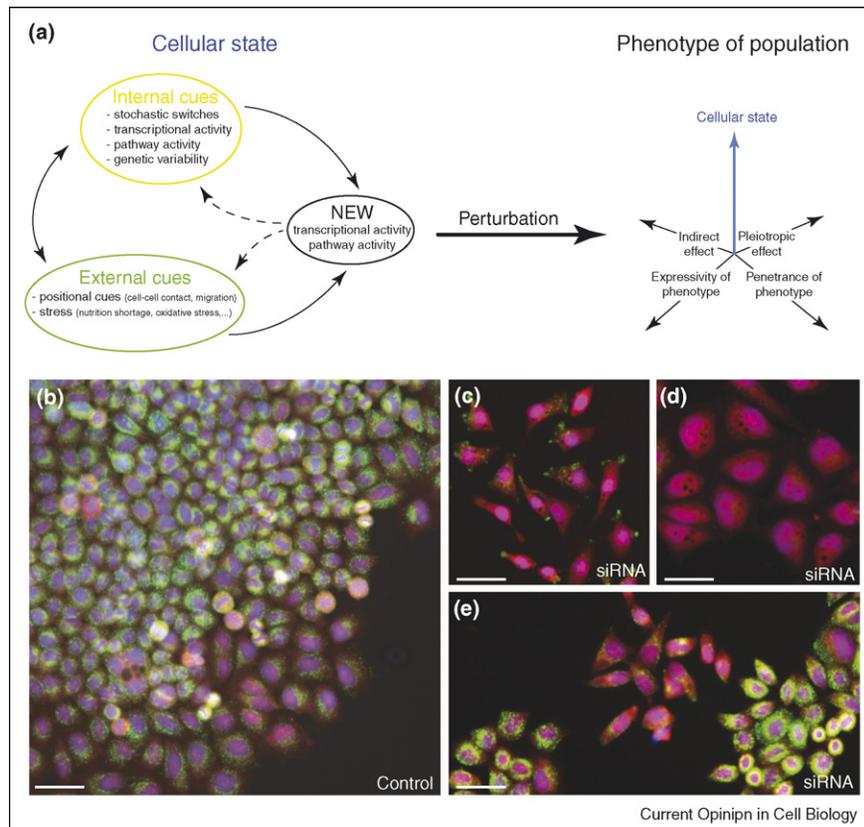
### Image-based phenotypic profiling

Currently, others and we are generating an integrated basis for high-throughput phenotypic siRNA (short-interfering RNA) screens of cellular processes in large populations of human cells. Depending on the particular focus of each study, either a candidate-based or a more global siRNA approach is performed, using three or more individual siRNA duplexes against each gene. After gene silencing in the cell line of choice (in our case cells in which the endocytic pathway under study is active), the endocytic assay is conducted. Besides imaging the endocytic membrane system as described above, additional

probes to visualize the various compartments of the secretory pathway, the cytoskeleton and the cell membrane, as well as the activity status of signaling pathways, are often essential for correct interpretation of the RNAi phenotype [17].

Applying such relatively complex assays to a high-throughput setup, in which a quantitative statement about each single cell is generated, needs sophisticated computational infrastructure and image analysis algorithms, and a high degree of method standardization. Currently, the methodology of 'high-throughput and high-content imaging' is at an early stage but rapidly developing. These activities hold high promises for the near future to increase speed, reliability, optical resolution, and sensitivity in molecular systems cell biology [18]. There are several open source [19,20] and commercially available image analysis software packages [21–23], actively improved and expanded by a growing scientific

**Figure 1**



**(a)** A schematic model of the development of variable phenotypes in a cell population. On the left hand, the interplay between different cues that lead to a cellular state is depicted. Arrows depict potential feedback mechanisms. On the right hand, parameters that could act on the different cellular states are represented in a multidimensional space. **(b–e)** Development of a phenotype following a perturbation. Unperturbed **(b)** and genetically perturbed (using siRNA) **(c–e)** populations of HeLa cells were incubated with fluorescent transferrin for 10 min. Cells were stained with CellTrace (molecular probes) to visualize the cell as a whole and nuclei were stained with DAPI. All images were taken with a high-throughput microscope at 20 $\times$  magnification and the channels were merged. Scale bars represent 50  $\mu$ m. **(b)** Unperturbed cells display a variety of phenotypes, **(c)** cells show different severities of a morphological phenotype and a pleiotropic effect of transferrin uptake, **(d)** cells show a pleiotropic effect with different severities and different penetrance (cell size (9/12), vacuoles (7/12), no transferrin uptake (12/12)). Note that different penetrance of phenotypes in one subpopulation is a sign of uncoupled pleiotropic effects. **(e)** Three clonal patches of cells display different phenotypes upon silencing of a single gene.

