quantum of conductance $G_Q = e^2/\hbar$, where e is the electron charge, and $\hbar$ is Planck's constant. As the gate voltage is moved to negative values, the conductance increases again, because holes are attracted. Because there is no energy gap, it is not possible to deplete the carriers completely and drive the conductance to zero. New ways must then be found to make transistors.

An even more unusual phenomenon occurs in graphene: The Klein paradox allows a relativistic particle to pass freely through a tall barrier of great width (*10*), whereas an ordinary particle would bounce backward, like a baseball after it hits a wall. As an electron approaches a potential barrier created by a gate, or by charged traps in the substrate, the electron's kinetic energy is reduced to zero as it hits the barrier. But instead of backscattering the electron, as one would expect, the barrier converts the electron into a hole, which is then attracted to the barrier potential and moves forward freely. When the hole exits the far side of the barrier, it turns back into an electron. The Klein paradox allows electrons to move through a graphene layer as if it were ideal. Graphene samples are not perfect, but are divided into pools of electrons or holes. Despite these imperfections, the charge carriers can move ~ 0.3 μm at room temperature with little scattering (*5*). This allows one to make ballistic transistors from such a simply prepared material.

Electrons in graphene show strong quantum behavior, even at room temperature, due in part becaue of their freedom of motion. In the quantum Hall effect, the Hall conductance is quantized in integer multiples of G_Q. To observe this effect, high quality samples, strong magnetic fields, and low temperatures ~4 K are usually needed. The quantum Hall effect has been observed in graphene (*11*, *12*)—the quality of the data is very high, despite the simple methods used to produce the samples. It is interesting that the quantization of Hall conductance steps is different for graphene: They are spaced by $4G_Q$, with the lowest steps at $\pm 2G_Q$. These changes are caused by graphene's unusual band structure in which the electrons and holes travel at a constant speed,. Recently, the quantum Hall effect in graphene has been observed at room temperature (*13*), which demonstrates its potential for quantum devices.

The freedom of motion associated with the Klein paradox creates a problem: How does one confine charge carriers inside a device? A simple approach is to cut the graphene layer into the right shape, as for the quantum dot in the figure. Quantum confinement of electrons can also be used to control their motion. For example, a narrow ribbon of graphene (*14*) can effectively create an energy gap between the electron and hole bands, the magnitude of the gap being inversely proportional to the ribbon width. With an energy gap, one can then deplete the carrier concentrations with a barrier or gate, as with a conventional device.

A quantum dot created by etching a graphene layer is shown on the left in the figure. A pool of electrons is confined inside a graphene disc that is connected to its leads by two narrow constrictions; a graphene side gate is used to tune the dot charge. Because they are very narrow, the conductance of these constrictions can be reduced to values $<G_Q$, which allow the dot to trap individual electrons at low temperatures. In this regime, the quantum dot acts as a single-electron transistor, and its conductance shows a periodic set of peaks as the gate voltage is increased. By making very small graphene dots with sizes <100 nm, Pomonarenko *et al.* (*1*) were able to enter a regime where the conductance peaks are no longer periodic, but are instead controlled by the energy of the individual quantum states of electrons trapped inside. The measured energy distribution showed the electrons behaved as a chaotic quantum system (*7*), as one might expect for a dot that is not perfectly round. By etching the dots even smaller, they were able to achieve transistor operation at room temperature.

Graphene is an exciting new material with unusual properties that are promising for nanoelectronics. The carriers move freely, ignoring barriers created by imperfections, and they show quantum effects at room temperature. Through advances in fabrication and characterization building on those of Ponomarenko *et al.* (*1*), it may become possible to make quantum dots so small that they approach the molecular scale (see figure, right panel). The future should be very interesting.

