High Speed, High Volume Laboratory Network for Infectious Diseases and Center for Rapid Influenza Surveillance and Research

September 22, 2008

Scott P. Layne, MD
University of California Los Angeles
School of Public Health
scott.layne@ucla.edu

Tony J. Beugelsdijk, PhD MBA
Los Alamos National Laboratory
Chemistry Division
beugelsdijk@lanl.gov

This white paper summarizes our activities, progress to date, and work to come on the High Speed, High Volume Laboratory Network for Infectious Diseases and Center for Rapid Influenza Surveillance and Research at the University of California Los Angeles (UCLA). It emphasizes that implementing near real-time, global and (what aspires to) transparent surveillance of infectious diseases is an essential next step for human health and security. It also explains how our efforts will enable this next step, first at one laboratory site and then at multiple networked sites. Cited on-line materials and references at the end provide additional information.

Critical Importance. Infectious diseases are the leading cause of acute human illness, lost productivity, and death. They are attributed to approximately 17 million deaths per year worldwide, mostly in children under five years of age. Infectious diseases outbreaks can appear unpredictably with significant social, economic, and political consequences. Because of such impacts, the far-reaching International Health Regulations 2005 were enacted to reorganize public health assets as security ones, and Homeland Security Presidential Directive/HSPD-21 was signed by the United States President to realign public health and medical preparedness assets with national security (1, 2).

Surveillance and Research Methods. In 2006, the UCLA School of Public Health and Los Alamos National Laboratory (LANL) began close collaborations to develop the *High Speed, High Volume Laboratory Network for Infectious Diseases*. The goal is to bring about a quantum leap in the surveillance, testing, and analysis of infectious agents from any source. It will vastly increase the rate at which infectious agents are submitted, tested, and analyzed (high speed), the total number of samples tested per day (high volume), and the amount of information available for critical decision making. In everyday situations, it will closely connect (within a few hours to days rather than weeks) surveillance teams in the field to testing and analysis teams in the laboratory. In combination with the real-time monitoring of incident reports, news reports, and Internet searches (words, phrases, their patterns and locations), such closer connections will enable new and more powerful prediction and prevention efforts (3). In emergency situations, it will provide near real-time, global, and enhanced surveillance for making actionable decisions, delivering effective drugs, and selecting optimal vaccines. The initial focus is on all influenza viruses, including H5N1 avian influenza (bird flu), which represent leading threats and have the potential to cause an overwhelming pandemic. However, the laboratory will also be capable of testing other bioagents, including viruses and bacteria (dual use).

As shown in **Figure 1**, key elements of high speed, high volume laboratory include: 1) worldwide surveillance; 2) rapid submission of samples; 3) high-throughput testing; and 4) data-driven analysis. The roles of UCLA are to operate the high-throughput laboratory systems, conduct data-driven analyses, and provide emergency capabilities in the event of an outbreak. It will assume these roles through its School of Public Health and educate new leaders in biological security. The roles of LANL are to build an operating system for the laboratory, design and validate laboratory systems in Los Alamos, install functioning

laboratory systems in Los Angeles, and conduct intensive analysis of influenza. It will assume these roles through its Bioscience, Chemistry and Theoretical Divisions and interdisciplinary Centers and educate new leaders in national security (4).

Figure 1. Key elements of high-throughput surveillance.

Major Funding. The High Speed, High Volume Laboratory Network program is supported by congressionally-directed DoD appropriations of $16 million and a California Office of Homeland Security grant of $9 million (5). In 2007, UCLA was awarded one of six NIH/NIAID Centers of Excellence for Influenza Research and Surveillance (CEIRS) contracts to establish the Center for Rapid Influenza Surveillance and Research (CRISAR). Over five years, the contract will provide $20 million to UCLA and subcontracting institutions including the University of California Davis (UCD) for surveillance, University of Alaska Fairbanks (UAF) for surveillance, Wildlife Conservation Society (WCS) for surveillance, and LANL for influenza dipstick development (6). At present, these affiliated researchers collect about 50,000 samples per year at foreign and domestic surveillance sites; this total number is expected to increase by several fold over time. Surveillance includes places where direct human-animal contacts are frequent and thereby constitute frontiers for animal viruses crossing into humans.

Facilities. As shown in **Figure 2**, UCLA is building a 6,300 sq. ft. BSL3-enhanced (BSL3e) facility to house the high-throughput automated systems. It is located in the California NanoSystems Institute (CNSI) and scheduled for completion and commissioning in May 2009. The BSL3e facility was designed by the global science and technology design firm CUH2A, which specializes in laboratory containment work (7). This firm has previously designed many BSL3 and BSL4 facilities, including NIH's Center for Biodefense and Emerging Infectious Diseases and CDC's Emerging Infectious Disease Laboratory. The BSL3e facility is being built by the construction firm PCL, which has prior experience in containment laboratories and a reputation for consistent high level service (8).

Incoming samples will be barcoded and contained in standard tubes that are compatible with automation. Dedicated rooms and areas will facilitate sample receipt, flexible BSL2/BSL3 operations, and waste disposal through double-door autoclaves. Showers (male/female), waste water holding and chemical sterilization tanks, and perimeter

hallways with video monitors and access controls will conform to BSL3e specifications as set forth in the *Biosafety in Microbiological and Biomedical Laboratories*, 5th edition (9).

The high-throughput laboratory facility will operate with stringent standard operating procedures and quality assurance/control protocols in a Good Manufacturing Practice-like manner. The Information Technology (IT) system will generate permanent records for tracking all infectious disease samples in a chain-of-custody manner. Since highly pathologic avian influenza (HPAI) is designated as a Select Agent, all laboratory personnel will undergo the required background checks and Select Agent registrations (10).

