

2<sup>nd</sup> Annual National Academies Keck *Futures Initiative* Conference  
*Designing Nanostructures at the Interface  
between Biomedical and Physical Systems*  
Arnold & Mabel Beckman Center, Irvine, California  
November 18-21, 2004

**A Micro System to Isolate, Sequence, and Identify DNA from a Small, Low-concentration Sample**  
**Focus Group Summary**

*Summary written by:*

Susan Brown, Science Communication Program, University of California – Santa Cruz

Focus group members:

- Rigoberto Advincula, Associate Professor, Department of Chemistry, University of Houston
- Rene Baston, Vice President, Business Development, New York Academy of Sciences
- Susan Brown, Graduate Student, Science Communication Program, University of California - Santa Cruz
- Yury Gogotsi, Professor, Department of Materials Science and Engineering, Drexel University
- Kimberly Hamad-Schifferli, Assistant Professor, Department of Mechanical Engineering & Biological Engineering Division, Massachusetts Institute of Technology
- Yue Kuo, Dow Professor, Department of Chemical Engineering, Texas A&M University
- Shuang Fang Lim, Postdoc, Princeton University
- Liviu Movileanu, Department of Physics, Assistant Professor, Syracuse University
- Robert Riehn, Research Associate, Princeton University
- Holger Schmidt, Department of Electrical Engineering, Assistant Professor, University of California – Santa Cruz
- Joel Schnur, Director, Center for Biomolecular Science and Engineering, Naval Research Laboratory
- Xing Su, Senior Staff Scientist, Intel Research, Intel Corporation

**Summary:**

Genomic information is expanding on an unprecedented scale. Pathogens, new organisms, and diseases are being identified on the basis of genetic sequences.

If genomic information could be harvested quickly and inexpensively, it could be applied widely in beneficial, practical applications.

If inexpensive, fast, and ubiquitous, the technology could be used to identify specific strains of illness-causing pathogens. In a clinic, for example, treatment could be immediately tailored to the infection. If cheap and robust, detectors could be placed in public places or carried into battlefields where they could identify harmful pathogens quickly enough to prevent widespread infection.

Imagining a way to realize this potential became the focus for our group. We wanted to design a device that would take an unprocessed biological sample – from the air, soil, a patient, a crime scene, ancient bone – and, within a single device, process, detect, and identify the genetic material found in the sample. To make the technology widely adoptable, we also aimed for a solution cheap enough to be disposable and as easy to use as a home pregnancy test.

By the end of the first session, our goal became “the unambiguous identification of any unknown DNA and/or RNA within minutes from a native sample at low cost.”

Currently, samples can be screened for known pathogens or genetic sequences in a matter of hours to days. The costs remain high, particularly for sequencing, at hundreds of dollars per sample, with initial capital investment for equipment in the tens of thousands.

## **Circling in on a solution**

Over the course of four sessions, we proposed a solution that combines current state-of-the-art technologies to achieve the goal.

Defining a solution was an iterative process, but the outline of what could be achieved was developed early. From there, we identified the bottlenecks in the process to determine how the goals could be accomplished more quickly and at lower cost.

Early in the process the group divided the problem into four basic steps:

- • separate nucleic acids from the sample
- • sequence or detect the nucleic acids
- • process the sequence information
- • report the outcome

## **Separating nucleic acids from a gunky sample**

Extracting genetic information quickly and cheaply from unpurified sample with no preprocessing using a single device challenged the group. It requires the ability to handle gunky samples, such as mucus or soil. The device would also have to separate all the other cellular components from nucleic acids.

Early in our discussion, we decided against amplifying the DNA because the required primers and thermocycling would greatly increase cost and time. New technologies that promise to sequence individual molecules of nucleic acid are on the horizon.

We considered a wide variety of solutions, such as binding nucleic acids to magnetic beads and sucking them out of solution, or using a microfluidics array in which posts deflect and slow larger molecules in a sort of nanoscale pachinko machine to fractionate cellular components by size.

In the end, we decided to minimize fluids and instead use multi-layer thin films with enzymes and other reagents impregnated in each layer to initially separate nucleic acids from other components of the sample. The sample would pass through a filter under pressure to remove macromolecules, and then through a film impregnated with enzymes, such as lysozyme, to open the cells.

The final film would contain chaotropic salts, just before the sample passed through to a binding matrix material, such as a glass fiber mat. In the presence of chaotropic salts, DNA binds to glass fibers. Other cellular components would then be eluted using a valved microfluidics device to pump ethanol through to rinse waste into a separate chamber, followed by water or a buffer solution to elute the DNA from the binding material.

The DNA could be fragmented for faster parallel sequencing by passing it rapidly through narrow microchannels to shear it into randomly sized pieces. The original sequence would be recovered by assembling the sequences of overlapping fragments.