community. In the interest of science, we believe it is important to support methods whose algorithms have been subjected to peer review and are freely shared among scientists.

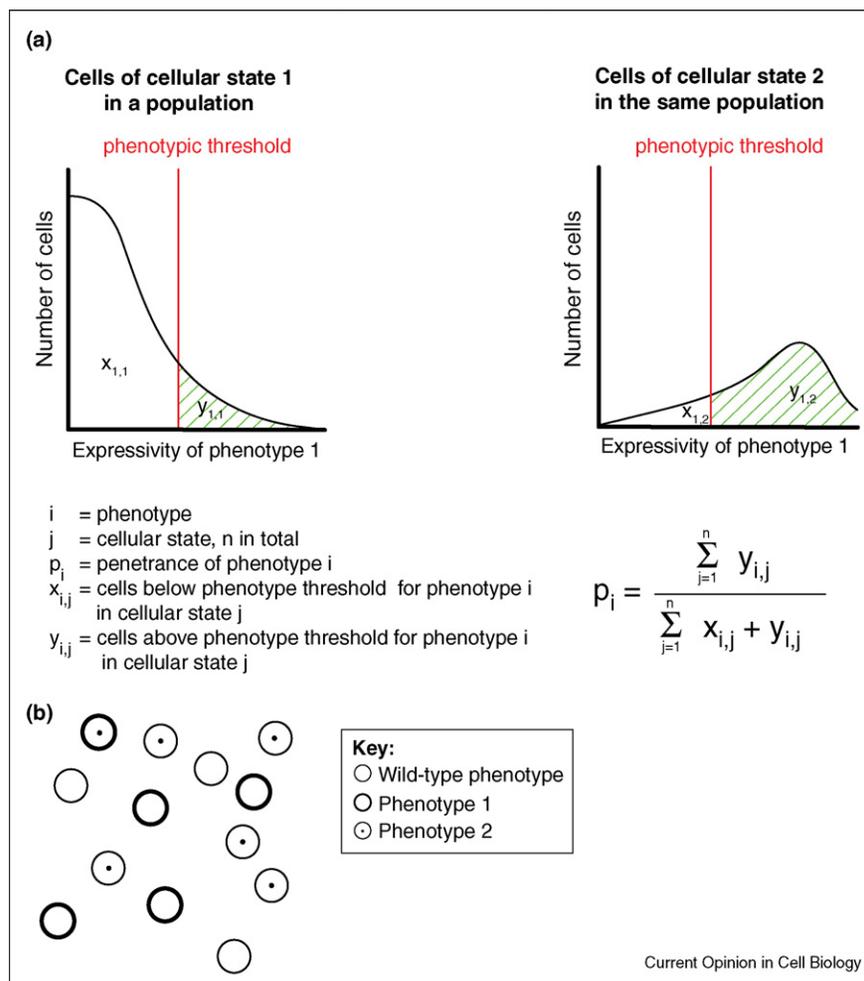
**Development of a phenotype in RNAi screens**

Before embarking upon identification and computer-aided classification of phenotypes from large imaged datasets, we believe it is important for scientists to consider the mechanisms by which gene loss-of-function phenotypes develop in a population of human cells. A cell population in human tissue culture is heterogeneous [21], because cells in the same population exist in different states [24]. Two obvious states are mitosis and apoptosis, but by continuously integrating extrinsic and intrinsic cues, a single cell in a population can have a variety of different states of gene expression and pathway activity. In Figure 1 we illustrate conceptually how this hetero-

geneity arises in an unperturbed cell population. We also give examples of how phenotypic heterogeneity (fluorescent transferrin uptake, cell size, presence of vacuoles) develops as a result of gene perturbation (Figure 1b–d). Extrinsic factors for this heterogeneity might be differences in cell–cell contacts, in cell extracellular matrix (ECM) contacts, or gaps in a monolayer, leading to specific gene expression programs [25]. Another source of heterogeneity in a population of cells could arise from genetic or epigenetic changes that are inherited during cell division, leading to the formation of genetic mosaics. Such clonal patches can differ morphologically, as well as in their gene expression and signaling pathway activity, as exemplified by clathrin-mediated uptake of transferrin (Figure 1e).