In everyday research mode, the facility will be capable of fully characterizing up to 10,000 influenza samples per year. In emergency response mode (outbreak and/or pandemic), the facility will be capable of processing and partially characterizing up to 10,000 samples per day and enable 24/7 surge operations (11).
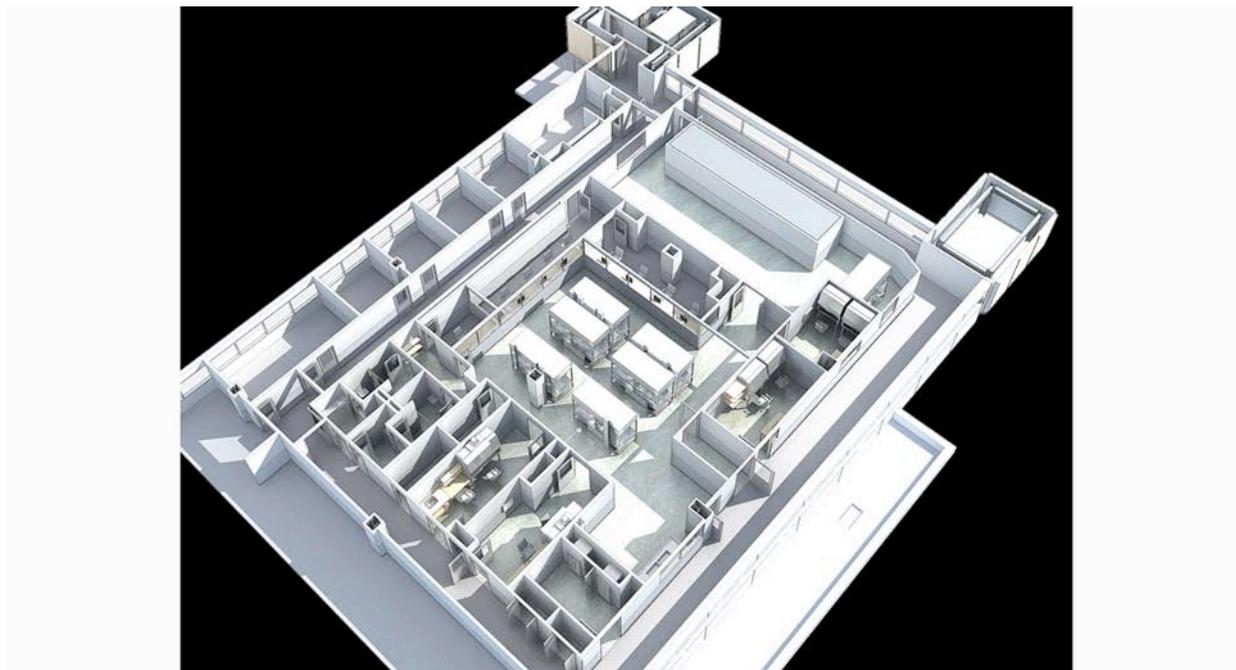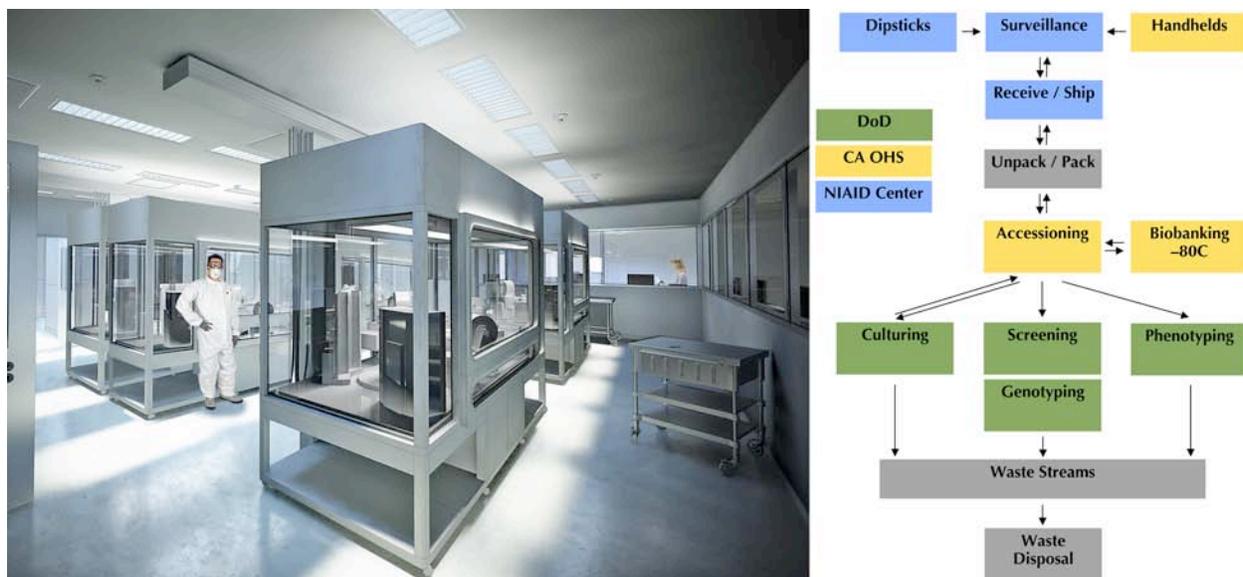

Figure 2. Aerial view of BSL3-enhanced facility.

Staffing. Operation and maintenance in everyday research mode will require a core staff of about eight people at UCLA. They include a highly experienced Laboratory Manager who oversees all activities; a highly experienced Biosafety Manager who trains staff and maintains select agent documentation; an experienced Quality Assurance Manager who assures good laboratory practices and procures certified materials and supplies; and five experienced Laboratory Technicians who have primary and secondary responsibilities for each system within the high-throughput laboratory. The core staff will also have responsibility for training, certifying, and monitoring all researchers and visitors who have need to work in the BSL3-enhanced facility.

Operation and maintenance in emergency response mode will require expansion of the core staff by at least several fold (11). The additional staff will be required to manage the markedly increased workflows, maintain 24/7 operations over many weeks to months, reduce work fatigue, and shortages due to attrition. We plan to recruit and cross-train surge staff from UCLA, LANL, and regional public health agencies. As summarized further below, we also plan to become part of the Laboratory Response Network (LRN) and National Critical Infrastructure.

High-Throughput Systems and Status. As shown in **Figure 3**, the high-throughput BSL3e facility will house several flexible automated laboratory systems as follows. The **accessioning** system will input, clarify, dispense, and repackage incoming samples. The **biobanking** system will store up to up to 800,000 samples at –80C. The **screening** system will use polymerase chain reaction-based assays (PCR) to detect and quantify viruses in samples. The **genotyping** system will use PCR amplification and capillary array-based analyzers to generate partial and/or whole genome sequences. The **culturing** system will use animal and human-based cell lines; it will maintain and monitor 50,000 to 100,000 individual test cultures at once. The **phenotyping** system will use viruses, sialated latex spheres, and World Health Organization (WHO) typing sera to perform hemagglutinin inhibition (HI) assays. The initial emphasis for these systems will be influenza but, because of their inherent flexible and programmable designs, protocols can be implemented for practically any BSL3e-level infectious disease agent (12 - 14).

The accessioning and biobanking systems have been acquired from the laboratory automation vendors Tecan and REMP, respectively (15, 16). A Request for Proposals (RFP) for the genotyping system has been completed and will soon be acquired from a competitive bidder. Design specifications for the screening, culturing, and phenotyping systems are being developed and their RFPs will soon be advertised in consecutive order. Following their acquisition, each automated system will undergo several months of initial testing and validation at LANL before final installation and operation at UCLA.

Figure 3. Interior view of the BSL3-enhanced facility (right) and sample workflows (left).

Handhelds. As shown in **Figure 4**, LANL has designed and delivered 150 portable handheld devices to UCLA that incorporate a GPS transceiver, digital camera/barcode reader, voice recognition, Wi-Fi and Bluetooth connectivity, and surveillance questionnaire (17, 18). These paperless devices can be used in everyday research or emergency response modes and their questionnaires can be easily adapted to any biological agents and/or outbreak situation. They are now being used by surveillance teams to capture epidemiologic data at various geographic locations. Convenient protocols and software are then used to upload data from handhelds to surveillance databases. With surveillance-based data from handhelds and laboratory-based data from high-throughput systems, it will become possible to address many important hypothesis and data-driven questions for the first time.

