

How Can Genomics Facilitate Vaccine Development?

WORKING GROUP DESCRIPTION

Background

Vaccines are the most efficacious means of minimizing the impact of infectious diseases on the human population. The challenges and importance of making vaccines that will meet FDA approval have never been greater. Genomics has the potential to improve the process of vaccine development substantially. Genome sequencing can help to identify genetic patterns related to the virulence of a disease, as well as genetic factors that contribute to immunity or successful vaccine response. All this information could lead to vaccines with better and more specific targets that elicit more successful protective immune responses. Comparing the genome sequences of viruses that cause infection with those that do not, may provide additional insights. In turn, genome manipulation can facilitate derivation of attenuated strains or other vehicles for delivery of the desired antigens to stimulate immune response. On the other end of the spectrum, analysis of host diversity can reveal effective immune responses and possibly the genetic basis for inappropriate response. The recent progress in definition of the innate immune system, necessary for acquired response, should facilitate the definition of this host diversity.

The Problem

Explore the ways these and future approaches in genomics might be applied to speed the development of vaccines. Targets include emerging threats from either natural reservoirs or terrorist activities; established targets for which there are either no or only ineffective vaccines; and pathogens with effective vaccines with unacceptably high rates of untoward reactions. Beyond genome sequence analysis, participants should consider related technologies, such as the analysis of gene expression by either the pathogen or host upon infection or vaccination; proteomic analysis, including protein-protein interactions within the pathogen or between host

and pathogen; pathogen and host rapid phenotyping, whole genome synthesis; and the design of more effective vaccine vehicles and adjuvants.

Initial References

Fauci, A.S., N.A. Touchette, and G.K. Folkers 2005. Emerging infectious diseases: a 10-year perspective from the National Institute of Allergy and Infectious Diseases. *Emerging Infectious Diseases* 11(4):519-525.

Rappuoli, R. 2004. From Pasteur to genomics: progress and challenges in infectious diseases. *Nature Medicine* 10(11):1177-1185.

Relman, D. , and S. Falkow. 2001. The meaning and impact of the human genome sequence for microbiology. *Trends in Microbiology* 9(5):206-208.

Scarselli, M., M.M. Giuliani, J. Adu-Bobie, M. Pizza, and R. Rappuoli. 2005. The impact of genomics on vaccine design. *Trends in Biotechnology* 23(2):84-91.

Segal, S. and A.V. Hill_. 2003. Genetic susceptibility to infectious disease. *Trends in Microbiology* 11(9):445-448.

WORKING GROUP SUMMARY

Summary written by:

Cecilia Dobbs, Graduate Science Writing Student, New York University

Focus group members:

- Dat Dao, Director, Life Sciences and Health Group, Houston Advanced Research Center
- Cecilia Dobbs, Graduate Science Writing Student, New York University
- Ananda Goldrath, Assistant Professor, Biology, University of California, San Diego

- Mui J. Khoury, Director, Office of Genomics and Disease Prevention, Centers for Disease Control and Prevention
- Corinne Lengsfeld, Associate Professor, Engineering, University of Denver
- Alan McBride, Researcher, Gonçalo Moniz Research Center, Oswaldo Cruz Foundation, Salvador, Brazil
- Catherine McCarty, Interim Director and Senior Research Scientist, Center for Human Genetics, Marshfield Clinic Research Foundation
- Fabienne Paumet, Associate Research Scientist, Physiology and Biophysics, Columbia University
- John E. Wiktorowicz, Associate Professor, Human Biological Chemistry and Genetics, The University of Texas Medical Branch
- Zhenhua Yang, Assistant Professor, Epidemiology, University of Michigan, School of Public Health

Summary

If the developing world could run classified ads, one ad would certainly read, “Needed: vaccines, fast and cheap.” In 2005 alone, if there had been a vaccine for malaria, it could have saved more than a million people’s lives. If there had been a vaccine for HIV, it could have protected more than 3 million people from infection. If there had been a method of manufacturing influenza vaccines that did not need 12 months and hundreds of millions of chicken eggs, many of the current fears over pandemic flu could be allayed. Each case is different, but the bottom line remains the same: if we are ever to triumph over infectious disease, vaccine development needs to become faster and cheaper.