Internal and external cues create a highly variable population of cells, with the full range of possible cellular states

Figure 2



**(a)** Formal description of the penetrance of a phenotype in a cell population consisting of multiple cellular states. The severity or expressivity of a phenotype and its subpopulation penetrance can differ depending on the cellular state. The sum of the cells of one phenotype in a cellular state ( $y_{i,j}$ ) divided by the total cells ( $x_{i,j} + y_{i,j}$ ) defines the penetrance of the phenotype in a population. **(b)** Pleiotropic effects can have different penetrance and can be coupled or independent of each other.

(Figure 1a, b). It is important to realize that this is the situation in an unperturbed population of cells. A genetic perturbation in a population of individual cells will therefore create a different distribution of cellular states with additional dimensions (Figure 1a). This observation raises important questions as to how RNAi phenotypes arise in human tissue culture cells, and we here discuss and speculate on the factors that contribute to the development of such phenotypes.

#### Phenotype penetrance and expressivity

In classical genetics, a gene loss-of-function phenotype is penetrant when it is manifest in a proportion of individuals within a genetically homogenous population [26]. We here expand on that by posing that the different cellular states in a population of isogenic cells form subpopulations with different penetrance (Figure 2a). Depending on the frequency of the particular cellular state in which the phenotype is apparent, the phenotype might have a very low penetrance. There are numerous examples from classical genetics where a low penetrance of a particular loss-of-function phenotype is still significant [27,28]. Moreover, in mutants that are genetically null (complete lack of gene product), it can occur that the penetrance is less than 100% [29–31].

A related phenomenon known from classical genetics is the expressivity or severity of a phenotype, which is the variability in the degree with which a phenotype becomes apparent in an individual organism, and by analogy, an individual cell. As with penetrance, expressivity can also be modified by genetic and environmental influences. Both parameters need to be taken into account when either qualitative or quantitative statements are made. In fact, expressivity can be used to define a phenotypic threshold (Figure 2a).

#### Genetic redundancy

In several cases, silencing of a gene via RNAi leads to no obvious phenotype, even though the gene is known to have an important role (usually concluded from overexpression of dominant, active or negative forms) [32,33]. The lack of a phenotype by RNAi can be explained by genetic or functional redundancy, where the activity of a gene is lost, yet compensated by the function of a different gene or a different pathway [34,35]. In model organisms, redundant genes are discovered by conducting synthetic screens and studied by performing double mutant analyses [36,37]. Genetic techniques in human cell lines are limited, but similar studies could be undertaken by combinatorial RNAi screens, or by combining RNAi with dominant-negative mutants or small molecule inhibitors [38,39].

#### Pleiotropic effects

Unlike genetic redundancy, gene loss-of-function can also result in multiple phenotypes. What classical geneti-

cists refer to as pleiotropy [26] can also be observed in a population of cultured cells, with different cells showing different phenotypes, and/or the same cells sharing more than one phenotype (Figure 1d, Figure 2b). While different cells showing different phenotypes could be explained by the perturbation acting on different cellular states, different phenotypes within one cell could be the result of the gene product participating in more than one biological process, or has signaling functions on multiple target molecules.

#### Indirect effects

Determining an RNAi phenotype will also be complicated by a potential indirect, although ‘on-target’ effect of gene silencing. A phenotype can be the result of affecting a general, essential cellular process, or affecting a process that ‘feeds’ into the one perturbed by the knockdown (Figure 1a). An obvious example of the first case is that genes regulating the translational capacity or the metabolic status of a cell are expected to have widespread influences on various biological activities. Examples of the latter might be imagined as effects that act through changing the distribution of cellular states. Even though such phenotypes are interesting and important for the understanding of a cellular process, they will not inform us about the components of the core machinery of the studied process.