Figure 4. Portable handheld surveillance devices.

Influenza Dipsticks. LANL is also developing influenza dipsticks that yield positive/negative test results in 30 to 60 minutes (19). These disposable and simple-to-use devices are planned to cost less than $15 each and are amenable to mass production and distribution. As shown in **Figure 5**, they work as follows. 1) Samples are introduced into dipsticks with standard swabs. Influenza viruses are captured by sialic acid-coated magnetic beads, aggregated by a small permanent magnet, and washed to remove unbound materials. 2) Captured viruses are heated to disrupt their envelopes and release their RNA genomes. 3) Isothermal reverse transcription and DNA-amplification are used to generate billions of copies of specific influenza gene sequences. 4) DNA-amplicons adhere to antisense probes that are bound to blue-colored beads and are next captured by antisense probes on the lateral flow membrane. Visible blue lines appear when influenza-specific amplicons are present in samples.

At present, temperature stabilization for isothermal reactions is achieved by an embedded microheater/microprocessor powered by two AA batteries. This configuration enables dipstick use in excessively cold or hot climates. The devices are intended to achieve detection thresholds (sensitivity) of 100 viral genomes and molecular discriminations (specificity) that exceed 99.9%. They offer direct visual readouts (lines appear for positives tests and positive controls), multiple flow channels for discriminating influenza types (A vs. B) and/or subtypes (H1, H3, H5, H7, etc), and the flexibility to add molecular tests for other bioagents.

Highly sensitive and specific influenza dipsticks will have many uses. Here are three important ones. For surveillance teams in the field, they can be used to determine whether influenza viruses are present in sick or dead animals. For emergency response teams at outbreaks, they can be used to make simple and rapid diagnostic assessments before samples are analyzed by the high-throughput laboratory. For health care workers during pandemics, they can be used to rule out influenza infection (viral shedding) before returning home to family members.

5
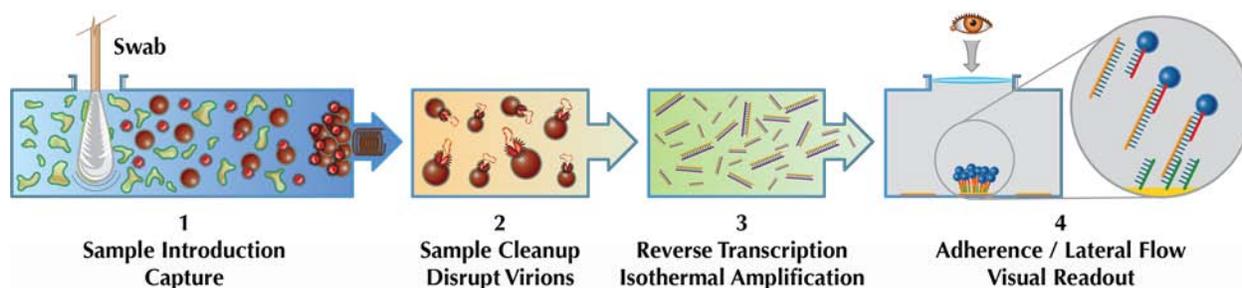
Figure 5. Influenza dipsticks.

Information Technology (IT). UCLA and LANL are implementing an IT system and infrastructure for everyday research and emergency response modes. The overall design is based on a Service Oriented Architecture (SOA) that maximizes system scalability, computing platform-independence, and also simplifies integration with heterogeneous (internal and external) system components and networks (20). The Web Services Description Language (WSDL) and SOAP are being used throughout to enable system flexibility and allow open data and communication standards such as the Extensible Markup Language (XML) (21 - 23). The required 24/7 fault tolerance is achieved with: redundant geo-spatially distributed servers at UCLA and LANL; a Storage Area Network (SAN) that incorporates a Redundant Array of Independent Disks (RAID); and high-speed database mirroring over a 1 to 10 gigabit per second optical wide-area network link between Los Angeles, CA and Los Alamos, NM.

As shown in **Figure 6**, the IT system integrates and controls most functions of high-throughput laboratory. It also manages epidemiologic, genotypic, phenotypic, quality control, and audit trail information on all samples with its Laboratory Information Management System (LIMS) (24). The web-based LIMS will serve a variety of functions as follows. It will enable compliance with 21 CFR Part 11 (Electronic Records and Electronic Signatures) and the Health Insurance Portability and Accountability Act (HIPPA) (25). It will manage complex laboratory workflows, manage laboratory consumables and supply chain inventories, and enforce Standard Operating Procedures (SOPs) and Quality Assurance/Quality Control (QA/QC) protocols. It will manage user data acquisition, instrument data acquisition, data archiving functions, and automated system interfaces. It will record device maintenance and calibration records, store personnel training records and laboratory certifications, and automate billing and reporting.

