Testimony of

Dr David Relman

May 11, 2004

STATEMENT OF DAVID A. RELMAN
ASSOCIATE PROFESSOR, STANFORD UNIVERSITY

BEFORE THE

COMMITTEE ON THE JUDICIARY
SUBCOMMITTEE ON TERRORISM, TECHNOLOGY & HOMELAND SECURITY
UNITED STATES SENATE

CONCERNING

EARLY RECOGNITION AND MANAGEMENT OF INFECTIOUS DISEASES: HARNESSING THE HUMAN GENOME IN AN ERA OF NEW AND EMERGING BIOLOGICAL THREATS

PRESENTED ON

MAY 11, 2004

STATEMENT OF DAVID A. RELMAN, MD
ASSOCIATE PROFESSOR OF MICROBIOLOGY & IMMUNOLOGY,
AND OF MEDICINE
STANFORD UNIVERSITY
BEFORE THE
COMMITTEE ON THE JUDICIARY
UNITED STATES SENATE

May 11, 2004

Good morning Chairman Kyl and Members of the Committee. Thank you for providing me, and others at this table, the opportunity to appear before you today and discuss a new and rapidly developing area of science that may revolutionize the way we detect and manage disease caused by emerging and unanticipated infectious agents. I am an infectious diseases clinician and a researcher based at an academic medical center whose interests are in the discovery of novel pathogens, the study of human microbial ecology and the use of genomic techniques for these purposes. I am also a member of the Board of Directors of the Infectious Diseases Society of America.

Introduction: Persistent difficulties in the diagnosis of infectious diseases

Disease-causing microorganisms were recognized and identified more than one hundred years ago. Methods for cultivating bacteria and viruses have been available since the early or mid twentieth century, and are widely established in clinical settings. These statements would therefore understandably lead one to believe that in today's clinical work-place the ability to diagnose human disease caused by microbial agents is well-established and reliable. However, this is not the case. Carefully-performed studies of various clinical syndromes suggestive of an infectious etiology fail to identify a possible microbial causal agent in more than 50% of cases. An investigation by CDC and academic scientists of unexplained critical illnesses and deaths with features strongly suggestive of infection found that all of the cases subsequently diagnosed using state-of-the-art laboratory methods were caused by well-known

microbial agents which failed to be detected with routinely-available tests. Furthermore, even when infectious disease agents are successfully identified, the answer often arrives late in the course of the illness, after the patient has either recovered or succumbed.

There are multiple reasons for this unsatisfactory state of affairs, including a heavy reliance on cultivation techniques (these techniques are often insensitive), difficulties in obtaining clinical specimens from an appropriate patient site and at an appropriate time (such that one can expect the agent to be present in the specimen), and the inherent delays and lack of specificity in many of today's diagnostic approaches. It is important to emphasize that the clinical features of different serious systemic infections often look identical to both the health care provider and the patient in the earliest phases of the disease. The consequences of this poor diagnostic capability in infectious diseases are profound. Clinicians are often compelled to institute antibiotic treatment in a broad, empiric manner without a definitive diagnosis, despite the fact that in many cases, a different or more specific drug would have indicated had the infectious agent been known. In addition, for many cases, antibiotics are inappropriate altogether because the causative agent is a virus, or the condition is not caused by an infectious agent at all. As a result, patients suffer from delayed or sub-optimal treatment, and the prevalence of antibiotic-resistant bacteria grows inexorably.

How the unraveling of the human genome sequence has presented a set of unprecedented opportunities

The best clinicians are known for their ability to listen carefully to the patient and extract subtle clues as to the correct diagnosis. We are now in a position to be able to "listen" to patients in a manner that is far more sensitive and comprehensive than any method has previously available. This capability has been enabled by the deciphering of the human genome sequence. One new approach that I wish to describe to you is based on highly-parallel measurements of human gene expression and the analysis of these patterns of expression, using a tool called a DNA microarray.

Although we all share the same set of genes, there is a great deal of variability over time within any individual, and between individuals, in how we use our genes. Each gene provides the blueprint for production of a protein. At times when a protein is needed, the corresponding gene is "activated" to make (express) copies of its own specific messenger RNA (mRNA), which then serves as the template for protein synthesis. When the protein is no longer needed, the gene is "repressed", its specific mRNA is no longer made, and the abundance of this mRNA decreases. DNA microarrays are high density arrays of DNA probes that are each specific for a different gene and its mRNA. These probes capture their target mRNA in a specific and reliable manner. If the starting pool of mRNA in a specimen is labeled with a fluorescent dye, the captured mRNA molecules can be detected bound to their matching probe on the microarray using a laser. By scanning a DNA microarray with this laser one can quickly measure the abundance of mRNA bound to each probe, and hence the amount of mRNA present in the specimen for each human gene. Thus, a human DNA microarray can be used to monitor the degree to which each and every human gene was active (being expressed) within any given patient specimen at the moment the specimen was obtained. Genes are activated and repressed on a minute-to-minute basis in response to environmental cues. Each gene responds in a unique manner to different stimuli. One can view the human genome (our complete collection of genes) as a living, dynamic entity! If we could learn to recognize the various patterns of gene expression associated with specific kinds of stimuli, such as disease caused by different microbial disease-causing agents, we might be able to "read" this form of host response and diagnose infectious diseases at the earliest stages of development.