References

1. L. A. Ponomarenko *et al.*, *Science* **320**, XXX (2008).
2. J. S. Bunch, Y. Yaish, M. Brink, K. Bolotin, P. L. McEuen, *Nano Lett.* **5**, 287 (2005).
3. F. Miao *et al.*, *Science* **317**, 1530 (2007).
4. C. Stampfer *et al.*, *Appl. Phys. Lett.* **92**, 012102 (2008).
5. A. K. Geim, K. S. Novoselov, *Nat. Mater.* **6**, 183 (2007).
6. P. R. Wallace, *Phys. Rev.* **71**, 622 (1947).
7. M. V. Berry, R. J. Mondragon, *Proc. R. Soc. London Ser. A* **412**, 53 (1987).
8. K. S. Novoselov *et al. Science* **306,** 666 (2004).
9. K. S. Novoselov, *Proc. Natl. Acad. Sci. U.S.A.* **102,** 10451 (2005).
10. M. I. Katsnelson, K. S. Novoselov, A. K. Geim, *Nat. Phys.* **2,** 620 (2006).
11. K. S. Novoselov *et al.*, *Nature* **438**, 197 (2005).
12. Y. Zhang, J. W. Tan, H. L. Stormer, P. Kim, *Nature* **438**, 201 (2005).
13. K. S. Novoselov *et al.*, *Science* **315**, 1379 (2007).
14. M. Y. Han, B. Ozyilmaz, Y. B. Zhang, P. Kim, *Phys. Rev. Lett.* **98**, 206805 (2007).

10.1126/science.1156936

MEDICINE

The Ultimate Model Organism

Atul J. Butte

A deeper understanding of disease requires a database of human traits and disease states that is integrated with molecular information.

This month, the scientific community celebrates the 25th anniversary of GenBank, the open access database of DNA sequences and the molecules they encode. Heralded as one of the earliest bioinformatics community projects, it has fueled our need to understand how this information can be linked to physiology and disease. Since then, biocomputational, informatics, and statistical methods have been used to relate sequences and molecules to diseases. But as highlighted in meetings such as last month's Summit on Translational Bioinformatics (*1*), the same high-bandwidth measurement style that has accelerated the molecular and genetic study of disease must be practiced in physiology if we are to gain a deeper understanding of normal and impaired health.

Within the last 5 years, systematic studies on the commonalities (*2*) and differences (*3*) across diseases have shown that particular biological signaling pathways and modules share similar properties. Other studies have shown that diseases that resemble each other can share genes with variants (*4*, *5*) or share genes coding for proteins that interact with each other (*6*). So many diseases have now been studied that publicly available data can be used to find genes with common changes in expression for each condition (*7*).

The difficulty with interpreting such analyses lies with how diseases are defined. The definition of a disease is often specified

Stanford Center for Biomedical Informatics Research, Department of Medicine and Department of Pediatrics, Stanford University School of Medicine, Stanford, CA 94305 USA, and Lucile Packard Children's Hospital, Palo Alto, CA 94304, USA. E-mail: abutte@stanford.edu

by a particular knowledge base and is thus subject to limitations and biases. For example, a network built from a knowledge base of monogenic diseases (those associated with a single gene) may not be generalizable to more common diseases caused by multiple genes (*5*). More recently described diseases, such as sudden infant death syndrome, may be less well characterized, so searches for gene variants by matching syndromes through clinical descriptions could yield false-negative predictions (*6*). In some studies, gene expression or genotyping samples are studied based on clinical disease labeling by physicians, and thus, genes found to be "associated" with a condition do not yet fully explain the observed traits present in a disease. Moreover, there is a growing movement toward direct-to-consumer testing, with the promise that consumer-provided DNA samples could be used for genome-wide association studies. It remains to be seen how such samples can be analyzed when the "assignment" of phenotype and disease is provided by a consumer.