In addition, the IT system will implement a role-based security model that enables push-and-pull plus publish-subscribe capabilities over the Internet. This feature will allow users (researchers, collaborators, customers, public health providers, government officials, and other federated laboratory systems) to upload and/or download information on a 24/7 basis. They will be able to use virtually any type of computing platform, including smart phones, laptops, desktops, and mainframes. By utilizing concepts that originated with Cloud Computing, the IT system will allow users to access technology-enabled services "in the cloud" and use them without specific knowledge, expertise, and control of the underlying technology infrastructures (26). (**Note:** Cloud Computing approaches allow much needed flexibility and scalability for the High Speed, High Volume Laboratory Network for Infectious Diseases and Center for Rapid Influenza Surveillance and Research. They will also allow us to harness the virtually infinite computing and data resources provided by Cloud Computing providers (such as Google, Microsoft, and Amazon) in an on-demand, pay-as-you-go fashion.)
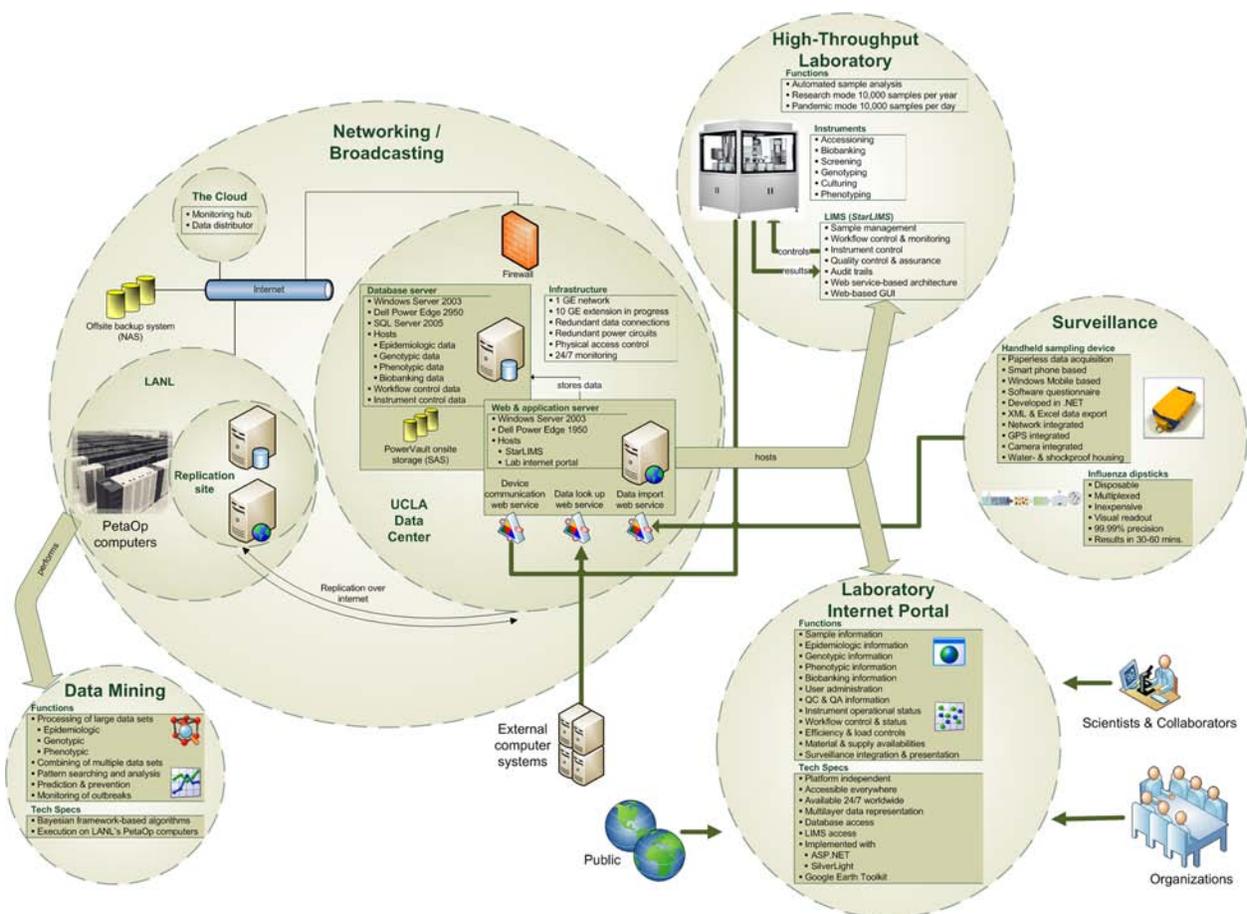
Figure 6. Information Technology (IT) system, architecture, and resources.

Important Scientific Questions.   The Severe Acute Respiratory Syndrome (SARS) and avian H5N1 influenza outbreaks are two examples of our limited means to deal with emerging infectious diseases threats.   For SARS, fortunately, a combination of luck and international cooperation controlled the outbreak.   Infected people became febrile and clinically ill before shedding the SARS virus, which allowed case identification and quarantine measures to work — although precariously.   For the avian H5N1 influenza outbreak, however, no such luck and cooperation have prevented the ongoing outbreak.

Since 2003, avian H5N1 influenza has spread from Asia to Russia, Europe, the Middle East, and Africa through wild bird migrations and animal transportations.   This ongoing outbreak has been attributed to over 385 human cases, 243 human deaths, a case fatality of 63% that is increasing, the destruction of over 180 millions of food animals, and human malnutrition in the underdeveloped world.   At present, spread to the Americas seems inevitable but its timing is unknown (27).

Yet scientists and health officials cannot estimate the probability that avian H5N1 influenza will cause a pandemic. We do not know which combinations of viral genotypes characterize efficient human-human spread.   The much needed databases that associate epidemiologic behaviors with genotypic and phenotypic features for thousands of influenza strains must still be created.   It therefore remains unknown whether circulating avian H5N1 influenza strains are capable, or even close, to starting a pandemic or not. Consequently, health officials must consider worst case scenarios.

Our current "unknown" assessments of influenza (and other infectious disease agents) must be improved. Such action will require enhanced animal and human surveillance, sample collection worldwide, high-throughput laboratory testing, and intensive data-driven analyses. It will result in a dynamic, empirical database of influenza strains that are emerging in animals and humans. At present, avian H5N1 influenza represents the most important biological threat but similar predictive approaches can also be applied to human immunodeficiency virus (HIV), Mycobacterium tuberculosis (TB), methicillin-resistant Staphylococcus aureus (MRSA) and as yet unknown infectious diseases. Although each disease presents different challenges and strategies for prevention, the need for predictions that link epidemiologic, genotypic, and phenotypic features remains the same (28).

The envisioned infectious diseases database must store information on thousands of diverse strains and bring together several types of information on each and every one. It must include information on multiple levels. *Epidemiologic* information that is collected from observations in the field with dates, locations, hosts, outcomes, histories, and exposures. *Genotypic* information that is generated from high-throughput laboratory testing with complete genetic sequences whenever feasible. *Phenotypic* information that is generated from high-throughput laboratory testing and with immunologic profiles for vaccines and antiviral drug sensitivity patterns for therapies. Data-driven computational biology methods that make use of Bayesian frameworks can then be used to mine the information, identify associations, and assign uncertainties to them. Such methods and algorithms are available and have been successfully applied by researchers at UCLA (29 - 32). For extremely large databases, they can be optimized, and run on supercomputers at LANL or elsewhere. In addition, multi-scale mechanistic mathematical models and risk assessment tools can be used to estimate the impact of various interventions and policies on potential outbreaks.

As shown in **Figure 7**, a few of the most important scientific questions for influenza viruses include:

- Which genotypes characterize human and/or animal-derived strains (host range)?
- Which genotypes characterize efficient human-human spread (airborne transmission)?
- Which genotypes characterize deadly outcomes (virulence)?
- How do genotypes map to specific immune phenotypes (vaccine selection)?
- How do genotypes map to specific drug phenotypes (antiviral selection)?
- How do these various associations relate to pandemic potential (threat ranking)?
- What strain will cause the next pandemic (unknown)?
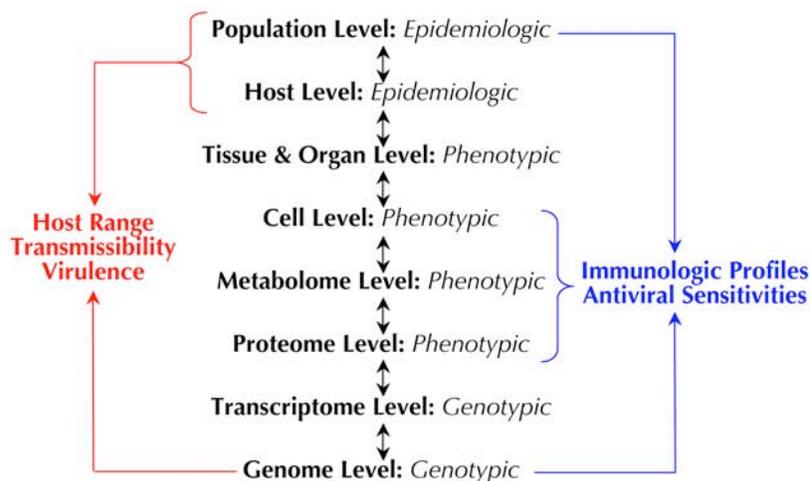
Figure 7. Data-driven questions and associations.

