EDITORIAL

Opportunities—And Hard Work—Ahead

Robert N. Hoover, Stephen J. Chanock

Correspondence to: Robert N. Hoover, MD, ScD, NCI-Shady Grove, Room 7E412, 9609m Medical Center Drive, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD 20890 (e-mail: hoover@mail.nih.gov) or Stephen J. Chanock, MD, NCI-Shady Grove, Room 7E412, 9609m Medical Center Drive, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD 20890 (e-mail: chanocks@mail.nih.gov).

In this issue of the Journal, Dr. Wild and his colleagues issue a clarion call to epidemiologists, other prevention-oriented disciplines, and cancer research funding sources to embrace the extraordinary opportunities in cancer etiology and prevention afforded by the recent "omic" revolution (1). New approaches to understanding underlying mechanisms of carcinogenesis in human populations will undoubtedly lead to an unprecedented expansion of our discovery of potentially preventable carcinogenic exposures, as well as opportunities to interdict the adverse biological effects of those exposures not practically preventable. The primary molecular tools to gain such insights include a range of "omics"—genome, epigenome, transcriptome, proteome, metabolome, and exposome. The authors point out the paradigm shift in our understanding of genetic susceptibility to cancer, one that occurred in less than a decade through agnostic screening of common variants over the entire genome. Indeed, using these approaches, we have gone from knowing five established common variants related to cancer risks to over 450 and still counting (2). Moreover, the identification of new susceptibility alleles has yielded a more complex understanding of the underlying genetic architecture of genetic susceptibility. The authors predict a similar application of other omics could extend this phase of discovery, but note that to make sense of new biomarkers, infrastructure will be needed to realize these lofty goals. They call for the establishment of large cohorts rich in biospecimens, the formation of multi-institutional and inter-national consortia, the development of more high-throughput technologies, and commitment to more multidisciplinary teams.

We couldn't agree more with the exciting prospects and new frontiers that lie ahead for discovery of the biologic basis for cancer causation and its possible prevention. We also agree that while the current enthusiasm for "precision" medicine is appropriate, extending the "precision" paradigm to prevention could have a greater overall impact on health, particularly in reducing the anticipated enormous burden of cancer worldwide.

We would, however, urge approaching these prospects with a healthy dose of realistic caution. To paraphrase Thomas Edison, these “opportunities…are dressed in overalls and look like work.” There are indeed formidable challenges ahead, ones that are likely to require considerable resources and time to overcome. The successes of mapping the human genome and leveraging it to discover germ-line risk variants and driver somatic mutations represent the beginning of a process, one in which the discoveries require further investigation to understand the biological underpinnings of the genetics, which in turn can provide the foundation for precision medicine and prevention. With the bulk of the newly discovered risk variants mapping to non-coding regions (3), it is likely to be some time before we uncover the underlying mechanisms. We now have a hint of how cancer arises because of disruption of the regulation of pathways and networks over years.

The early success of genomics does not imply success with the other omics will occur over a comparably short time interval. For the genome, we knew the approximate number of base pairs and could measure them with reliability. We learned the variants’ relationships to each other, established by millennia of assortative mating, and could control for this at a fine level. What we measured was always the same for a specific person, not influenced by time, eating, exercise, or any other exogenous or endogenous exposures. Essentially none of this is true for the other omics. Not only is accuracy and relevant measurement difficult, for most of these omics, we do not even know how many there are. Even for the genome, the debate rages on as to how many noncoding RNA species exist and what fraction have biological consequences (4). For metabolomics, we have no reliable estimate of how many small molecules there are. Currently, using mass spectroscopy, several hundreds can be measured reliably; thousands of as yet unidentified “peaks” can also be observed, but would be difficult to replicate or validate (5). Certainly there are hundreds of thousands of metabolites still unmeasured. Moreover, measurement of metabolites often exhibits major within person variability, with levels influenced by age, recent diet, activity, time of day, health status, and many other factors (6). And if the operative molecules are actually

Received: October 26, 2014; Accepted: October 28, 2014
© Published by Oxford University Press 2014.
metabolites of metabolites, then the measurement challenges are substantially increased (7). For each of the nongenetic omics, we are also challenged with the question of how well assays of available stable biospecimens (eg, blood and urine) reflect the relative concentrations in the tissues of interest.

A celebrated advantage of mapping cancer susceptibility alleles or somatic driver genes has been their “apparent” immu-
nity from the sources of bias and confounding that challenge most epidemiologic investigations of environmental and lifestyle risk factors. Once confounding by population structure was incor-
porated (accomplished with the genetic data), no other sources of confounding appeared to be important for discovery of the main effect, namely susceptibility to a type of cancer. This allowed for the creation of large databases, including case series, generic control groups, studies of high and low epidemiologic quality, and those containing no other information than genetics. The resulting increase in statistical power also enabled the identification of associations with very low relative risks, including below 1.1, occurring consistently across studies, and being accepted as likely causal. With the other omics being substantially influenced by exogenous and endogenous exposures, potential bias and confounding become relevant and likely limit credible analyses to higher quality studies with more available information.

As noted, these caveats are not meant to undermine the enthusiasm for the remarkable opportunities that the new molecular science and technology are providing. However, we do think that it is important to recognize that the revolutionary multiplicity and importance of new leads coming from genomics is a result of agnostic screening of the entire genome, allowing it to tell us which genetic regions, harboring candidate genes, are important, rather than relying on our own speculations, which led to decades of failure pursuing “candidate genes.” We do not yet have this capability for most of the other omics, and it will take a considerable methodological effort in terms of commitment, time, and resources to reach this point.

In the near future, we are optimistic that we will be able to interrogate important portions of each of the other omics that could lead to new actionable findings in etiology and prevention. Indeed, it could be that we will be fortunate and have multiple preventive avenues to pursue. If we are so lucky, it will fan the flames of an old controversy, how and when to initiate primary and secondary prevention strategies that don’t involve the investment of randomized clinical trials of prevention. If we are to counter the upsurge in cancer globally, then we have to begin to develop strategies based on the synthesis of observational data.

As others have emphasized (8), we actually do know the causes of a substantial proportion of the overall cancer burden in the population, notably including tobacco use and obesity, and we should be refining and accelerating implementation of current prevention strategies targeted at reducing their impact. While clearly correct, it is also likely that using new molecular science to understand the underlying mechanisms of established carcinogens, including tobacco addiction and the specific pathways responsible for obesity-related cancers, we are likely to uncover preventive opportunities that complement current behavioral and societal interventions that have their own limitations.

Note
The authors have no conflicts of interest to declare.

References
logic and functional implications of genome-wide associa-
4. The ENCODE Project Consortium. An integrated encyclo-
6. Sampson JN, Boca SM, Shu XO, et al. Metabolomics in epide-