Lee Hood

Lee Hood outlines his vision of personalized medicine for the next 10 years.

As a pioneer in new technologies that affect genome science and systems biology, Lee Hood of the Institute for Systems Biology in Seattle has set his sights on transforming the practice of medicine. Here he provides his vision of how new approaches will change the way we view health and disease.

What is P4 medicine?

Lee Hood: P4 medicine stands for medicine that is predictive, preventive, personalized and participatory, but the basic idea is that P4 medicine looks at medicine from an informational point of view. In the next ten years, patients are likely to be surrounded by virtual clouds of billions of data points and we will use IT [information technology] to reduce this data dimensionality to form simple hypotheses about health and disease. Individuals differ by 6 million nucleotides on average and are exposed to different environmental stimuli. The old population-based methods gave you bell curves for traits and if you were on either tail of the curve, you were sick. In contrast, with P4 medicine one treats each person individually and not as a part of a group.

Can you describe this approach in more detail?

LH: There are basically two major types of biological information, the digital information of the genome and environmental signals arising from outside the genome. The information structures that connect these two types of information with the phenotype in health and disease are the biological networks that capture, integrate and modulate information and then pass it off to molecular machines that execute the information. The disease-perturbed dynamics of these networks lie at the heart of understanding disease mechanisms. For example, in our study on prion disease in mice (*Mol. Syst. Biol.* **5**, 252, 2009), we identified about 300 disease-perturbed genes that are involved in four major and six minor biological networks. The dynamics of these networks explained virtually every aspect of this neurodegenerative disease. The networks became disease-perturbed in a sequential fashion. If you want to think about early diagnosis and therapy, you should focus on the first disease-perturbed network both to identify biologically relevant molecules secreted into the blood for diagnosis and to find drugs that can reengineer the disease-perturbed network to make it behave normally, thus abrogating the progression of the disease. We also identified organ-specific blood markers, that make blood a window through which we can distinguish health from disease. All organs have specific markers that are secreted into the blood and constitute a molecular fingerprint that reports, by concentration changes, shifts from a normal to a disease-perturbed state.

Why do so many published biomarkers never make it to the clinic?

LH: Success can be improved by rationally choosing biomarkers rather than doing a random shotgun search to detect biomarker changes between disease and health. I would guess that 99% of those biomarkers that are discovered by random searches are not going to be very useful. Most likely they just represent biological noise.

> "The real issue with large data sets is that the data sets have enormous signal-to-noise problems."

What technologies are needed to make P4 medicine a reality?

LH: A key advance is that we are now able to do complete genome sequencing of families to identify genes that are involved in simple genetic diseases. We are now beginning to apply genome sequencing to families with more complicated genetic diseases. Those studies look promising as well. In addition, the genomes of individuals will increasingly provide insights into the future health trajectory of the individual. A second transformative technology is the development of targeted proteomic assays for essentially all human proteins that my institute has recently announced. A third area is the use of microfluidic chips to be able to quantify not tens of proteins, but eventually thousands of proteins from a droplet of blood in just a few minutes. Making these devices is relatively straightforward, except for one thing: we need better protein-capture agents, such as aptamers or peptide binders, for protein assays. Fourth, single-cell analysis will be incredibly important for assessing distinct quantized populations of cells. The idea that you can learn a lot about biology by looking at the individual cell rather than averaging populations of cells will provide fundamental new insights into cancer, development and physiology.

How will this translate into a shift from disease treatment to disease prevention?

LH: Systems thinking about disease gives you an entirely new strategy for identifying drug targets. The drug companies are good at making drugs once they have the target, but they are really bad in choosing the target. If you understand the nature of disease-perturbed networks, you can reengineer disease-perturbed networks to be normal. In most cases, this is clearly going to take multiple drugs and we will need good biomarkers to detect the early changes in these networks. It is a short step to design drugs that prevent potentially disease-perturbed networks (predicted from your genome sequence) from ever becoming disease-perturbed—true preventive drugs.

What will this mean for the provision of healthcare in the future?

LH: The focus of healthcare will shift over the next ten years from disease to wellness. We are developing metrics for assessing an individual's wellness. There will be a wellness industry that in time could dwarf the healthcare industry. Medicine will also be focused entirely on the individual in the future. We will all have the equivalent of iPods that will be recording enormous amounts of personal data and transmitting it to servers for analyses that will monitor your wellness status and report developments that are a cause for alarm by sending you a signal, such as "Slow down on eating."

Data analysis tends to lag behind data generation in biology—do you see this changing?

LH: I think the real issue with large data sets is that the data sets have enormous signal-to-noise problems. If you measure a given phenotype response, it could be the sum of a number of different biological phenomena. If you are only interested in one of them, you have got to be able to subtract away the others. Learning to do that biological subtraction is one of the grand challenges in P4 medicine. That is the reason the genome-wide association studies have only been marginally effective. The signal-to-noise issues are overwhelming.