

# CANCER'S HETEROGENEITY: Modeling Tumors' Diversity

By Alexander Gelfand

Cancer might spring from a single cell gone awry, but tumors are not monolithic collections of clones. Far from it: They contain many different types of cancer cells, all with their own mutations, proliferation rates, metastatic capacities, and drug responses.

This diversity pushes the limits of current diagnostic and treatment capabilities. A biopsy might miss a crucial subpopulation of tumor cells; and a treatment that works for one set of cells might be ineffective against another set within the same tumor. Moreover, greater heterogeneity is associated with worse outcomes for several types of cancer.

"We want to better understand how to treat tumors more effectively," says **Kornelia Polyak, MD, PhD**, a breast cancer researcher at the Dana-Farber Cancer Institute. And that's going to require learning a lot more about heterogeneity—including how it affects the way a tumor will respond to treatment, and how treatment itself may change the tumor.

To get a handle on cancer's heterogeneity, some researchers are using computational modeling and simulation. They aim to illuminate how the variety of cell types in a tumor influences cancer progression, and to predict the most effective course of treatment for a given tumor. To that end, they are using an assortment of tools almost as diverse as cancer itself, drawing upon fields ranging from machine learning to digital circuit design.

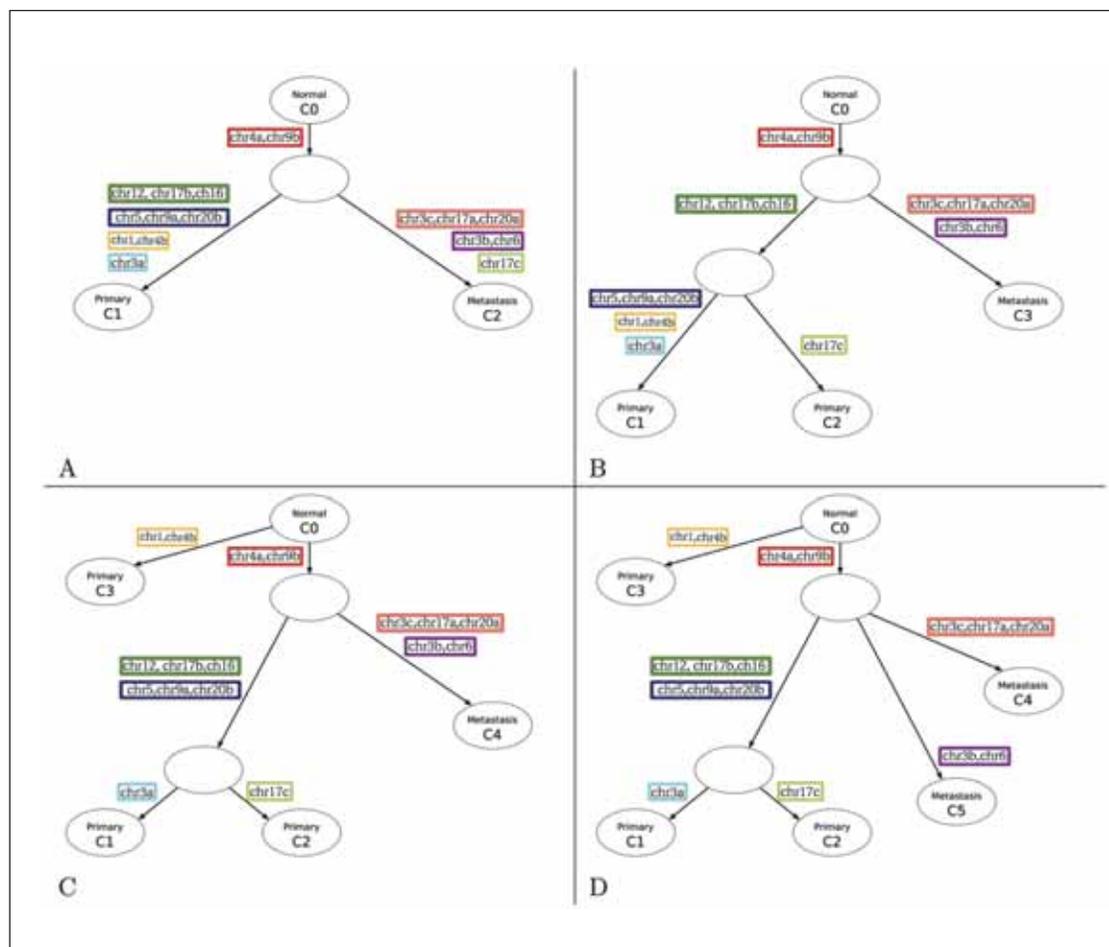
## Defining the Clones

Ignoring the presence of even small clonal subpopulations within a tumor can allow them to flourish; so defining the number and nature of clones from limited tissue samples is extremely important, though difficult. Statistician **Daniela Witten, PhD**, and her colleagues at the University of

Washington succeeded in doing just that—in work published in *PLoS Computational Biology* in July 2014—by applying some very sophisticated statistical techniques to some very rich next-generation sequencing data.

Witten and her collaborators began with multiple tissue samples from a 44-year-old breast-cancer patient. Some were from the patient's primary and metastatic tumors,

others from her healthy breast tissue. Using DNA sequencing, the team identified normal alleles found in both the healthy and cancerous samples, and abnormal ones (mutations) found only in the cancerous samples. But knowing which mutant alleles are present in the cancerous tissue is not enough. Clinicians need to know the genotypes of the cancerous clones. So Witten