#### Note on transfection efficiency and off-target effects

Attempting to infer siRNA transfection efficiency from phenotype penetrance could be misleading. Although transfection efficiency is typically high, some or occasionally most cellular states might not show a phenotype because the targeted gene is either not involved, or not expressed. Multiple siRNAs against the same gene can show different phenotypes, an observation that is often interpreted as off-target effects. An additional explanation could be that depletion levels of the same gene can have different penetrance. Such differences can well be imagined to give rise to different, population-averaged mRNA profiles obtained from microarray experiments [40,41], and might in the future even be harnessed to understand phenotypic heterogeneity.

Taken together, if we want to reveal the full spectrum of a gene’s function in a particular cellular process, which includes the heterogeneity of that process in a population, we must be able to account for and integrate the phenomena described above. It is thus important to be careful with conclusions drawn from observations made with small numbers of cells, or to treat a cell population as a homogeneous unit. Therefore, what has been originally described in classical genetics, can be largely adopted by cell biologists using genetics in mammalian cell populations with single-cell resolution.

## Classification methods

Classical genetics can, however, suffer from small sampling size and non-quantitative, potentially biased classifications of phenotypes. Computer-aided, image-based RNAi screening with single-cell resolution overcomes these problems. It has therefore the potential to mechanistically explain the phenomena from classical genetics, and to reveal the underlying molecular and physical principles, such as stochastic effects and regulatory circuits in cell fate determination [42\*,43,44\*,45]. For most cellular systems, however, such as clathrin-mediated endocytosis, these are completely unexplored aspects of systems behavior. It will be important that we change the methods by which we identify, classify, and integrate gene loss-of-function phenotypes in large populations of cells. We therefore introduce and discuss some of the principal concepts from computer science and statistics that we expect to be particularly important for this development.

Visual classification of phenotypes is a proven and powerful method to find different phenotypes, especially when the phenotype is complex, subtle, or of low penetrance [46–48]. It is, however, clear that for formal and comprehensive phenotypic descriptions we need unbiased and quantitative methods. These methods should not employ population averaging, should be able to distinguish complex phenotypes at least as good as a human expert, and should be able to do that for many phenotypes in a short period of time. Both supervised and unsupervised machine learning will be essential techniques for computer-based classification. The basis for both is an image analysis that allows the segmentation and detection of cells, and the extraction of many (>100) quantitative features (e.g. shape, intensity, and texture features) from these cells.

### Supervised classification using machine learning

Individual numerical features or parameters from one identified object (e.g. a cell) can be combined in an  $n$ -dimensional vector [49], generally called a feature vector. Supervised learning approaches use such feature vectors to create functions defining the boundaries between phenotypic classes from a training dataset. The training data consist of cells with prototype phenotypes that are defined by an expert. The two most commonly used computational methods to mathematically define phenotype boundaries are neural networks or Support Vector Machines (SVM) [50]. A well-known example is the use of machine learning to detect different states of the mitotic cycle on the basis of morphology, intensity, and texture features from DAPI-stained nuclei [51\*]. In addition, machine learning can be employed to detect particular cell shape phenotypes, using features from fluorescent probes that stain the cell membrane [52\*\*].

Generally speaking, supervised machine learning can transform non-comprehensive numerical values from

cells into biological meaningful phenotypes. Once the training of a prototype phenotype is complete, the separating function allows a subsequent unbiased phenotype classification in a quantitative manner.

There is, however, a note of caution for feature selection: Not all features are useful for classification, and some features will disturb the classification outcome [52\*\*,53\*\*]. Furthermore, we do not yet have sufficient features to separate between more sophisticated phenotypes, such as complex intracellular staining patterns. Here, new image analysis algorithms need to be developed. Finally, supervised methods are believed to miss unknown phenotypes. However, by doing multiple rounds of computer-guided classification, one might in the future be able to accurately classify most phenotypes. Thus, the approach might allow not only the discovery of new genes involved in known phenotypes, but also the comprehensive analysis of all possible phenotypic states of a cellular system.

### Unsupervised methods

If features have been carefully selected, feature vectors can be hierarchically clustered in an unsupervised manner to generate phenoclusters [54\*]. Phenoclustering has been applied to group genes with similar function and to map previously uncharacterized genes or chemical compounds (drugs) to an already known group [53\*\*].