The analysis of human gene expression patterns in clinical specimens ("gene expression profiling") began about 5-7 years ago. The vast majority of these analyses, so far, have focussed on the study of human cancers, and the results have been impressive. To summarize a growing body of published data, expression profiling has been used successfully to diagnose various forms of cancer that could not be otherwise diagnosed, and to predict patient clinical outcome (favorable or unfavorable course, response to therapy, etc) when these predictions could not have otherwise be made. In addition, novel, critical mechanisms of cancer have been deduced from these expression patterns, which in turn have led to the development of novel effective therapies. It should be mentioned that the same kinds of new insights have been gleaned from a different type of genomic analysis: instead of analyzing patterns of mRNA abundance, patterns of protein expression have been measured directly. Additional details of this work will be presented by the next discussant in this session.

The application of gene expression profiling to infectious diseases, and to related diseases caused by biological agents (e.g., toxins), is still in an early stage of exploration. However, early findings are encouraging. It appears that

these profiles differ in the blood of different sick patients as a function of the specific disease-causing agent, thus, indicating that humans do discriminate between different infectious agents at the level of gene expression patterns. Gene expression profiles in the blood of healthy persons also differ between individual but much less so than in disease; however, these lesser degrees of variation may be sufficient to reveal important physiological differences between individuals. At the same time, there are inadequate data with which to determine the level of resolution offered by these patterns in the identification of a causal infectious agent. And while we expect that patient outcome might well be predicted one day based on these patterns, we do not yet have enough experience to identify and recognize prognostic "signatures" in a wide spectrum of different humans.

What is needed?

One of the most important needs for advancing this technology application is a much more extensive, well-annotated set of expression profile data from diverse individuals at varying stages of different infectious diseases, and after exposure to a wide variety of different biological agents. While it may not be possible to collect many (or any) specimens from humans after exposure to agents that are important bio-threats in the context of malevolent use, but uncommon causes of disease in a natural setting, it may be possible to collect relevant data from surrogate non-human hosts in an experimental laboratory setting. Furthermore, it will be critical to collect a large amount of additional data from humans exposed to agents that cause clinical syndromes close to, or indistinguishable from those caused by bio-threat agents, both before and after onset of clinical findings. One goal will be to recognize critical disease processes during the incubation period, prior to signs and symptoms, or else during an early clinical stage when the patient has mild, nonspecific signs and symptoms, and is not yet debilitated. In order to acquire these large, new sets of data, coordinated multi-center clinical studies will be needed. Standardized methods and tools will be needed in order that the resulting data from different subjects at different geographic sites and at different times, are comparable.

From these expanded clinical studies we will learn how and where gene expression profiles can be applied for early detection and prognosis of infectious diseases. Candidate signatures will be identified, and will then need to be validated with independent sets of clinical specimens and patients. This effort should ideally be an international venture, with the goal of identifying signatures that are useful for human populations of diverse origins. Other needs include the development of more automated methods and miniaturized devices for measuring gene expression patterns and other genomic patterns that reflect human response to disease. Blood may not be the most appropriate type of specimen for large scale monitoring of human populations; thus, the utility of other specimens types, such as saliva and urine, should be explored. Finally, it is expected, but necessary that this technology become less expensive.

Future prospects for the use of gene expression profiling in defense against the threats of biological agents to human health

The impact of gene expression profiling on our management of individuals with possible exposure to a bio-threat agent, who might be in the earliest phase of a potentially serious illness, may be enormous. The potential value of this approach includes early indication of an imminent important illness, such as a serious infectious disease of either natural or malevolent origin. This indication may be provided before the onset of significant signs and symptoms, and might contain specific information about etiology and the likely future clinical course for that individual. Microbial agents with disease-causing capabilities often colonize humans without inducing disease. By examining the response pattern of the host (i.e., is the host "perturbed?") one can establish whether the presence of the microbial agent is clinically significant. Thus, in the setting of a large disease outbreak, a diagnostic gene expression signature might distinguish the many-fold greater numbers of "worried-well", from the otherwise indistinguishable people with incipient serious infectious disease. The latter would then be able to receive early, specific treatment, thereby directing what might be scarce health care resources to those in true need. These early diagnostic patterns would also provide valuable information to those responsible for coordinating and planning emergency medical care on a regional or national level.

Gene expression profiling and other genome-wide measurements of human response to disease should be viewed as complementary to more traditional approaches based on direct detection of the infectious agent. We may discover that human gene expression patterns lack the degree of microbiological specificity provided by direct microbial detection approaches. This weakness will probably be most evident when the goal is a forensic one and focussed on

establishing attribution. Concerns about maintaining privacy that are raised by genomic approaches such as this must also be addressed. On the other hand, one advantage to host expression profiling that has not been mentioned so far is its ability in theory to identify hosts who have been affected by novel, genetically-engineered microbial agents. This broad degree of "coverage" across the biological "threat space" is a strong relative advantage for the host response approach. Other potential uses of this approach might include the discovery of novel early markers of host protective immunity following immunization. By using a host expression pattern as a surrogate indicator of protection, new vaccines against bio-threat agents might be developed and tested much more effectively.

At some point in the future, one might imagine a system that would permit minimally-invasive routine monitoring of each person's genome-wide response pattern, perhaps on a daily basis. The establishment of a individual-specific baseline would maximize our ability to recognize significant events at a pre-clinical stage, and alert the individual to seek medical attention, as well as alert the health care system to early signs of a more widely-distributed problem.

Conclusion

We stand on the verge of acquiring novel capabilities for recognizing and characterizing disease caused by a wide variety of biological agents at an early phase of the illness. These capabilities are brought about by discoveries and advances in the field of genomics. Clinicians and other point-of-care providers desperately need these kinds of capabilities. In order to bring to fruition the promises raised by these advances, we will need to address important, as yet unanswered scientific questions, conduct carefully-designed large-scale clinical studies, and promote further maturation of the associated technology. I would be pleased to answer any questions you may have.