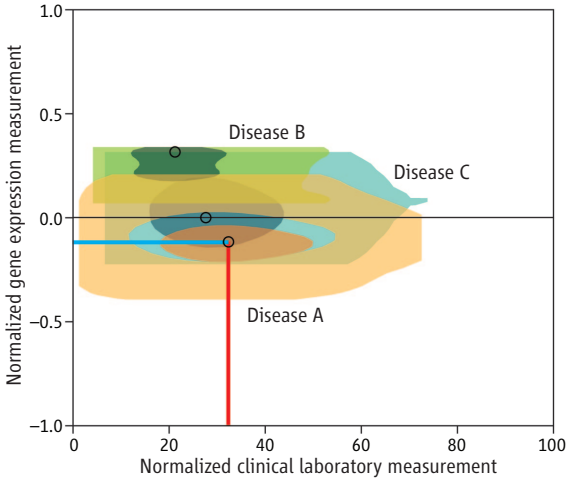
On the other hand, parallel measurements of physiological variables have been successfully linked to genetic markers in animal models of diseases, such as hypertension (*8*). These efforts have driven the organization of efforts such as the Physiome Project (*9*), an international collaboration to model the human body through computational methods that integrate biochemical, biophysical, and anatomical information about cells, tissues, and organs. There have also been calls for a Human Phenome Project (*10*), whose goal is to establish databases of phenotypes that are associated with physiology, to determine their relation with genes and proteins. Data for some complex human physiological traits are already publicly available for analysis at resources such as PhysioNet (*11*).

Yet, the current approach for defining phenotypes for molecular discovery is not adequate, and therefore, doesn't optimize the use of physiological data. Phenotypes—traits ranging from height and weight to glucose metabolism, predisposition to disease, and behavioral characteristics—and their differences between individuals can be due to environmental influences and/or genetic variation. One solution for defining richer phenotypes is to take advantage of clinical measurements, which are born from physiological measurements. Enormous numbers of clinical tests are performed each year, and are increasingly being captured in electronic health records, along with patient interventions (medications or procedures). These kinds of data could be used to answer basic biological questions (*12*, *13*). Mathematical arrays of such data have already been assembled from hospital-based clinical measurements or epidemiological information and have successfully identified biomarkers for human maturation and aging (*14*, *15*). Connections between clinical findings and molecular measurements can also now be tested across a large set of findings and molecules. For example, gene expression profiles of individual liver cancer samples have been predicted by prior radiological findings on abdominal computed tomography scan (*16*).

How can we take advantage of the petabytes of clinical measurements on patients for whom genetic or genomic measurements may not yet have been obtained? The same broad consideration across diseases used successfully in molecular studies could also be applied to clinical measurements. For example, suppose three diseases are separately considered by a quantitative clinical laboratory test measurement (obtained from an electronic health record), and a gene expression measurement, (from a public repository of gene expression data) (see the figure).Within a disease, the distribution of gene and clinical measurements can be shown, but whether the clinical and gene measurements correlate cannot be determined, as the measurements were not taken from the same patients. But trends might be observed across the three diseases. For instance, as a disease shows more or less of a clinical measurement in patients, microarray samples of the disease may show more or less expression of a particular gene. Thus, associations could be discovered between molecular and clinical measurements, even when these measurements are not made using the same samples or patients. Instead of studying samples or patients as data points in the traditional reductionist manner, one could study and plot diseases.

But there are challenges to using clinical data as physiological measurements. Access issues to patients' private health information can dissuade basic researchers from using clinical measurements. Although patient data could be deidentified and patients approached for informed consent, much clinical data exists as documents that are difficult to deidentify and/or sift through using automated processes. Even as these challenges are addressed, purely numerical quantitative clinicians measurements could be used to start, as these are the easiest to deidentify and analyze.

Whereas there are public international repositories for many molecular measurements, we do not yet have an equivalent for deidentified clinical measurements. There are multiple reasons for this. The fear that personal medical information could be inappropriately released is a powerful disincentive for sharing. Clinical data may also be viewed by clinical and hospitals as a "trade secret," and only recently are data on performance and quality being published. This fear could be averted if health care networks pooled deidentified data sets, thus deidentifying the source of care as well. Clinical researchers are also justifiably protective of the resources they create and might fear missing a discovery within their own patient cohort. Availability agreements could address retention of rights, intellectual property, and publication embargoes. Instead of viewing data availability as a disadvantage, clinical researchers and institutions should be encouraged to look at the success of resources such as GenBank to see how the public availability of deidentified data can yield many more discoveries when shared.