Opportunities for Partnership and Collaboration. Our first goal is to build and operate a high-throughput laboratory at UCLA and have it serve as a powerful and persuasive model of robust surveillance (34). As shown in **Figure 8**, about 25,000 samples per year will be provided by the Center for Rapid Influenza Surveillance from California, Alaska, Russia, Japan, Mongolia, Cambodia, Cameroon, Congo, Kenya, Tanzania, Uganda, and Democratic Republic of Congo (6). Another 25,000 samples per year will be provided by the UCLA Institute of the Environment's Center for Tropical Research from across the United States, Mexico, Guatemala, Panama, Honduras, Columbia, Ecuador, and Bolivia (33). These samples are now being tested and analyzed by traditional laboratory methods, which often takes many weeks to months. The high-throughput laboratory, however, will reduce this time to days.

Our next goal is to build and operate high-throughput laboratories at additional geostrategic locations and network them into a fully integrated surveillance system as called for by the International Health Regulations 2005 and ranking U.S. government officials (1, 35). To achieve this more significant and challenging goal, laboratory automation, information technology, workflows, and quality assurance/controls will be selected and implemented with interconnectivity and interoperability in mind. Inherent flexibilities in the overall design will enable various high-throughput laboratories to develop and share methods for testing and analyzing various other bioagents, including bacteria and viruses (14). In the event of a major infectious disease outbreak, however, this network of high-throughput laboratories will have the ability to shift focus and work together.

The first high-throughput laboratory at UCLA will serve as a "gold standard" for validating promising new technologies for infectious diseases surveillance and research. It will also serve as an important center for mentoring the next generation of leaders in biosecurity. Below we summarize important opportunities for partnership and collaboration in a number of areas.

Figure 8. Current national and international surveillance by country.

Next Generation Biobanking. The receipt of about 50,000 samples per year from surveillance and their subdivision into multiple aliquots (typically five) by the accessioning system will yield about 250,000 subdivided samples per year. They will be placed into uniquely coded, hermetically-sealed cryotubes to enable tracking and chain-of-custody and to eliminate cross contamination. One will be sent to the screening system for influenza detection and the other four will be sent to the biobanking system for storage at –80C. The automated biobanking system (*REMP Sample Safe*) will perform sample input/output, storage/retrieval, and inventory management functions (16). It will incorporate redundant cooling for sample protection and several automated –80C storage compartments that are accessed by a robotic arm and automated picking mechanism operating within a –20C compartment. The biobanking system will reliably maintain up to 800,000 0.3 ml cryotubes within a narrow temperature range (–80 ± 5C) and, at current surveillance rates, will reach capacity within three to five years.

The biobanking system will store samples from around the world and offer an on-line inventory lookup. It will present information on sample availability, sample quantity, sample sharing, import/export restrictions and linked epidemiologic, genotypic, and phenotypic records. This integrated inventory, tracking and record system will constitute a new, scalable "biological wiki." It will provide essential services for the coming high-throughput laboratory network and offer an open, adoptable, and flexible biobanking model for the infectious diseases surveillance and research communities.

There will be several opportunities for partnership and collaboration in this markedly improved biobanking effort. The first pertains to technical development and refinement of the scientific search engine. The second pertains to mining the inherent value of influenza-negative samples for the discovery of other infectious diseases agents. The third pertains to development of broad, cost-effective sample screening methods and inventory management practices to preserve valuable biobanking space.

Next Generation Sequencing. On average, about 5% of surveillance samples yield positive polymerase chain reaction (PCR) tests for influenza viruses. At the current collection rate of 50,000 samples per year, this corresponds to the opportunity to whole genome sequence (genotype) about 2,500 influenza viruses per year. This number will increase with better guided surveillance efforts (more positives from the field), improved viral stabilization and transportation methods (less loss from the field), and expansion to other countries and sites (more samples from the field). Conceivably, the total number of whole genome sequences could reach 10,000 per year, which will present both cost and technical challenges. There is therefore a great need to utilize best practices with current sequencing technology and facilitate next generation sequencing technology for infectious disease agents.

The first genotyping system in the high-throughput laboratory is based on mature Sanger technology. It uses 192 overlapping influenza-specific primers to amplify all eight gene segments, PCR to amplify products from primers, and capillary array DNA analyzers (Applied Biosystems 3730xl) to sequence these products. The overall method is proven and robust — it sequences over 60% of whole influenza genomes on the first attempt and over 90% on the subsequent attempt — but it also has several drawbacks. First, influenza-specific primers must be continuously updated to achieve better success on the first attempt. Second, primer-based methods are best suited to detect known or similar sequences, but they are not suited to detect unknown or divergent ones. Third, large primer sets require large amounts of viral RNA and, consequently, influenza viruses must usually be grown in cultures to provide sufficient starting material. This culture step, however, adds time and cost and fundamentally limits influenza discovery to those viruses that can be grown in chicken eggs or animal cells.

Given the above limitations, the next generation sequencing technology for infectious disease agents must be primerless and capable of rapidly testing samples from the field. In order to maintain high-throughput laboratory workflows, it must also work with conventional automated liquid handling processes and procedures. One of the most promising new technologies that will meet these requirements is from Pacific Biosciences for example (36). It has the potential to produce sequences form single strands of DNA or viral RNA that is reverse transcribed to complimentary DNA in real time.

There will be several opportunities for partnership and collaboration in this markedly improved sequencing effort. The first pertains to side-by-side comparisons of current sequencing technology to next generation sequencing technology and its validation for influenza. The second pertains to development of laboratory methods for other infectious disease agents. The third pertains to expansion of affordable and simplified sequencing technology for other laboratory network sites.

Dipstick and Laboratory Integration. Our program is developing two types of infectious disease diagnostics. One is represented by portable influenza dipsticks that will provide simple positive/negative test results within 30 to 60 minutes (19). The other is represented by high-throughput laboratory resources that will provide complete epidemiologic, genotypic and phenotypic results within days (13, 14). These two diagnostics will offer faster outbreak testing, greater outbreak testing, and higher confidence for critical decision making in first and third-world settings. Their integration is therefore important.