After using statistical methods to predict the genotypes of clones in breast cancer samples, Witten and her colleagues honed in on the possible number of clones by reconstructing possible phylogenetic trees for 3, 4, 5 or 6 clones (A-D here). Nodes correspond to inferred clonal populations, with C0 corresponding to the normal clone; edges are annotated with mutations that occur between the parent and child clones. Mutations are grouped into a colored box if they occur on the same branch in all four phylogenies. The team concluded that the three-clone model (A) was too simple to explain the data while four clones (B) explained the data quite well. Adding further clones (C and D) increases the level of detail in the trees but provides little additional predictive value, and can exaggerate the significance of minor fluctuations in the data. Reprinted from H Zare, J Wang, A Hu et al., *Inferring Clonal Composition from Multiple Sections of a Breast Cancer*, *PLoS Comp Biol*, July 2014, doi:10.1371/journal.pcbi.1003703.g006.

used an approach called statistical machine learning to determine probabilistic estimates of the clone genotypes as well as estimates of the frequencies of those clones in the different tumor subsections. The researchers then modeled different numbers of possible clones (e.g., 3, 4, 5, 6) to see which one would best explain the genetic variety observed in the sequencing data.

In the end, the model worked best with four clones that mapped the clone genotypes and frequencies to the tumor subsections in a way that corresponded both to the physical anatomy of the tumors, and to phylogenetic trees that described how the clones could have evolved from normal tissue, accumulating mutations and branching off from one another over time. That kind of evolutionary insight matters, since knowing how and when one clone gives rise to another could potentially help inform treatment decisions, such as when to introduce

“We know that cancer changes fast, and we are really just trying to get a handle on how it’s changing, and why,” Witten says.

a particular anti-cancer drug. “We know that cancer changes fast, and we are really just trying to get a handle on how it’s changing, and why,” Witten says.

### What Doesn’t Kill Them Makes Them Stronger

Getting a handle on how and why cancer changes over time was in fact the primary goal of a series of computer simulations conducted by **Eleftheria Tzamali, PhD**, and her colleagues in the Computational Medicine Laboratory at the Institute of Computer Science, part of the Foundation for Research and Technology–Hellas, in Greece. In particular, Tzamali wanted to understand how different cell types combine with microenvironmental factors to influence the morphology and progression of tumors. For example, invasive behavior in glioma (a type of brain cancer) and breast-cancer cells has been tied to low oxy-

gen levels—though it has also been observed regardless of oxygen level. Many treatments specifically target proliferative tumor cells by modifying their vasculature and starving them of oxygen, and Tzamali wondered how that might affect a tumor that had more than one kind of cell in it.

