showing both descriptive (perturbed pathways) and performance metrics (classifier accuracy) for frequent driver mutations from the Cancer Genome Atlas dataset would be a useful resource in target identification and drug discovery processes. Finally, inferring intricate oncogenic transcriptional states could benefit from an integrative framework that accounts for (1) epigenetic changes that result in transcriptional plasticity and drug resistance (Flavahan et al., 2017), and (2) that harnesses single-cell data to facilitate the measurement of clonal diversity and crosstalk between cancer cells and tumor microenvironment (Datlinger et al., 2017). Nevertheless, the ability of Onco-GPS to delineate the transcriptional effects of tumor-driving mutations provides an important step forward in defining the precision of our medicines.

## REFERENCES

Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehár, J., Kryukov, G.V., Sonkin, D., et al. (2012). Nature *483*, 603–607.

Chen, J.C., Alvarez, M.J., Talos, F., Dhruv, H., Rieckhof, G.E., Iyer, A., Diefes, K.L., Aldape, K., Berens, M., Shen, M.M., and Califano, A. (2014). Cell *159*, 402–414.

Dang, C.V., Reddy, E.P., Shokat, K.M., and Soucek, L. (2017). Nat. Rev. Cancer *17*, 502–508.

Datlinger, P., Rendeiro, A.F., Schmidl, C., Krausgruber, T., Traxler, P., Klughammer, J., Schuster, L.C., Kuchler, A., Alpar, D., and Bock, C. (2017). Nat. Methods *14*, 297–301.

Flavahan, W.A., Gaskell, E., and Bernstein, B.E. (2017). Science *357*, 357.

Hyman, D.M., Puzanov, I., Subbiah, V., Faris, J.E., Chau, I., Blay, J.-Y., Wolf, J., Raje, N.S., Diamond, E.L., Hollebecque, A., et al. (2015). N. Engl. J. Med. *373*, 726–736.

Iorio, F., Knijnenburg, T.A., Vis, D.J., Bignell, G.R., Menden, M.P., Schubert, M., Aben, N., Gonçalves, E., Barthorpe, S., Lightfoot, H., et al. (2016). Cell *166*, 740–754.

Kim, E. (2017). Cell Syst. *5*, this issue, 105–118.

Rizos, H., Menzies, A.M., Pupo, G.M., Carlino, M.S., Fung, C., Hyman, J., Haydu, L.E., Mijatov, B., Becker, T.M., Boyd, S.C., et al. (2014). Clin. Cancer Res. *20*, 1965–1977.

Seashore-Ludlow, B., Rees, M.G., Cheah, J.H., Cokol, M., Price, E.V., Coletti, M.E., Jones, V., Bodycombe, N.E., Soule, C.K., Gould, J., et al. (2015). Cancer Discov. *5*, 1210–1223.

# When Seeing Isn't Believing: How Math Can Guide Our Interpretation of Measurements and Experiments

**Paul Macklin**[1,*]
[1]Intelligent Systems Engineering, Indiana University, Bloomington, IN 47408, USA
*Correspondence: Paul.Macklin@MathCancer.org
http://dx.doi.org/10.1016/j.cels.2017.08.005

**Mathematical thought experiments probe the meaning and pitfalls of experimental measurements and demonstrate that caution is in order when measuring heterogeneity.**

As powerful new technologies allow us to probe single-cell dynamics in high throughput, are we at risk of reaching the wrong conclusions? An intriguing paper by Schumacher, Maini and Baker in this issue of *Cell Systems* uses a mathematical model of collective cell migration—a 21st-century computational thought experiment—to examine the potential pitfalls of measuring heterogeneity (Schumacher et al., 2017). In the process, they highlight the potential for mathematical modeling to help us better understand the meaning and limits of experimental measurements.

Cellular heterogeneity has emerged as an important factor in studying clonal dynamics in cancer, predicting therapeutic response, and assessing the limitations of molecular profiling (e.g., Fisher et al. (2013)). An accurate appreciation of heterogeneity is clinically important: undersampling a heterogeneous tumor, for example, runs the risk of making treatment decisions from unrepresentative subclones or missing aggressive clones altogether. Cellular heterogeneity allows a cell population to explore the space of possibilities, increasing the odds of finding a strategy that circumvents a treatment. Hence, there is an increasing interest in quantitating cellular heterogeneity at a variety of levels, including genomic, metabolic, motility, and other characterizations.

Single-cell molecular analyses and single-cell tracking in time course experiments have opened fascinating new possibilities in examining cell-to-cell variability, but we need to develop quantitative measures of heterogeneity to make full use of these data and to relate heterogeneity to disease phenotypes and clinical outcomes. Essentially, this is a challenge of relating molecular-level heterogeneity within the genome or proteome to observable variations in cellular behavior. One behavior that has received much attention is cell motility because it is clinically important for understanding metastasis and can be directly quantified. For example, the peak delay time—the time difference at which two cells' velocities are most aligned—has emerged as a metric to quantitate collective cell migration; measure this across a migrating collective of cells, and we can quantitate heterogeneity in collective cell migration (Sharma et al., 2015).