Information intersection. Three diseases may be separately considered by a quantitative clinical laboratory test measurement and a gene expression measurement (from a public repository of gene expression). Associations can be discovered between molecular and clinical measurements, even when these measurements are not made using the same samples or patients. For example, Disease A, when studied across all patients and time points, shows a high average level of a clinical test (red line), and a low level of a gene (blue line). The distribution of gene and clinical measurements are shown by sampling from both independent data sets (colored regions). The trend across the three diseases shown is that as a disease shows less of a clinical measurement in patients, it shows more expression of a particular gene.

CREDIT: DAVID CHEN

A population of well-supported and trained scientists and physicians must be nurtured to relate the enormity of physiological and clinical measurements to molecular measurements. The multiscale models of health they build will finally yield an understanding of disease that is more than just the sum of its parts.

References

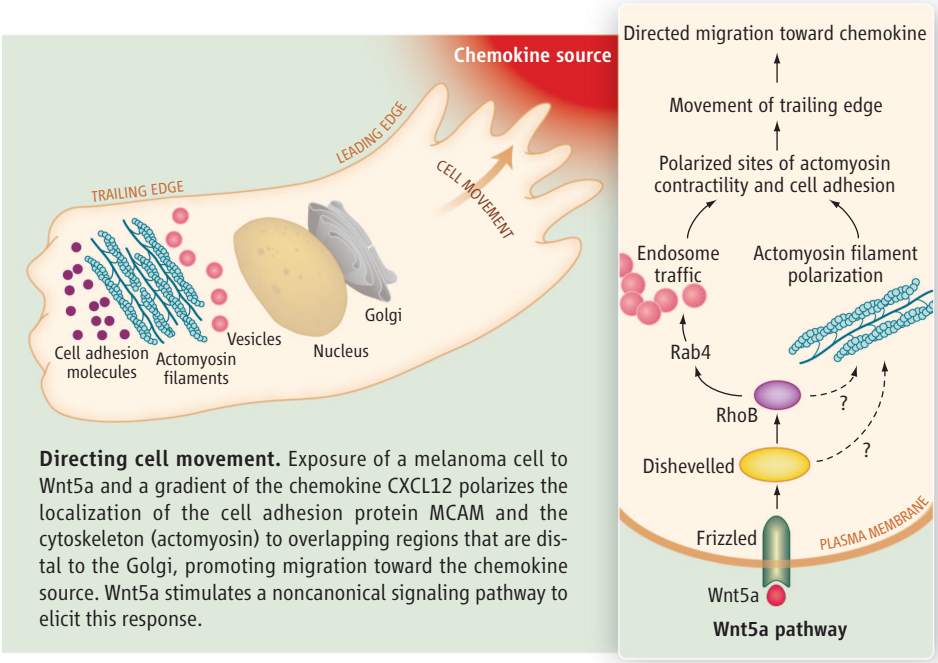
1. www.amia.org/meetings/stb08/
2. D. R. Rhodes *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 9309 (2004).
3. E. Segal, N. Friedman, D. Koller, A. Regev, *Nat. Genet.* **36**, 1090 (2004).
4. The Wellcome Trust Case Control Consortium, *Nature* **447**, 661 (2007).
5. K. I. Goh *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 8685 (2007).
6. K. Lage *et al.*, *Nat. Biotechnol.* **25**, 309 (2007).
7. A. J. Butte, I. S. Kohane, *Nat. Biotechnol.* **24**, 55 (2006).
8. M. Stoll *et al.*, *Science* **294**, 1723 (2001).
9. P. J. Hunter, T. K. Borg, *Nat. Rev. Mol. Cell Biol.* **4**, 237 (2003).
10. N. Freimer, C. Sabatti, *Nat. Genet.* **34**, 15 (2003).
11. http://www.physionet.org/
12. N. S. Sung *et al.*, *JAMA* **289**, 1278 (2003).
13. P. R. Payne *et al.*, *J. Investig. Med.* **53**, 192 (2005).
14. D. P. Chen *et al.*, *Pac. Symp. Biocomput.* 2008, 243 (2008).
15. A. Fliss, M. Ragolsky, E. Rubin, in *Summit on Translational Bioinformatics Proceedings*, San Francisco, CA, 10 to 12 March 2008 (AMIA, Bethesda, MD, ISCB, La Jolla, CA, 2008), p. 11.
16. E. Segal *et al.*, *Nat. Biotechnol.* **25**, 675 (2007).

