

KARIN A. REMINGTON, PHD



It's not what you know, but what you don't know...

Bioinformatics and computational biology have told us a lot about biology—primarily that we know so little. Advances have led to many more unanswered questions, suggesting we know less and less all the time. For example, during the half dozen years since the human genome sequence was first published, the biomedical research community has been pouring through a variety of computational annotations on the human genome, including regulatory and protein coding sequences, evolutionary conserved regions, and a broad collection of potentially significant single-nucleotide polymorphisms (SNPs). Still, we have a very poor understanding of even these basic genetic elements, let alone the potential role of other new and yet to be discovered players. Similarly, modeling and computational simulations have advanced our understanding of the structural and chemical properties of the DNA molecule, revealing new significance for genomic regions long dismissed as “junk.” What undiscovered control mechanisms might be hidden in the non-coding regions of the genome? Just this small sampling of recent findings aided by computational approaches reveals tantalizing new awareness of all that we don't know, and surprising new perspectives:

■ A bioinformatics analysis, coupled with hypothesis driven lab work, led **Paul Pease, Oren Levy and Jerod Ptacin** at University of California, Berkeley to the discovery that specific short, asymmetric DNA sequences control the bi-directionality of the DNA translocase FtsK, a molecular motor involved in bacterial chromosome dynamics (*Science*, 2005).

■ The ENCODE pilot project's recently published results (*Nature*, 2007), relying heavily on computational analysis and comparative sequence analysis, have demonstrated that even our most basic understanding of transcription must be revisited.

■ **Eran Segal** and a team from the Weizman Institute and Northwestern University constructed a model of nucleosome-DNA interactions in yeast with a validation strategy that revealed a genomic encoding of the nucleosome organization (*Nature*, 2006).

■ Computational simulations by **Jory Ruscio** and **Alexey Onufriev** at Virginia Tech have been targeted at the role of flexibility in nucleosomal DNA packaging, generating new experimental directions of inquiry that could elucidate nucleosome dynamics (*Biophysical Journal*, 2006).

■ Work in **Carlos Bustamante's** lab at UC, Berkeley led to the development of new mathematical models of the DNA double helix from a mechanical perspective, inspired by their experimental observations of DNA winding and stretching properties that confounded earlier models and intuitions (*Nature*, 2006).

As Director of Bioinformatics and Computational Biology at the National Institute of General Medical Sciences, which has been privileged to fund many of these studies, I can sense a growing awareness within NIH of the extraordinary potential of asking new questions and revisiting long held assumptions of what we “know.” The Institutes' recently announced, and aptly named, Exceptional Unconventional Research Enabling Knowledge Acceleration RFA (EUREKA, <http://grants.nih.gov/grants/guide/rfa-files/RFA-GM-08-002.html>), reflects this sense of promise. In fact, the new world of science, where advances increasingly require large interdisciplinary teams, provides the perfect opportunity to thrive on what we don't know. Most of us realize, from first hand experience, the difficulty in establishing communication across disciplinary boundaries. This difficulty, though, ought to be its very strength. It gives us a rare opportunity to ask colleagues about what we don't know and what doesn't make sense to us. In turn, if we thoughtfully consider our responses to similar questions posed to us, we will allow ourselves to be enlightened by the questions that challenge our assumptions of what we know, and energized by the exciting paths to discovery that may result. □

DETAILS

Karin A. Remington, PhD
 Director, Center for Bioinformatics
 and Computational Biology
 National Institute of General Medical Sciences
 National Institutes of Health