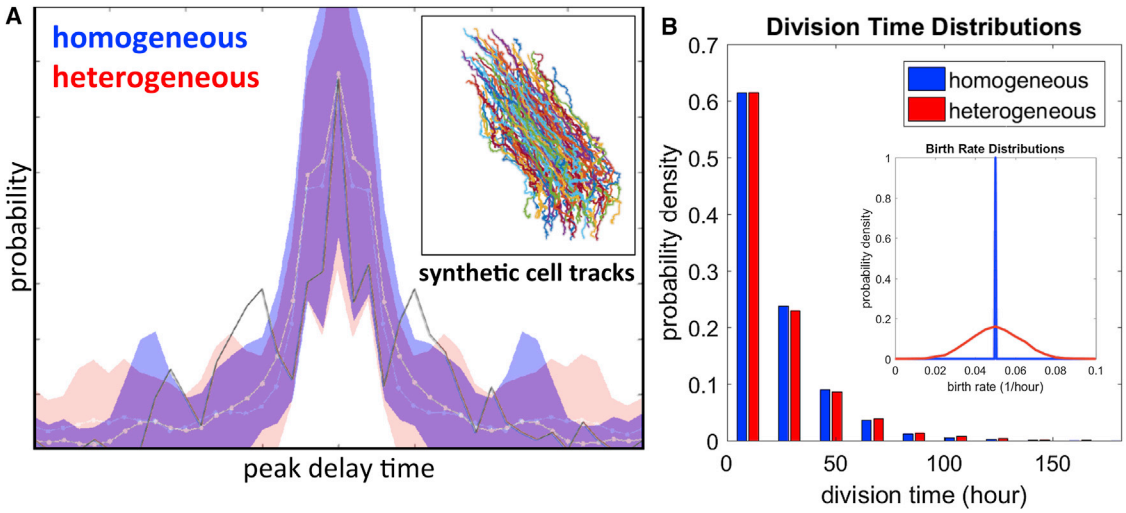


**Figure 1. Mathematical Models Show That Homogeneous Cells Can Exhibit Apparent Heterogeneity**

(A) Schumacher et al. (2017) used a mathematical model of collective cell motion to generate synthetic cell tracks (inset, modified here from Figure 3b in Schumacher et al. (2017)) and analyze them for heterogeneity. Simulations with completely identical cells (blue, modified and recolored from their Figure 4a) could be fitted to match available experimental datasets, just as well as simulations with 10% leader cells and 90% followers (red, modified and recolored from their Figure 4c).

(B) Simple mathematical models can also show this trend. Tracking the time for 10,000 identical cells to divide (blue) produces the same exponential distribution as for 10,000 cells with true heterogeneity in their proliferation characteristics (red). The code to generate this figure is available at http://MathCancer.org/CS-preview-2017.

Assessing whether such metrics truly measure cell heterogeneity requires a ground truth, and this is where the mathematical modeling comes in. Schumacher and colleagues developed a general mathematical model of migrating cells—with mechanical interactions, external directional cues, and random fluctuations—to generate collections of synthetic cell migration tracks. (See the sample tracks in Figure 1A inset.) Because the team had complete control over the parameters and behavioral rules of the model, they had perfect knowledge of the ground truth. This allowed them to run computational thought experiments to assess the impact of cellular heterogeneity (variation in the parameter values) on multicellular systems behavior, and ultimately metrics of heterogeneity like peak correlation times. And more broadly, this allowed them to ask, if we vary the cells' parameters from homogeneous to heterogeneous, what would the data look like? Would we recognize heterogeneity if it was there?

Schumacher, Maini, and Baker used their framework to generate cell tracks for homogeneous cell populations (cells with identical governing rules and parameters—the "genes" of the model), computed the distribution of peak delay times for the synthetic migrating cohorts, and calculated the heterogeneity (the standard deviation of the peak delays). Strikingly, this homogeneous set of cells showed significant heterogeneity in the migration metric (Figure 1A, blue data). That is, if we look for heterogeneity (by computing this metric), we are likely to find it, even if the underlying cells are completely homogeneous. Indeed, Schumacher and colleagues were able to fit homogeneous cell populations to two separate migration datasets published by Sharma et al. (2015).