Direct visual readouts from influenza dipsticks will permit point-of-care testing and reporting of test results to databases. The appearance of lines for positives tests and positive controls can be reported by satellite, Internet, and cellular based communications in the form of pictures, keystrokes, and voice recognition. The incorporation of unique serial numbers on each dipstick can also be used for a variety of purposes including incentives to report test results and quality assurance. In both everyday research mode and emergency response mode, the results from dipsticks will enhance and complement results from the high-throughput laboratory.

There will be several opportunities for partnership and collaboration in this diagnostics integration effort. The first pertains to accelerating influenza dipstick development, validation, scale up, production, and distribution. The second pertains to creating a simple and robust reporting portal for influenza dipsticks and devoting sufficient global bandwidth so that it functions during an outbreak and pandemic. The third pertains to devising optimal strategies for influenza dipstick use and extending this PCR-based technology to other infectious disease agents.

Next Generation Culturing. Influenza-positive surveillance samples from animals and humans contain anywhere from 10^3 to 10^8 viruses per ml, with most containing about 10^5 viruses per ml. Such concentrations are sufficient for PCR-based detection and qualification by screening assays, but insufficient for whole genome sequencing, antiviral sensitivity/resistance testing, and immunologic profiling. To perform these assays, influenza viruses must be grown in embryonated chicken eggs and/or cultured cell lines.

At present, trivalent influenza vaccines are made by growing genetically-engineered strains in embryonated eggs. Likewise, influenza viruses from surveillance samples are often grown in such eggs. These rather old and cumbersome processes are now being replaced by relatively new and convenient ones that use cultured cell lines (37). For the high-throughput laboratory, there are several good reasons to follow this trend. 1) Embryonated eggs restrict the discovery of new influenza viruses to those strains that grow in them. Whereas cultured cells can be used from multiple lines and avoid this fundamental restriction. 2) Embryonated eggs are not often used for sensitivity/resistance

testing of antivirals such as Oseltamivir, Zanamivir, Amantadine and Rimantadine. Whereas cultured cells are used for such antiviral testing. 3) Embryonated eggs are used to grow relatively few different viruses (viral families) in the laboratory. Whereas cultured cells can be used to grow many different viruses in the laboratory (SARS, HIV and West Nile for example). 4) In the event of an influenza pandemic, the supply of embryonated eggs will likely become scarce or uneven. Whereas cultures cells will always be available from preserved stocks on site.

Given the above limitations, the next generation culturing system for influenza viruses (and other infectious agents) must be based on multiple cell lines. In order to maintain high-throughput laboratory workflows, various steps in the process (cell dispensing, viral inoculation, incubation, growth monitoring, viral harvesting, and packaging) must also work with conventional automated liquid handling processes and procedures.

There will be several opportunities for partnership and collaboration in this markedly improved culturing effort. The first pertains to development of new (universal, high yield) cell lines for influenza virus culture by genetic engineering of molecular/cellular pathways. (**Note:** This effort is already underway at UCLA). The second pertains to direct comparisons of manual methods for antiviral sensitivity/resistance testing to fully automated methods and its validation for influenza. The third pertains to development and standardization of automated cell-based methods for other infectious disease agents.

Next Generation Phenotyping. Influenza is a rapidly-mutating RNA virus that constantly changes its appearance to the human immune system. The nonstop process of reproduction, mutation, and selection of immune-escaping strains requires the biannual production and distribution of new vaccines. Since influenza cases in humans peak during the fall and winter months, the WHO recommends new vaccine strains every six months for the northern and southern hemispheres, respectively (38).

The surfaces of influenza viruses are covered by a mosaic of three proteins. These surfaces include about 500 hemagglutinin (H), 100 neuraminidase (N), and 20 channel (M2) proteins per virus and, for humans, protective immunity to influenza viruses follows in the same order. Antibodies that bind to hemagglutinin are the most protective. Antibodies that bind to neuraminidase are intermediate. And antibodies to channel M2 are the least protective. Consequently, influenza vaccine strains are selected using Hemagglutinin Inhibition (HI) assays, which measure the immunologic relatedness among influenza strains. Because of their fundamental importance to public health and pandemic preparedness, there are strong demands to improve them.

HI assays are based on similarities between sialic acid receptors on respiratory cells and sialic acid ligands on red blood cells, which gives influenza the ability to agglutinate red blood cells in suspension. HI assays are performed by mixing red blood cells and influenza viruses in a series of wells, and next adding typing sera in two-fold dilutions to these wells. After the reactions stabilize, laboratory workers score assay wells (+/–) by looking for agglutinated red blood cells (+) which do not form bottoms at the bottom of plates versus non-agglutinated cells (–) which do form buttons, as shown in **Figure 9**. In most situations, panels of six to ten typing sera are used to match sample strains against vaccine and reference strains and, during any given influenza season, over 99% of sample strains are successfully matched by routine HI assays (39). Sample strains are categorized according to immunologic pedigrees such as: A/Brisbane/59/2007 (H1N1)-like virus; A/Brisbane/10/2007 (H3N2)-like virus; and B/Florida/4/2006-like virus. (40).

HI assays currently used by WHO Collaborating Centers and Reference Laboratories are based on avian red blood cells from chickens and turkeys. Unfortunately, the biological properties of these cells vary from laboratory to laboratory, which makes it difficult

(sometimes even impossible) to compare phenotyping results across the WHO network. Furthermore, in the event of an influenza pandemic, the supply of avian red blood cells will likely become scarce or uneven. Given these limitations, the next generation phenotyping system for influenza viruses must be based on markedly improved HI assays.

Over the past two years, LANL has developed methods to synthesize 2,3- and 2,6-sialic acids for capturing influenza viruses, and methods to attach these synthetic molecules to any surface (41). These biochemical methods are now ready for scale up to commercial production quantities. In their magnetic bead form, these synthetic sialic acids are being used to capture viruses in influenza dipsticks as described previously. In their non-magnetic bead form, they can be used to mimic and replace avian red blood cells in HI assays. (**Note:** Sialated beads can be freeze-dried and are highly stable at room temperatures.)

There will be several opportunities for partnership and collaboration on this markedly improved HI assay and phenotyping effort. The first pertains to performing side-by-side comparisons of old red blood cell-based assays to new sialated bead-based assays and its validation for vaccine strain selection purposes. The second pertains to developing a worldwide distribution and supply chain for these sialated beads. The third pertains to developing further improved phenotyping methods based on flow cytometry or optical waveguide biosensors for example (41).
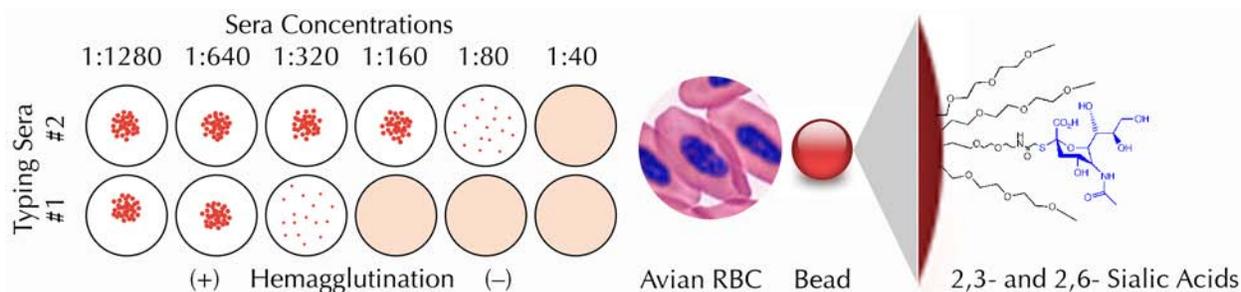
Figure 9. Hemagglutinin Inhibition assays (left) and sialated beads for HI assays (right).