To answer that question, Tzamali designed her simulations to match the physical structure of gliomas, which are known to contain subpopulations of proliferative cells concentrated toward the center and invasive cells toward the edges. She ran separate simulations using two different kinds of invasive cells: one that is activated by low oxygen levels (i.e., hypoxia), and one that is invasive regardless of how much (or how little) oxygen it gets. And she varied the availability of oxygen to the tumors to mimic different levels of vascularization.

In work published in *PLoS One* in August 2014, Tzamali’s simulations recapitulated the physical structure of a real glioma, with proliferative cells clustered in a compact core and invasive cells forming long, finger-like extensions along the rim. They also showed that changing the oxygen levels in a tumor can have unintended, and potentially undesirable, consequences. For example, establishing normal oxygen levels at the outset prevented the hypoxia-driven invasive tumor cells from establishing dominance; but it also accelerated the rise to dominance of the other, more generally aggressive cell type. In general, Tzamali says, the model suggests that

drugs that target tumor vasculature might simply favor one phenotype in relation to another, or at best change the rate at which

one clone overtakes its competitors.

Just as significantly, the model also predicted that invasive cells would appear at the outer edges of the simulated tumors at

densities too low for an MRI scan to detect. Since MRIs are commonly used to diagnose gliomas, this argues for performing multiple physical biopsies well beyond a tumor’s core in order to find any lurking killers, and against relying too much on drugs that only focus on the proliferative cells that tend to cluster closer in.

In the future, Tzamali and her colleagues plan to ramp up the complexity of their model, integrate more experimental data—and validate their results with lab animals.

### Complexity and Chemo

How tumors respond to chemotherapy also reveals the importance of cellular heterogeneity, according to a study by Polyak and an international team of researchers published in *Cell Reports* in February 2014. The team tracked changes in genotype (chromosome copy number), phenotype (four types with different proliferation and migration traits), and spatial coordinates for tens of thousands of tumor cells from 47 different breast-cancer patients who were given chemotherapy to reduce tumor size prior to surgery. They found that the patients with less pre-treatment diversity were more likely to have the tumor completely disappear, leaving nothing behind for the post-chemo analysis. For tumors that showed no or only partial response to chemotherapy, there was very little change in intra-tumor genetic diversity, but the frequencies of the different phenotypes changed, with the more-proliferative types typically declining—something that could happen because chemotherapy tends to target proliferative cells in particular, or because the cells are actually switching between one phenotype and another. Cells of similar phenotype also tended to cluster together after treatment, even when they

In general, Tzamali says, her team's model suggests that drugs that target tumor vasculature might simply favor one phenotype in relation to another, or at best change the rate at which one clone overtakes its competitors.

were genetically different.