It takes between seven and fifteen years and \$200 million to \$600 million to develop a vaccine. In many cases, that financial investment is well worth it. Vaccines still represent one of the most cost-effective medical interventions for preventing death from disease, and hundreds of millions of lives have been saved over the decades because of them. Yet the basic approach to development has changed very little during that time, and scientists have had little or no success

in combating diseases like malaria, TB, and HIV. Clearly, new approaches are needed for old problems.

One promising source for innovation is the burgeoning field of genomics and proteomics. Current and future technologies in these fields will enable scientists to study an organism's entire set of genes and proteins simultaneously, instead of working with one gene or protein at a time. This could help identify genes or proteins that play a key role in a pathogen's ability to infect and the host's immune system response, and ultimately lead the way to better vaccines. Exactly how much could the fields of genomics and proteomics enhance the process of vaccine development? At the third annual National Academies Keck *Futures Initiative* a group of nine researchers, with backgrounds ranging from bioengineering to immunology, came together to tackle this very question.

In the case of infectious diseases, of course, there are actually three genomes or proteomes to consider: that of the host, the pathogen, and the vector (like the mosquito in the case of malaria). When it comes to studying the host genome, some people are naturally more resistant to disease than others, and scientists may be able to identify genes or portions of genes that play a role in that immunity. Alternatively, data from the pathogen genome may help scientists identify the genes that play a key role in the pathogen's ability to infect and harm the host. Vaccine development begins with epidemiology and pathogenesis—information gathering about the disease itself. Before researchers can begin to develop a vaccine, they need to know what type of pathogen causes the disease, how it is transmitted, what cells it targets in the body, and what negative effects it has on the host. Currently, most epidemiology efforts, particularly with emerging diseases, focus entirely on the characteristics of the pathogen, and do not take genetic variables in the host into account.

The more data collected, the easier it is to develop an appropriate and properly targeted vaccine. Some of the major stumbling blocks to vaccine development could be overcome if better and more complete information existed on the genetic variability of pathogens, as well as on the genetic basis for host susceptibility or resistance to disease. On the proteomics side, more data could enable scientists to identify cellular responses to infection, as well as biomarkers for infection and for successful vaccine response. The group concluded that vaccine development research needs a better infrastructure than there is right now for data collection that would include information in populations on both pathogen and host, and specifically one that would

record the type of immune response, if any, elicited in the host, as well as any genomic or immunological markers that may aid in properly targeting the vaccine.

A second shortcoming in the basic research stage occurs in the laboratory. The pathogen needs to be culturable in the laboratory, but the conditions for growth are often very specific and difficult to determine. Sometimes this means that an organism simply will not grow in culture, but in other instances the culture conditions actually create selection pressures that result in mutated strains of the pathogen that do not exist in nature and are therefore useless for developing a vaccine.

A similar challenge exists for scientists who are trying to crystallize proteins to study their shape, because the conditions for crystallization vary widely with different types of proteins and are also difficult to achieve. A member of the group suggested that a technology used in protein crystallization could be applied to culturing pathogens, and the concept of the nanoarray incubator was born. This incubator would have several intersecting reservoirs, each inputting a different component of the growth environment, thus creating multiple compartments, each with a unique growth environment. Thus, a variety of growth environments could be tested simultaneously and the correct conditions could be identified more quickly.

Once an organism has been successfully cultured in the laboratory, the next step in developing a vaccine is identifying which of the pathogen's proteins should be the target antigens. This is a monumental task; there are many, many possibilities, and there is no easy way to determine which antigens will elicit the right type of immune response or a strong enough immune response. This involves a lot of trial and error, but the group determined several possible ways that genomics might be able to improve both the speed and the accuracy of the process.