Data Mining and Visualization. The September 4, 2008 issue of *Nature* highlights the era of "BIG DATA" and "SCIENCE IN THE PETABYTE ERA." Its first page editorial further emphasizes that "researchers need to adapt their institutions and practices in response to torrents of new data — and need to complement smart science with smart searching" (42). Our high-throughput laboratory program will not only generate torrents of new data, it will also need to rapidly mine and effectively visualize this data in combination with other relevant data.

For example, enormous datasets that have use for monitoring and early warning of infectious diseases outbreaks are now being generated by Internet searches worldwide (3). And, sooner or later, enormous datasets from a major infectious diseases outbreak or influenza pandemic will be generated by the global surveillance, research, and response communities.

There is therefore great need to develop and integrate a variety software tools with different purposes in mind. We need tools that extract, integrate, mine, and visualize multiple datasets at once. We need tools that enable infectious disease scientists and epidemiologists to analyze and look for patterns in epidemiologic, genotypic, and phenotypic datasets. We need tools that assess policies and interventions and present big pictures to health officials and policy makers for effective decision making and action. We need simple data collection and awareness tools that run on cell phones and/or PDAs as well as more powerful analytic tools that run on high-performance (PetaOp) computers.

At present, we are developing the Information Technology (IT) system that will provide a framework and foundation to develop such tools. It stores epidemiologic, genotypic, phenotypic, quality control, and audit trail information on all samples processed by the high-throughput laboratory. It makes use of an open-source architecture that enables improvements, upgrades, and scalability. In addition, as mentioned above, we are developing data-driven computational biology tools and algorithms that utilize Bayesian frameworks to mine information, identify associations, and assign uncertainties to them (29 - 32). Such tools and approaches, however, are merely the beginning of what is needed for near real-time, global, and (what aspires to) transparent surveillance of infectious diseases.

There will be several opportunities for partnership and collaboration on this data mining and visualization effort. The first pertains to adapting existing tools or developing new tools to extract, integrate, mine, and visualize multiple datasets at once. The second pertains to developing HIPAA-compliant tools that work together for human and animal surveillance (25). The third pertains to revising, maintaining, and documenting tools for a wide user base.

Nanotechnology and Infectious Diseases. The UCLA high-throughput laboratory is situated in the California NanoSystems Institute (CNSI), which facilitates interdisciplinary research across many disciplines (43). Our program will open new opportunities for applied research, development and delivery of nanotechnology in several ways. It will provide biological expertise to help guide the selection of promising nanotechnology. It will provide well-characterized infectious disease samples to use and appropriate biological safety level space to test promising nanotechnology. It will provide high-throughput laboratory systems that serve as "gold standards" to compare commercial macrotechnology against promising nanotechnology.

At present, there is significant promise but no commercially available nanotechnology for infectious diseases related work. In theory, nanotechnology can significantly increase sample and cost savings by reducing sample volumes and consumable volumes. In practice, however, there are unavoidable signal-to-noise relationships that apply. Very small working volumes equate to very few bioagent copies per volume. Practical nanotechnology for infectious disease detection must therefore take advantage of this relationship.

Educational and Training of Leaders. The high-throughput laboratory program will provide unique opportunities to educate and train the next generation of infectious disease leaders. For doctoral students and postdoctoral fellows, it will provide education on data-driven laboratory and computational biology methods that are being developed by UCLA faculty and LANL scientists. For visiting scientists and officials, it will provide first-hand training on how to build and operate high-throughput facilities and how to organize and manage near real-time surveillance and research efforts. These unparalleled opportunities will invite partnerships and collaborations with governments, foundations, businesses, and academic institutions.

Laboratory Response Network and National Critical Infrastructure. The Department of Health and Human Services (HHS), Centers for Disease Control and Prevention (CDC), have established the Laboratory Response Network (LRN) to ensure effective laboratory responses to bioterrorism or major infectious diseases outbreak (44). At present, over 130 laboratories are members of this network. In addition, HHS has established the National Infrastructure Protection Plan (NIPP) to strengthen the United States' critical infrastructure and key resources. The Plan is based, in part, on public-private partnerships and makes prudent security investments throughout the country (45). The high-throughput laboratory

is a strong candidate for membership in these two federal programs, which we are seeking to join before our UCLA facility opens in May 2009. In the event of an influenza pandemic or major outbreak, sustained 24/7 operations by the high-throughput laboratory will require the following resources that can be better assured by LRN/NIPP memberships.

- Functional transportation services and infrastructure.
- Functional communications infrastructure.
- Functional power infrastructure.
- Adequate and timely influenza sample streams.
- Adequate and timely funding to purchase reagents, materials, and supplies.
- Adequate and timely reagent, material, and supply streams from vendors.
- Adequate personnel who are trained and available for pandemic mode operations.
- Adequate maintenance support by vendors in case of laboratory system failures.

High-Throughput Laboratory Network. Establishing a network of high-throughput laboratories worldwide will present technical, social, political, and economic challenges. The experience gained from building and operating the first high-throughput laboratory at UCLA will provide the technical know-how. But it will not overcome the various social, political, and economic challenges that are posed by our paradigm shifting efforts.

As shown in **Figure 10**, it seems reasonable to imagine that high-throughput laboratories would be established at geostrategic locations and address the region's need for health security. In the Americas, another high-throughput laboratory would likely be located at a major U.S. Government facility, Canada, or South America. In Asia, where new viruses often emerge, such locations would include Singapore, Hong Kong, India, or Japan for example. On other continents — Africa, Europe, and the Middle East — the locations and infectious diseases focus remain to be determined.

In moving forward, we are seeking partnerships, collaborations, and support to address several key questions. Who builds and operates a network of high-throughput laboratory facilities? Who maintains and updates standards between facilities? What are economic model(s) to sustain these facilities? What role(s) do and/or should governments, foundations and academic institutions play?

Figure 10. Next site in the high-throughput laboratory network.