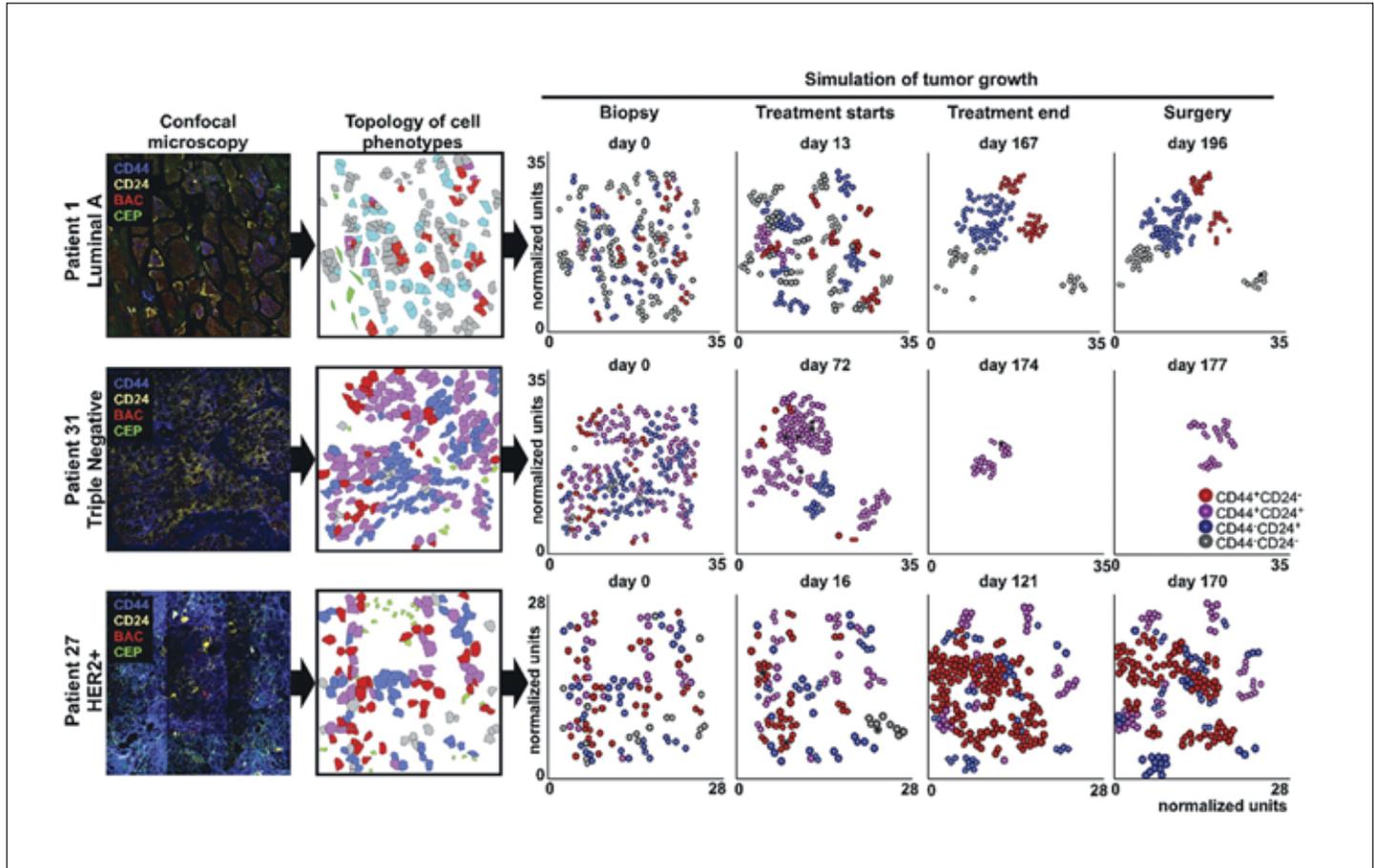
To understand what might be driving these growth patterns and evolutionary dynamics, Polyak’s colleague, the compu-

tational biologist **Franziska Michor, PhD**, developed an agent-based model to simulate the proliferation and death of each patient's tumor cells before, during, and after treatment. Because the researchers knew the actual phenotypes, proliferation rates,

giving the researchers some idea of what was driving changes at the level of basic biology. In the future, Polyak hopes that such a model could be used to predict the likelihood that a specific individual would respond well to a specific drug, giving doctors

kinase (MAPK) signal transduction network, which plays a key role in cell growth, as a series of Boolean logic gates (AND, OR, etc.).

The MAPK network also possesses a number of pathways that, when dysregu-



*Michor and her colleagues built these computer simulations of tumor growth using data from actual patients. They show changes in cellular phenotype and topology during chemotherapy. Reprinted from V Almendro et al., Inference of Tumor Evolution during Chemotherapy by Computational Modeling and In Situ Analysis of Genetic and Phenotypic Cellular Diversity, Cell Reports 6, 514–527, (2014).*

and spatial coordinates of the cells at the beginning and end of treatment, they could vary parameters for each patient-specific simulation to see if the model could account for what had actually occurred.