High-throughput methods, in which multiple antigens can be screened simultaneously, could dramatically speed the process. In this way, a scientist can study and characterize many antigens in the same amount of time that it previously would have taken to study just one antigen. An existing process, called reverse vaccinology, begins with the sequenced genome of a pathogen, and then uses statistical analysis to identify the genes that are most likely to influence the pathogen's ability to infect the host. The proteins that these genes code for become the target antigens, and a vaccine is created from this information. None of the vaccines currently available were developed this way, but the group felt that using genomic information this way has the potential to dramatically streamline the vaccine development process in the future. A similar

approach could be used with proteomics: sequence the entire proteome of the pathogen and then target that whole protein combination, instead of just a single antigen or two.

After selecting the target antigens, the next step in the process is testing the vaccine in animal models to see whether it is both safe and effective. Unfortunately, animal models are far from perfect. It is possible, even likely, that a good result in animals will not translate into a good result in humans. The group suggested that with genomic analysis, future animal trials may be able to use biomarkers to quickly identify whether a vaccine is effective or toxic, or maybe even identify genetic markers that predict how effective a vaccine will be in humans. Other markers might be able to identify whether the vaccine will convey protective immunity.

Exploring even further outside the box, genomic advances may someday make animal models obsolete. If scientists can use genomic data to engineer human tissue in the laboratory, they may be able to study host-pathogen interactions and vaccine efficacy directly in human tissue and eliminate the guesswork of translating results from animal to human models.

After all the epidemiological studies, searches for antigens, and animal modeling, a vaccine is finally ready for clinical trials. This is by far the most expensive part of the process, and often trials go on for some time before discovering that the vaccine is ineffective or has unforeseen side effects. Here, too, the group felt genomics could improve the process. It may be possible to identify and test for biomarkers that alert researchers to toxicity or efficacy issues early in the clinical trial process. It may also be possible to identify the most genetically susceptible populations for a particular disease and thereby reduce the sample size needed for an effective clinical trial, which would also reduce the cost.

If clinical trials determine that a vaccine is both safe and effective, it is approved and released to the market. But problems often arise in the real world that did not show up in the controlled experiment. A vaccine may work only for one particular population group, or may be harmful for some other population group. The result is usually that the baby is “thrown out with the bathwater” and a vaccine that could have been very beneficial to some people is discarded entirely because of negative effects in other people. This happened in 1999 when a vaccine for rotavirus was pulled from the market after a small number of children who received the vaccine developed intussusception, a form of bowel obstructions. The vast majority of children who received the vaccine never developed this intestinal problem, so the vaccine was clearly good for some people but not for others.

Here again, the group felt genomic information could help. There may be genetic markers that could identify children for whom the vaccine would be effective and those for whom it would likely cause problems. But very little data is gathered once a vaccine hits the market, meaning that both successes and failures in vaccine development do not inform future work. Just as more data collection is needed at the beginning of the vaccine development process, more information is also needed at the end of the process (postmarket surveillance).

There are several existing databases that could be expanded and used to gather important data on vaccines. For instance, the Vaccine Safety Datalink is a program run by a network of HMOs in conjunction with the National Immunization Program at the Centers for Disease Control and Prevention. It already collects data on about 5 percent of the U.S. population, and could be expanded to collect genomic information regarding immune response to a vaccine. If a larger percentage of the population were registered with this database, researchers could track the effectiveness of vaccines in the real world after they hit the market. The information gleaned from this monitoring could feed back into the start of the process, informing and improving the development of future vaccines.

At nearly every stage of vaccine development, the group was able to identify a number of ways that genomic information could accelerate the production of good vaccines while keeping costs low. Scientists have been searching for vaccines for malaria, tuberculosis, and HIV for many years now, and the classic approaches have yielded few results. But the group was confident that with new information from genomics and proteomics, better systems of development are on the horizon, and we hope new, successful vaccines are as well.