## **Rapidly sequencing small amounts of unknown DNA**

This is a hot field with both private industry and government agencies, such as DARPA and NIH, investing millions in its solution. We assumed rapid development of each of these sequencing options would lead to several fast and inexpensive options. To expand on the group's expertise, we imported several visitors, including Andrew Ellington, to share specialized knowledge in various areas.

We focused on two similar approaches based on the sequencing of overlapping fragments of DNA.

One option would be to sequence fragments of DNA that are immobilized on a solid surface. Biochemical reactions would be used to read the bases in each individual molecule. For example, DNA polymerase or exonuclease could be engineered to generate an optical signal as each base is added or deleted.

The sequence information could be recorded for all the DNA fragments simultaneously using a sensitive

charge coupled device (CCD) to detect the optical signals. The polymerase and exonuclease reactions are fast, about 10,000 bases per minute. The challenge will be to design a detection system that can operate at the speed of the reaction.

Alternatively, the DNA could be passed through nanochannels and the sequences detected electrochemically. An array of nanochannels could be used to increase efficiency.

Finally, resequencing chips that hybridize DNA are currently available. But the chips require known sequences, and processing takes hours to days. The processing time is likely to fall with incremental improvements in the technology. Resequencing chips will be most useful for focused applications that discriminate between several known options, such as different strains of influenza virus.

### **Identifying sequences**

For focused applications, in which sequences would be compared to a known and limited set, a lexicon could be stored on a chip. In that case, sequencing and identification could occur in the same step.

In the more ambitious case, in which the sample is completely unknown, the sequences may need to be compared to a database, such as GenBank, using a remote device. This last part would likely be a small reusable radio-frequency or Wi-Fi handheld instrument in the field, a desktop computer connected to the Internet in a clinic or laboratory, or a satellite communication device in a remote place.

### **Reporting signals**

The end users will need to know what to do once the sequence is identified. For home-test applications the message needs to be simple and clear. Group members envisioned tissues that, when you blow your nose into them, revealed a message based on your illness, such as “go to the emergency room,” “call your doctor,” or “have chicken soup.” The same device that analyzed a nasal swab could pop out a pill tailored to the particular strain of pathogen, for example.

Sensors deployed in public places could be programmed to deliver similar messages warning people not to enter an area if intentional release of a pathogen is detected.

### **Gaps between current science and technology and realizing this vision**

Current sequencing methods are not sensitive enough to allow rapid and reliable sequencing of single molecules of DNA. If more sensitive detection systems are developed, such as the ones we have proposed here, the high cost of DNA amplification by PCR could be eliminated. Highly parallel sequencing of single DNA molecules is the key to achieving our goal.

Success will require other technological advances as well. For example, the multi-layer films suggested for the first step have excellent permeability and filtering properties, but still need to be optimized to work in the robust, field-based kits the group has proposed.

Finally, despite rapid advances, the database of gene sequences remains incomplete. It is also inaccurate, containing wrong and sometimes mislabeled sequences. More information about sequences and better retrieval schemes will allow the system to find answers in an imperfect database.

### **Potential benefits and dangers to society**

**Information** – More ubiquitous sequencing would help to clean up the genetic databases. If widespread sensing of DNA sequences reveals something novel, the sequence could be submitted to a temporary database. If confirmed, for example by additional sensing, it could be added to the database. Errors in sequences could also be corrected if a mechanism for comparing the differences detected in the field were incorporated. Linking development of this new technology to mechanisms for checking, correcting, and updating the database could rapidly accelerate the acquisition of new genetic information.

**Health** – The greatest potential benefit would be a vast improvement in human health and safety. Early detection of infectious agents could lead to an end to epidemics, even an end to infectious disease. Rapid and accurate diagnosis, especially in a home-based kit, would minimize the impact of minor illness

for the individual, but also for society in fewer lost workdays and a reduced burden on the health care system.

Deployment of sensors in public places or gatherings that might be vulnerable to biological attack could lead to early detection and save lives.

These benefits may come at a cost. Identifying pathogens is of no use if treatments are not available. The greater good to society will come only if most people with infectious disease comply with treatment. Would we force people to comply? A conflict between society and individual rights may arise.

Finally, fast and inexpensive sequencing will lead to the sequencing of individuals' genomes. Patients' own cells will be included in any biological samples, thereby making this likely. Good will come of that if drug treatments can be tailored to their individual genetic makeup – relieving patients of treatments with harmful side effects that are unlikely to work and more rapidly identifying those most likely to do good.

But genetic information about individuals could lead to discrimination in employment or lack of access to medical insurance. Genetic information could be psychologically harmful to individuals if there is nothing they can do about their condition. And doctors who withhold information from their patients for that reason may be vulnerable to liability suits.

If scientists keep these concerns in mind as they develop this new technology, safeguards can be put into place as it is implemented. Doing so will go a long way toward reassuring the public and preventing the kind of opposition that might obstruct the introduction of a potentially widely beneficial new approach.