Intriguingly, they found that proliferation alone could not produce the phenotypic clustering they observed in the patients' tumors. To do that, the model also had to include phenotype switching. Permitting cellular migration, as occurs in metastatic breast cancer, also increased the amount of phenotype switching that was required to explain the patient data. "We knew those were possibilities," Polyak says of their findings, "but it wasn't really expected."

After phenotype switching and migration were added to the model, it proved capable of predicting changes in tumor-cell distributions on a patient-by-patient basis,

the power to run *in silico* clinical trials for their patients and helping them develop better treatment strategies in general.

### Divining the Logic of Cancer

Many researchers aim to tailor treatments to address cancer's troublesome heterogeneity. That's certainly what **David Basanta, PhD**, and **Aniruddha Datta, PhD**, are after, albeit through very different means.

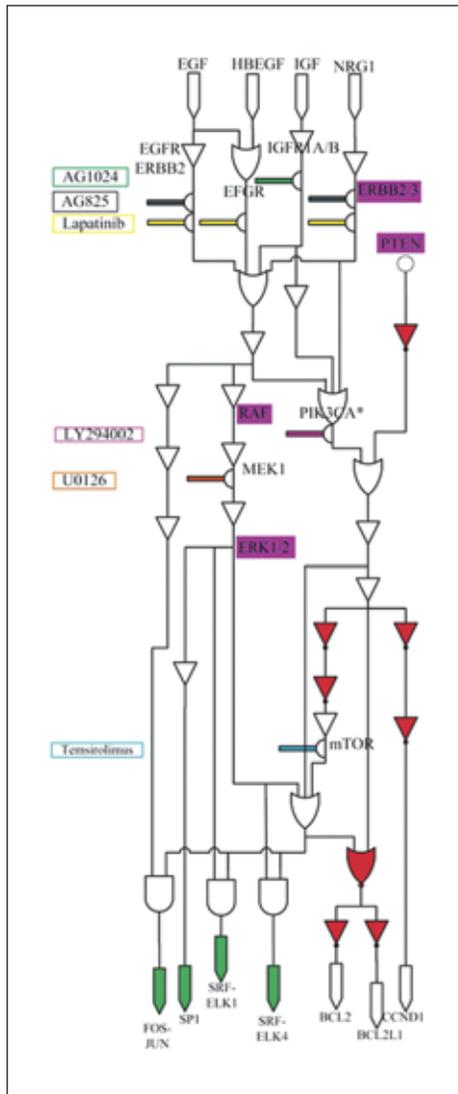
Datta, who directs the Center for Bioinformatics and Genomic Systems Engineering at Texas A&M University, has a background in control-systems engineering, and a fondness for modeling cancer with the kinds of Boolean networks that are used to study digital circuits. For example, with the help of his colleague, the computational biologist **Michael Bittner, PhD**, Datta has represented the mitogen-activated protein

lated, can lead to cancer. Normally, cell proliferation results when growth factors bind to a cell, triggering a signal transduction cascade that eventually activates the genes involved in proliferation. But mutations in the genes within the cascade can cause those cell-proliferation genes to be switched on permanently, or turn off the genes that would normally inhibit cell proliferation. Both kinds of mutations can be represented in a circuit by something that engineers call a "stuck-at fault," in which a particular logic gate is stuck in the on or off position. The effects of anti-cancer drugs that target these mutations can be superimposed on the same circuit.

In a real patient with a heterogeneous tumor, however, each clonal subpopulation will require its own representative circuit. The trick, according to Datta, is to determine exactly where all of the various faults are occurring; sort out the relative influence that each circuit exerts on the others and on the tumor as a whole; and predict the efficacy of different drug combinations given

all of the above.