Such clustering brings the advantage that it provides a statistical value for the similarity of each phenotype with every other phenotype [55]. By applying a threshold below which the similarity is significant, the sets of phenotypes will fall into phenotypic networks, which represent the functional relations between the perturbed genes. In addition, the phenotypic distances can be compared or combined with protein–protein interaction distances, synthetic lethal interaction distances, or gene expression profiling distances to find significant functional modules of molecular cellular components with a similar loss-of-function phenotype [56].

Unsupervised methods have, however, a large drawback. Usually, each feature in the feature vector has the same weight in determining the distance between feature vectors. If a small difference in the value of one particular feature represents a clear boundary between two phenotypes, this will not be detected. By contrast, supervised approaches can select different features and can put different weight on specific features every time a specific phenotype class needs to be defined.

### From classifying individual cells to cell populations

It would be a great loss of information if each carefully analyzed and classified single cell would next be population-averaged to obtain a single readout of a population of cells. Since we have the possibility to assign each cell to

more than one phenotype class, either by supervised methods or with unsupervised phenoclustering, we can define phenotype indices. These represent the fraction of cells in the population with a certain phenotype or a certain phenotype combination. Such indices are a feature vector themselves, where each feature constitutes a phenotype index. These index vectors can be used in supervised and unsupervised approaches to group genes not only according to the actual loss-of-function phenotype they induce in a single cell, but according to the variety of phenotypes they induce in a population of cells. This information, both from individual cells and from cell populations might be used to create comprehensive hierarchical schemes of genetic interactions that regulate cellular processes and coordinate them within a population of cells.

## Conclusions

Current accomplishments have advanced our understanding of the cell on the molecular level, but at the same time revealed more of its complexity. Phenotypes arising from gene silencing are good witnesses for this complexity. Quoting Sydney Brenner in his 1974 paper “How genes might specify the complex structures found in higher organisms is a major unsolved problem of biology”, the paradigm for today’s efforts should be to carefully collect quantitative data from phenotypes and utilize them for the generation of mathematical and physical models. These models could, on the one hand, describe biological processes of interest and, on the other, predict their dynamic behavior over time, not only in a single cell, but also in a population of cells.

## Acknowledgements

We would like to thank the members of the Pelkmans lab for discussions, and especially Berend Snijder, Herbert Polzhofer, and Dr Pauli Rämö for critical comments. LP is supported by the Swiss National Science Foundation (SNSF), SystemsX.ch, the European Union, and the ETH Zürich. LS is a Bonizzi-Theler Foundation fellow. RS is part of the Molecular Life Science PhD program, Zürich and is supported by the SNSF. We apologize to colleagues whose work has not been cited owing to the space limitations.