CELL SIGNALING

Wnt Moves Beyond the Canon

A secreted protein associated with the metastasis of a melanoma harnesses cell adhesion proteins and the cytoskeleton to direct cell motility.

Bruce Bowerman



Directing cell movement. Exposure of a melanoma cell to Wnt5a and a gradient of the chemokine CXCL12 polarizes the localization of the cell adhesion protein MCAM and the cytoskeleton (actomyosin) to overlapping regions that are distal to the Golgi, promoting migration toward the chemokine source. Wnt5a stimulates a noncanonical signaling pathway to elicit this response.

The cellular signaling pathways that respond to Wnts, a highly conserved family of secreted proteins, control numerous and different animal developmental processes, particularly those that govern cell fate and patterning. Often, components of Wnt signaling pathways malfunction and foster deadly cancers (*1*). Although intensely studied, much about Wnt signaling remains enigmatic. Whereas most signaling events that target gene expression in response to Wnt are referred to as canonical pathways (the more commonly observed sequence of molecular events), noncanonical Wnt signaling often targets the cytoskeleton. On page xxx of this issue, Witze *et al.* (*2*) advance our understanding of a noncanonical Wnt signaling pathway (Wnt5a) that promotes the invasive migration of melanoma cancer cells. Their findings advance our understanding of cancer metastasis and highlight the diversity of mechanisms through which Wnt signals influence cells.

The canonical Wnt signaling pathway was deduced from genetic studies in the fruit fly *Drosophila melanogaster* and is widely conserved in other animals (*1*). Wnt binds to Frizzled, its transmembrane receptor, prompting cytosolic Dishevelled proteins to prevent the otherwise constitutive proteolytic destruction of another cytosolic protein, β-catenin. Stabilized β-catenin then transits to the nucleus, where it converts transcriptional repressors called T cell factors into activators, driving the expression of target genes that specify cell fate. By contrast, one cannot so easily summarize noncanonical Wnt signaling. The best-understood example was discovered in *Drosophila*, where it polarizes epithelial cells along a common axis (*3*, *4*). This planar cell polarity requires the canonical Wnt signaling components Frizzled and Dishevelled as well as proteins that do not participate in the canonical pathway. Although related signaling pathways that respond to Wnt influence cell migration during early vertebrate embryogenesis (gastrulation), the full extent of their conservation remains unclear. More recently, genetic studies in the nematode Caenorhabditis elegans have identified noncanonical Wnt signaling pathways that bear little or no resemblance to the one that controls planar cell polarity (*5*–*9*). The only unifying theme is that all of these examples somehow differ from the canon.

Nevertheless, a shared feature of many noncanonical Wnt pathways is their targeting of the cytoskeleton, with important implications for cancer. For example, expression of noncanonical Wnt5a correlates with metastatic melanoma invasiveness in humans, and exposure to Wnt5a, but not other Wnts, promotes invasiveness in melanoma cell lines (*10*). Witze *et al.* investigated how Wnt5a polarizes the cytoskeleton to promote directional motility in cultured melanoma cells. They show that this response is permissive, requiring an independent directional cue (a chemokine). Dispersed cells within a chemokine gradient require exposure to Wnt5a for polarization and migration, but exposure to Wnt5a alone has little effect.

To understand this noncanonical and permissive response, Witze et al. used fixed- and live-cell imaging to examine the localization and movement of proteins within single cells shortly after exposure to Wnt5a. They examined several proteins, including a melanoma cell adhesion molecule previously implicated in metastatic mela-noma (*11*), a Frizzled

Institute of Molecular Biology, University of Oregon, Eugene, OR 97403, USA. E-mail: bbowerman@molbio.uoregon.edu

CREDIT: ADAPTED FROM BRUCE BOWERMAN