Toward that end, Datta and doctoral student **Anwoy Kumar Mohanty** used statistical methods to develop a multilevel, hierarchical model of cancer tissue that can accommodate a number of possible networks as well as multiple drugs. They also developed an algorithm that uses probabilistic techniques (including Bayesian ones) to estimate the frequencies of a tumor's various subpopulations and their relative influence



**Datta and his colleagues constructed this Boolean network of the MAPK transduction network. Fault locations (i.e., genes stuck in the “on” or “off” position) are shown in purple boxes. Target locations of inhibitory drugs are indicated by solid colored rectangles, with the names of their molecular targets (PTEN, MEK1) printed nearby. The names of the drugs themselves (Lapatinib, U0126) appear in color-coded boxes on the left-hand side of the diagram. Reprinted with permission from A Mohanty and A Datta, A Model for Cancer Heterogeneity, IEEE Transactions on Biomedical Engineering, 61(3) (2014).**

on overall tumor behavior. The work, which is described in a paper that appeared in *IEEE Transactions on Biomedical Engineering* this past March, is still in the early stages: so far, the algorithm has only been validated using synthetic data and experimental data derived from normal, healthy cells. But Datta plans to begin testing the model against three different cancer cell lines within the next year. And he envisions a day when a patient's biopsy could be algorithmically analyzed to produce a drug regimen tailored to his or her particular tumor. “That’s my goal,” he says. “That’s the finish line.”

### From Genetic Mutations to Genetic Algorithms

It’s a finish line that Basanta hopes to cross as well. Basanta and his fellow modelers in the Cancer Evolutionary Dynamics research group at the Moffitt Cancer Center in Tampa, Florida, work with biologists and clinicians to understand the evolutionary dynamics of cancer progression and treatment resistance. Recently, Basanta, **Arturo Araujo, PhD, Jill Gallaher, PhD**, and other modelers in the Mathematical Oncology department used differential equations to model a heterogeneous, metastatic prostate cancer tumor located in the bone—a tumor whose cells could possess any permutation of three different mutations, for up to eight possible cancer cell phenotypes with unique proliferation rates and drug responses. The model, which is described in a paper that appeared last August in *Clinical and Experimental Metastasis*, includes two sets of equations: one representing the tumor and its various phenotypes, which grow and respond differently to five distinct therapies (e.g., hormone deprivation therapy, chemotherapy, experimental therapies targeting specific pathways); and one representing the bone microenvironment. The two equations are coupled so that the tumor affects the bone microenvironment, and vice versa.

Basanta and Araujo fed the model with data drawn from the literature, and from lab experiments conducted by a team of Moffitt biologists led by **Conor Lynch, PhD**. And they looked to clinicians at the center for information on how particular treatments were applied to real patients—and how those treatments affected cancer cell growth rates and bone behavior. Once fully parameterized, the model was able to simulate a virtual patient; and each simulation could include a different tumor with its own par-

ticular mix of mutations and cell types.

The outputs of each patient-specific simulation were then fed into a genetic algorithm that produced 1,000 successive generations of treatment options using the five drugs in the model. With each generation, the algorithm dropped the worst performing treatments and kept the best, until it was left with only those that kept the cancer at bay the longest. For each patient, the algorithm came up with an optimal single-drug regimen, and another that used more than one therapy in a particular sequence. And it did it all pretty quickly: for a virtual patient with the most heterogeneous tumor possible, the algorithm arrived at the current real-world standard of care—continuous hormone deprivation therapy—in less than 15 minutes. It also yielded some surprising results. In a couple of cases, for example, the algorithm recommended sequential courses of only two of the five therapies. And as Basanta notes, the first therapy didn’t even have to kill as many cancer cells as possible; it just had to prime the tumor for the second one.

“It really defies logic,” Araujo says. “You’d think that if you threw everything you had against the tumor, it would work.”

“It really defies logic,” Araujo says. “You’d think that if you threw everything you had against the tumor, it would work.”

But, he explains, the tumor is evolving and developing resistance according to a genetic algorithm of its own. “How can you tackle that? How can you keep one step ahead of what the tumor’s going to do? The only way is to simulate how the tumor might evolve and anticipate all of its moves using the same genetic algorithm.”

Like Datta, Basanta and Araujo hope for a day when a patient can walk into the clinic, have his tumor sampled and analyzed, and receive a personalized treatment regimen based on its specific composition. “A lot of these decisions are being made in the dark,” Araujo says of the current approach to treating tumors that may contain multitudes. “But mathematics says there is a better way.” □