Cited On-Line Materials and References

1. World Health Organization. International Health Regulations (2005). http://www.who.int/csr/ihr/en/
2. The White House. Homeland Security Presidential Directive/HSPD-21 (October 18, 2007). http://www.whitehouse.gov/news/releases/2007/10/20071018-10.html
3. Google.org. Predict and Prevent. http://www.google.org/predict.html
4. Los Alamos National Laboratory. National Security Mission. http://www.lanl.gov/natlsecurity/
5. UCLA School of Public Health. Awarded $9 Million Grant from the State of California for New Biological Lab (September 14, 2006). http://www.ph.ucla.edu/pr/newsitem091406.html
6. NIH/NIAID. Centers of Excellence for Influenza Research and Surveillance (CEIRS). http://www3.niaid.nih.gov/research/resources/ceirs/
7. CUH2A. http://www.CUH2A.com
8. PCL Construction. http://www.pcl.com/
9. HHS, CDC, NIH. Biosafety in Microbiological and Biomedical Laboratories (BMBL) 5th Edition. Washington, DC: US Government Printing Office (2008). http://www.cdc.gov/OD/ohs/biosfty/bmbl5/bmbl5toc.htm
10. CDC. Select Agent Program. http://www.cdc.gov/od/sap/
11. Layne, SP. CRISAR Pandemic Research Response Plan: Working Draft (February 25, 2008). http://www.ph.ucla.edu/epi/layne/PUBLICATIONS/PRRP2008.pdf
12. Layne SP, et al. A Global Lab Against Influenza. Science 2001; 293: 1729. http://www.sciencemag.org/cgi/reprint/293/5536/1729.pdf
13. Layne SP. Human Influenza Surveillance: the Demand to Expand. Emerging Infectious Diseases 2006; 12: 562-569. http://www.cdc.gov/ncidod/EID/vol12no04/pdfs/05-1198.pdf
14. Layne SP, Beugelsdijk, TJ. Laboratory Firepower for Infectious Disease Research. Nature Biotechnology 1998; 16: 825-829. http://www.ph.ucla.edu/epi/layne/PUBLICATIONS/NBT.pdf
15. Tecan AG. http://www.tecan.com/
16. REMP AG. http://www.remp.com/
17. Staab, T. AvianTracker.org Resource Site. http://aviantracker.org/
18. Owens, M. Bird in the Hand. Los Alamos National Laboratory Currents 2008; May: 4-5. http://www.lanl.gov/news/currents/docs/currents_05_2008.pdf
19. Foiling the Flu Bug. 1663 Los Alamos Science and Technology Magazine 2007; January. http://www.lanl.gov/science/1663/flu.php
20. Service Oriented Architecture. http://en.wikipedia.org/wiki/Service-oriented_architecture
21. Web Services Description Language. http://en.wikipedia.org/wiki/Web_Services_Description_Language
22. SOAP. http://en.wikipedia.org/wiki/SOAP
23. Extensible Markup Language. http://en.wikipedia.org/wiki/XML

24. STARLIMS corp. http://www.starlims.com/
25. Health Insurance Portability and Accountability Act. http://www.hhs.gov/ocr/hipaa/
26. Cloud Computing. http://en.wikipedia.org/wiki/Cloud_computing
27. Layne SP. California's Level of Preparedness for an Outbreak of Avian Influenza. Testimony to the Chair of the Assembly Budget Committee on Health and Human Services California Assembly (November 4, 2005). http://www.ph.ucla.edu/epi/layne/TESTIMONY/Layne_Testimony.pdf
28. Layne SP, Beugelsdijk TJ, Patel CKN (editors). Firepower in the Lab: Automation in the Fight Against Infectious Diseases and Bioterrorism. Washington, DC: Joseph Henry Press (2001). http://books.nap.edu/catalog.php?record_id=9749
29. Suchard M, et al. Bayesian selection of continuous-time Markov chain evolutionary models. Molecular Biology and Evolution 2001;18:1001-1013. http://mbe.oxfordjournals.org/cgi/reprint/18/6/1001
30. Suchard M, et al. Hierarchical Phylogenetic Models for Analyzing Multipartite Sequence Data. Systematic Biology 2003; 52: 649-64. http://www.jstor.org/pss/3651067
31. Kitchen CMR, et al. Evolution of Human Immunodeficiency Virus Type 1 Coreceptor Usage during Antiretroviral Therapy: A Bayesian Approach. Journal of Virology 2004; 78: 11296-302. http://jvi.asm.org/cgi/reprint/78/20/11296
32. Kitchen CMR, et al. Two-way Bayesian hierarchical phylogenetic models: an application to the co-evolution of gp120 and gp41 during partial treatment interruption of enfuvirtide. Computational Statistics and Data Analysis (June 12, 2008). http://www.sciencedirect.com/science?_ob=ArticleListURL&_method=list&_ArticleListID=783667284&_sort=d&view=c&_acct=C000050221&_version=1&_urlVersion=0&_userid=10&md5=961ec0b3ca9ffd4e69155e7bea289bc3
33. UCLA Institute of the Environment – Center for Tropical Research. http://www.ioe.ucla.edu/CTR/research/AvPath/avian_influenza_main.html
34. Yamada T, et al. Ready for avian flu? Nature 2008; 454: 162. http://www.nature.com/nature/journal/v454/n7201/pdf/454162a.pdf
35. Vanderwagen, WC RADM. Safeguarding our Nation: HHS Emergency Preparedness Efforts. Testimony before the House Committee on Homeland Security Subcommittee on Emerging Threats, Cybersecurity, Science, Technology and Government Affairs (July 22, 2008). http://www.homeland.ca.gov/pdf/testimony/house/hhs/072208/Testimony_Vanderwagen-USDepartmentofHealthandHumanServices.pdf
36. Pacific Biosciences. http://www.pacificbiosciences.com/
37. Biomedical Advanced Research and Development Authority (BARDA). http://www.hhs.gov/aspr/barda/
38. CDC. Selecting the Viruses in the Influenza (Flu) Vaccine. http://www.cdc.gov/flu/professionals/vaccination/virusqa.htm
39. Ziegler T, et al. Type- and subtype-specific detection of influenza viruses in clinical specimens by rapid culture assay. Journal of Clinical Microbiology 1995; 33: 318-321. http://jcm.asm.org/cgi/reprint/33/2/318
40. WHO. Epidemic and Pandemic Alert and Response. http://www.who.int/csr/disease/influenza/vaccinerecommendations/en/index.html
41. Kale RR, et al. Detection of Intact Influenza Viruses using Biotinylated Biantennary *S*-Sialosides. Journal American Chemical Society 2008;130: 1869-1871. http://pubs.acs.org/cgi-bin/sample.cgi/jacsat/2008/130/i26/pdf/ja800842v.pdf?isMac=425994
42. Editorial. Community cleverness required. Nature 2008; 455: 1.
43. California NanoSystems Institute (CNSI). http://www.cnsi.ucla.edu/
44. Laboratory Response Network Partners in Preparedness. http://www.bt.cdc.gov/lrn/
45. National Infrastructure Protection Plan. http://www.hhs.gov/aspr/opeo/cip/nipp.html

###