## References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
  - of outstanding interest
1. Reddien PW, Horvitz HR: **CED-2/CrkII and CED-10/Rac control phagocytosis and cell migration in *Caenorhabditis elegans***. *Nat Cell Biol* 2000, **2(3)**:131-136.
  2. Lundquist EA, Reddien PW, Hartwig E, Horvitz HR, Bargmann Cl: **Three *C. elegans* Rac proteins and several alternative Rac regulators control axon guidance, cell migration and apoptotic cell phagocytosis**. *Development* 2001, **128(22)**:4475-4488.
  3. Kinchen JM, Cabello J, Klingele D, Wong K, Feichtinger R, Schnabel H, Schnabel R, Hengartner MO: **Two pathways converge at CED-10 to mediate actin rearrangement and corpse removal in *C. elegans***. *Nature* 2005, **434(7029)**:93-99.
  4. Emery G, Knoblich JA: **Endosome dynamics during development**. *Curr Opin Cell Biol* 2006, **18(4)**:407-415.
  5. von Zastrow M, Sorkin A: **Signaling on the endocytic pathway**. *Curr Opin Cell Biol* 2007, **19(4)**:436-445.
  6. Jékely G, Sung HH, Luque CM, Rørth P: **Regulators of endocytosis maintain localized receptor tyrosine kinase signaling in guided migration**. *Dev Cell* 2005, **9(2)**:197-207.
  7. Vaccari T, Bilder D: **The *Drosophila* tumor suppressor vps25 prevents nonautonomous overproliferation by regulating notch trafficking**. *Dev Cell* 2005, **9(5)**:687-698.
  8. Lu H, Bilder D: **Endocytic control of epithelial polarity and proliferation in *Drosophila***. *Nat Cell Biol* 2005, **7(12)**:1232-1239.
  9. del Pozo MA, Balasubramanian N, Alderson NB, Kiosses WB, Grande-García A, Anderson RG, Schwartz MA: **Phospho-caveolin-1 mediates integrin-regulated membrane domain internalization**. *Nat Cell Biol* 2005, **7(9)**:901-908.
  10. Pelkmans L, Helenius A: **Insider information: what viruses tell us about endocytosis**. *Curr Opin Cell Biol* 2003, **15(4)**:414-422.
  11. Damm EM, Pelkmans L, Kartenbeck J, Mezzacasa A, Kurzchalia T, Helenius A: **Clathrin- and caveolin-1-independent endocytosis: entry of simian virus 40 into cells devoid of caveolae**. *J Cell Biol* 2005, **168(3)**:477-488.
  12. Pelkmans L: **Secrets of caveolae- and lipid raft-mediated endocytosis revealed by mammalian viruses**. *Biochim Biophys Acta* 2005, **1746(3)**:295-304.
  13. Marsh M, Helenius A: **Virus entry: open sesame**. *Cell* 2006, **124(4)**:729-740.
- A description of the endocytic pathways used, as well as the signaling cascades activated by viruses to infect host cells.
14. Le Roy C, Wrana JL: **Clathrin- and non-clathrin-mediated endocytic regulation of cell signalling**. *Nat Rev Mol Cell Biol* 2005, **6(2)**:112-126.
  15. Mills IG: **The interplay between clathrin-coated vesicles and cell signalling**. *Semin Cell Dev Biol* 2007, **18(4)**:459-470.
  16. Pelkmans L, Fava E, Grabner H, Hannus M, Habermann B, Krausz E, Zerial M: **Genome-wide analysis of human kinases in clathrin- and caveolae/raft-mediated endocytosis**. *Nature* 2005, **436(7047)**:78-86.
  17. Moffat J, Sabatini DM: **Building mammalian signalling pathways with RNAi screens**. *Nat Rev Mol Cell Biol* 2006, **7(3)**:177-187.
  18. Pepperkok R, Ellenberg J: **High-throughput fluorescence microscopy for systems biology**. *Nat Rev Mol Cell Biol* 2006, **7(9)**:690-696.
  19. Abramoff M, Magalhaes P, Ram SJ: **Image processing with ImageJ**. *Biophotonics Int* 2004, **11**:36-42.
  20. Lamprecht MR, Sabatini DM, Carpenter AE: **CellProfiler: free, versatile software for automated biological image analysis**. *Biotechniques* 2007, **42(1)**:71-75.
- In this paper the CellProfiler software is introduced to the scientific community. CellProfiler is an open source, easy-to-use software that enables biologists to comprehensively and quantitatively analyze imaged cells and other biological samples without programming knowledge. It offers a starter set of segmentation and feature extraction algorithms, and it can handle batch processing for large datasets <http://www.cellprofiler.org/>.
21. Perlman ZE, Slack MD, Feng Y, Mitchison TJ, Wu LF, Altschuler SJ: **Multidimensional drug profiling by automated microscopy**. *Science* 2004, **306(5699)**:1194-1198.
  22. Lindblad J, Wählby C, Bengtsson E, Zaltsman A: **Image analysis for automatic segmentation of cytoplasm and classification of Rac1 activation**. *Cytometry A* 2004, **57(1)**:22-33.
  23. Zhou X, Cao X, Perlman Z, Wong ST: **A computerized cellular imaging system for high content analysis in Monastrol suppressor screens**. *J Biomed Inform* 2006, **39(2)**:115-125.
  24. Levsky JM, Singer RH: **Gene expression and the myth of the average cell**. *Trends Cell Biol* 2003, **13(1)**:4-6.
  25. Müller EJ, Williamson L, Kolly C, Suter MM: **Outside-in signaling through integrins and cadherins: a central mechanism to control epidermal growth and differentiation?** *J Invest Dermatol* 2008, **128(3)**:501-516.