Moreover, when Schumacher et al. introduced a truly heterogeneous cell population (10% are "leader cells" who follow their own directional cues, while the remaining 90% follow their neighbors), they were able to fit the experimental data equally well (Figure 1A, red data). That is, experimental measurements can exhibit apparent heterogeneity even when none is actually present, and cannot easily differentiate between homogeneous and heterogeneous cell populations. In the clinic, the results mean that cancer cells can demonstrate significant phenotypic heterogeneity without any underlying genomic heterogeneity. The good news is that a clinician may not need to seek a rare subclone to target the motile behavior; the bad news is that any cancer cells surviving a treatment could potentially give rise to the original motile behaviors. More broadly, it implies that even single cancer clones could explore the phenotypic space of possibilities, allowing a cancer to evade treatments without need for mutational events. Schumacher, Maini and Baker close their paper with extra insights on the implications for measuring heterogeneity and a call for caution when selecting null models for testing if a measurement of heterogeneity is significant.

Such mismatches between phenotypic observations and underlying molecular reality may not be rare, even for simpler measurements. As a simple computational thought experiment, consider a common mathematical model of proliferating cell populations: each virtual cell (with birth rate $r$—the cell's abstracted "oncogene") has probability of $r\Delta t$ of dividing between now ($t$) and a small time from now ($t+\Delta t$). If we track the time it takes for 10,000 cells (with identical $r$) to divide, the times are exponentially distributed (Figure 1B, blue data). (See the supplementary material in Macklin et al. (2012) for more on the mathematics underlying this thought experiment.) This homogeneous cell

population demonstrates apparent heterogeneity: its division times span over two orders of magnitude. Moreover, if we repeat this experiment with a normal distribution of *r* (a truly heterogeneous cell population), the cell division times follow an *identical* distribution (Figure 1B, red data).

Schumacher, Maini, and Baker's work highlights one of the key strengths of mathematics in biology: by creating systems where the ground truth is truly known—where virtual experiments are run on cells with fully understood "biology," using zero-error instruments with infinite sampling resolution—we can better understand the implications of measuring real-world cells using real-world instruments with errors and sampling limits. Poleszczuk et al. (2015) used mathematical models of clonal dynamics to understand the impact of limited biopsies in characterizing clones. We simulated the impact of variability in pipetting, seeding conditions, and limited observations on fitting cell proliferation rates; in some cases, measuring two identical lines with differing seeding conditions can yield significant ($p < 0.05$) but false differences in the estimated growth rates (Friedman and Macklin, 2017). Baker's group has used simulations to investigate the impact of assay geometry on estimates of cell motility and proliferation (Treloar et al., 2014), and more recently to evaluate the impact of temporal sampling limits on measuring motility and biotransport (Harrison and Baker, 2017).

Schumacher et al.'s work points to an exciting future. Mathematicians are banding together to create general-purpose modeling frameworks that can simulate the tissue-level behaviors that emerge from cell-scale hypotheses (e.g., Swat et al. (2012) and Mirams et al. (2013)). High-throughput computing can run thousands of simulations simultaneously, allowing us to broadly sample the space of possible hypotheses and ask which simulations best match the form and function of real tissues. If we could quantitate the structural and functional mismatch between a simulation and a tissue—find an error metric—we could perform automated, high-throughput thought experiments to seek the hypotheses that minimize the mismatch. In the long run, we could run the ultimate computational thought experiments: what cell rules do we need to engineer a purpose-built tissue or cell colony to remediate an aquifer, repair a tissue, or metabolize a xenobiotic? And can emerging techniques in synthetic biology, gene editing, and 3-D tissue printing bring our designs to reality?

## REFERENCES

Fisher, R., Pusztai, L., and Swanton, C. (2013). Br. J. Cancer *108*, 479–485.

Friedman, S.H., and Macklin, P. (2017). bioRxiv *175703*, http://dx.doi.org/10.1101/175703.

Harrison, J.U., and Baker, R.E. (2017). ArXiv e-prints *1708*, 1708.01562.

Macklin, P., Edgerton, M.E., Thompson, A.M., and Cristini, V. (2012). J. Theor. Biol. *301*, 122–140.

Mirams, G.R., Arthurs, C.J., Bernabeu, M.O., Bordas, R., Cooper, J., Corrias, A., Davit, Y., Dunn, S.J., Fletcher, A.G., Harvey, D.G., et al. (2013). PLoS Comput. Biol. *9*, e1002970.

Poleszczuk, J., Hahnfeldt, P., and Enderling, H. (2015). PLoS Comput. Biol. *11*, e1004025.

Schumacher, L.J., Maini, P.K., and Baker, R.E. (2017). Cell Syst., S2405-4712(17)30234-X.

Sharma, Y., Vargas, D.A., Pegoraro, A.F., Lepzelter, D., Weitz, D.A., and Zaman, M.H. (2015). Integr Biol (Camb) *7*, 1526–1533.

Swat, M.H., Thomas, G.L., Belmonte, J.M., Shirinifard, A., Hmeljak, D., and Glazier, J.A. (2012). Methods Cell Biol. *110*, 325–366.

Treloar, K.K., Simpson, M.J., McElwain, D.L., and Baker, R.E. (2014). J. Theor. Biol. *356*, 71–84.