26. Griffiths AJF, Miller JH, Suzuki DT, Lewontin RC, Gelbart WM: **An Introduction to Genetic Analysis**. New York: WH Freeman and Company; 1997.
27. Desai CJ, Krueger NX, Saito H, Zinn K: **Competition and cooperation among receptor tyrosine phosphatases control motoneuron growth cone guidance in Drosophila**. *Development* 1997, **124**:1941-1952.
28. Howard RM, Sundaram MV: **C. elegans EOR-1/PLZF and EOR-2 positively regulate Ras and Wnt signaling and function redundantly with LIN-25 and the SUR-2 Mediator component**. *Genes Dev* 2002, **16**:1815-1827.
29. Lackner MR, Kim SK: **Genetic analysis of the Caenorhabditis elegans MAP kinase gene mpk-1**. *Genetics* 1998, **150**(1):103-117.
30. Eichers ER, Abd-El-Barr MM, Paylor R, Lewis RA, Bi W, Lin X, Meehan TP, Stockton DW, Wu SM, Lindsay E *et al.*: **Phenotypic characterization of Bbs4 null mice reveals age-dependent penetrance and variable expressivity**. *Hum Genet* 2006, **120**(2):211-226.
31. Laub F, Dragomir C, Ramirez F: **Mice without transcription factor KLF7 provide new insight into olfactory bulb development**. *Brain Res* 2006, **1103**(1):108-113.
32. Tautz D: **Redundancies, development and the flow of information**. *Bioessays* 1992, **14**(4):263-266.
33. Thomas JH: **Thinking about genetic redundancy**. *Trends Genet* 1993, **9**(11):395-399.
34. Ferguson E, Horvitz HR: **The multivulva phenotype of certain C. elegans mutants results from defects in two functionally redundant pathways**. *Genetics* 1989, **123**:109-121.
35. Ellis RE, Jacobson DM, Horvitz HR: **Genes required for the engulfment of cell corpses during programmed cell death in Caenorhabditis elegans**. *Genetics* 1991, **129**:79-94.
36. Jorgensen EM, Mango SE: **The art and design of genetic screens: Caenorhabditis elegans**. *Nat Rev Genet* 2002, **3**(5):356-369.
37. Ceol CJ, Stegmeier F, Harrison MM, Horvitz HR: **Identification and classification of genes that act antagonistically to let-60 Ras signaling in Caenorhabditis elegans vulval development**. *Genetics* 2006, **173**(2):709-726.
38. Castanotto D, Sakurai K, Lingeman R, Li H, Shively L, Aagaard L, Soifer H, Gatignol A, Riggs A, Rossi JJ: **Combinatorial delivery of small interfering RNAs reduces RNAi efficacy by selective incorporation into RISC**. *Nucleic Acids Res* 2007, **35**(15):5154-5164.
39. Sahin O, Löbke C, Korf U, Appelhans H, Sülthmann H, Poustka A, Wiemann S, Arlt D: **Combinatorial RNAi for quantitative protein network analysis**. *Proc Natl Acad Sci U S A* 2007, **104**(16):6579-6584.
40. Jackson AL, Bartz SR, Schelter J, Kobayashi SV, Burchard J, Mao M, Li B, Cavet G, Linsley PS: **Expression profiling reveals off-target gene regulation by RNAi**. *Nat Biotechnol* 2003, **21**(6):635-637.
41. Kittler R, Surendranath V, Heninger AK, Slabicki M, Theis M, Putz G, Franke K, Caldarelli A, Grabner H, Kozak K *et al.*: **Genome-wide resources of endoribonuclease-prepared short interfering RNAs for specific loss-of-function studies**. *Nat Methods* 2007, **4**(4):337-344.
42. Kussell E, Leibler S: **Phenotypic diversity, population growth, and information in fluctuating environments**. *Science* 2005, **309**(5743):2075-2078.  
A study that describes the growth of clonal populations in relation to their phenotypic switching. The type of switching depends on information acquired from fluctuation environments.
43. Süel GM, Garcia-Ojalvo J, Liberman LM, Elowitz MB: **An excitable gene regulatory circuit induces transient cellular differentiation**. *Nature* 2006, **440**(7083):545-550.
44. Süel GM, Kulkarni RP, Dworkin J, Garcia-Ojalvo J, Elowitz MB: **Tunability and noise dependence in differentiation dynamics**. *Science* 2007, **315**(5819):1716-1719.  
The authors used a differentiation process in *B. subtilis* to study the effect of stochastic fluctuations or 'noise' on cellular behavior. Noise levels correlated with the frequency of differentiation events and contributed to alternative and dynamic behaviors of the system.
45. Acar M, Mettetal JT, van Oudenaarden A: **Stochastic switching as a survival strategy in fluctuating environments**. *Nat Genet* 2008, **40**(4):471-475.
46. Sönnichsen B, Koski LB, Walsh A, Marschall P, Neumann B, Brehm M, Alleaume AM, Artelt J, Bettencourt P, Cassin E: **Full-genome RNAi profiling of early embryogenesis in Caenorhabditis elegans**. *Nature* 2005, **434**(7032):462-469.
47. Kittler R, Pelletier L, Heninger AK, Slabicki M, Theis M, Miroslaw L, Poser I, Lawo S, Grabner H, Kozak K *et al.*: **Genome-scale RNAi profiling of cell division in human tissue culture cells**. *Nat Cell Biol* 2007, **9**(12):1401-1412.
48. Bai J, Binari R, Ni JQ, Vijayakanthan M, Li HS, Perrimon N: **RNA interference screening in Drosophila primary cells for genes involved in muscle assembly and maintenance**. *Development* 2008, **135**(8):1439-1449.
49. Joachims T: **Text Categorization with Support Vector Machines: Learning with Many Relevant Features** Berlin/Heidelberg: Springer; 1998. pp. 137-142.
50. Vapnik VN: **Statistical Learning Theory**. New York: Wiley-Interscience; 1998.
51. Neumann B, Held M, Liebel U, Erfle H, Rogers P, Pepperkok R, Ellenberg J: **High-throughput RNAi screening by time-lapse imaging of live human cells**. *Nat Methods* 2006, **3**(5):385-390.  
A pilot study that combines RNAi with time-lapse imaging. Several technical hurdles are overcome to give rise to a scalable platform to study mitosis over time. It is the basis for the MitoCheck project <http://www.mitocheck.org/>.
52. Bakal C, Aach J, Church G, Perrimon N: **Quantitative morphological signatures define local signaling networks regulating cell morphology**. *Science* 2007, **316**(5832):1753-1756.  
In this study a quantitative morphological profiling of perturbed cells leads to the generation of local signaling networks. Genes in these networks act to regulate cell protrusion, adhesion, and tension.
53. Loo LH, Wu LF, Altschuler SJ: **Image-based multivariate profiling of drug responses from single cells**. *Nat Methods* 2007, **4**(5):445-453.  
The authors present a multivariate method for classifying human cells on the basis of approximately 300 single-cell phenotypic measurements obtained after the application of chemical compounds. They minimize the feature set to roughly 20 informative features for each compound and marker set. This reduction enhances human interpretability and improves detection of cellular phenotypic changes with little loss of classification accuracy.
54. Piano F, Schetter AJ, Morton DG, Gunsalus KC, Reinke V, Kim SK, Kemphues KJ: **Gene clustering based on RNAi phenotypes of ovary-enriched genes in C. elegans**. *Curr Biol* 2002, **12**(22):1959-1964.  
The authors present a gene clustering approach based on RNAi-induced phenotypes in the *C. elegans* embryo. The obtained 'phenoclusters' contain genes with a similar loss-of-function phenotype, show a good correlation with functional predictions, and can, therefore, be used to predict functions of uncharacterized genes.
55. Grupe A, Germer S, Usuka J, Aud D, Belknap JK, Klein RF, Ahluwalia MK, Higuchi R, Peltz G: **In silico mapping of complex disease-related traits in mice**. *Science* 2001, **292**(5523):1915-1918.
56. Walhout AJ, Rebol J, Shtanko O, Bertin N, Vaglio P, Ge H, Lee H, Doucette-Stamm L, Gunsalus KC, Schetter AJ *et al.*: **Integrating interactome, phenome, and transcriptome mapping data for the C. elegans germline**. *Curr Biol* 2002, **12**(22):1952-1958.