BioCoder

DIY/BIO NEWSLETTER   JULY 2015

DIY Open Source Biomaterials
Daniel Modulevsky & Andrew E. Pelling

Homegrown Neuroscience: Backyard Brains and the RoboRoach
Glen Martin

Open Source Biomaterials for Regenerative Medicine
Daniel Modulevsky, Charles M. Cuerrier, & Andrew E. Pelling

Molecular Diagnostics on the Smartphone: The Mobile Health Revolution
Marc DeJohn

Did you enjoy this issue of BioCoder?

Sign up and we'll deliver future issues and news about the community for FREE.

http://oreilly.com/go/biocoder-news

BioCoder #8

JULY 2015

Contents

Foreword

Nina DiPrimio

Following the Synthetic Biology Leadership Excellence Accelerator Program (LEAP) showcase, I met with fellows Mackenzie Cowell, cofounder of DIYbio.org, and Edward Perello, cofounder of Desktop Genetics. Mac and Eddy both wanted to know what processes in laboratory research are inefficient, and how do we eliminate or optimize them? Pairing software developers and hardware engineers with biologists in academic labs or biotech companies to engineer small fixes will result in monumental increases in research productivity.

An example of an inefficient lab process that has yet to be automated is fruit fly manipulation. Drosophila handling and maintenance is laborious, and Dave and Matt Zucker from Flysorter are developing a technology to automate these manual tasks. This technology uses computer vision and machine learning software to do so. They are currently engineering prototypes, and this is a challenging problem to solve, but this is a perfect of example of engineers developing a technology to automate a completely manual and extremely tedious laboratory task. Check them out, and stay tuned for an article from them in the October issue of *BioCoder*.

This issue of *BioCoder* highlights a few examples of collaborations between individuals with complementary expertise to provide solutions for problems across diverse areas of biological research. Biomeme is a smartphone-based diagnostic for the on-site detection of DNA. ABioBot is a smart robot, using vision, sensing and feedback to automate encodable laboratory experiments. The Pelling Lab provides information on open source biomaterials as well as a tutorial on how to make your own scaffolds for tissue engineering, sprinkled with some excellent hardware hacks. These projects were developed through collaborations between hardware engineers, software engineers and biologists.

Biology researchers, what laboratory processes cause you distress? Is it your data analysis tool or lack their of, experiment design, expensive reagents, tracking of experiment information, a tedious technique that could benefit from a little

automation, or a piece of hardware that could use some customization? Tweet specific examples to @100ideas (Mac Cowell), @EdwardPerello or @NinaDiPrimio. We want to hear from you.

Homegrown Neuroscience: Backyard Brains and the RoboRoach

Glen Martin

Is there a single development in the DIY realm cooler than the cyborg bug? There are certainly legitimate competitors: glowing plants, for example, and human cheese. But there's something about seeing a big death's head cockroach (*B. discoidalis*) outfitted with a minute, circuit board backpack and responding to cues from a human overseer via smartphone that makes your jaw drop. Could it be that our command of the entire insect kingdom is at hand?

Of course not. But it's a damned neat trick nonetheless. What's more, it's one that any intelligent and inquisitive person with average digital dexterity can accomplish, thanks to the handy RoboRoach Surgery Kit from Backyard Brains, an educational products firm focused on DIY neuroscience.

The package, explains the company's cofounder and engineer Tim Marzullo, contains the PCB back, which carries the Bluetooth low-energy wireless receiver necessary for smartphone communication; an assortment of resistors, capacitors, and light-emitting diodes; a small battery; and three sets of electrodes (sufficient to implant three cockroaches). Each electrode has three wires: one is used for the right antenna, another for the left antenna, and the third, functioning as the ground, is (gently) inserted into the dorsal surface of the insect's thorax.

Cockroaches, says Marzullo, have neurons quite similar to that of human beings, just fewer of them. They cluster in decentralized ganglia running down the body (a larger ganglion located in the head serves as the insect's "brain").

"Implanting the tips of the electrodes into the antennae creates a neural interface," says Marzullo. "That allows you to send electrical signals similar to the ones naturally produced by the sensory neurons in the roach's antennae."

The signals sent via the PCB induce the neurons to fire electrical messages, known as "spikes" in neuro-speak, to the ganglia running down the roach's body. By stimulating the neurons on the right antenna, you can make the roach turn left, and vice versa for the left antenna. Voilà: the roach will do your bidding. Resistance is futile.

"It's our most popular product," Marzullo says of the cyborg bug kit. "We have a RoboRoach TED talk up on our site that has had over two million hits."

But the RoboRoach, of course, is more than a neat parlor trick involving a hefty insect. It's a means to a larger end: teaching neuroscience to the young, from grade-school students through college undergraduates.

"I met our cofounder, Greg Gage, at the University of Michigan," says Marzullo. "He's an electronics engineer, and I'm in biotech, but we were both deeply involved in neuroscience. We often found ourselves at outreach events with school kids who were eager to learn what we do, and we wanted to demonstrate spiking activity, but we had no equipment. Its cost was way too high. So we began using off-the-shelf electronics to build kits that were both easy to use and illustrated how nervous systems work."

Along with the RoboRoach kit, Backyard Brains sells several other DIY neuroscience products, among them the SpikerBox, a device that enables users to record cockroach neuron spikes; the EMG SpikerBox, an advanced version of the original SpikerBox that records neuron spikes in humans; the RoachScope, which allows students to prepare quality slides for high-powered magnification; the Heart & Brain SpikerShield Bundle, a kit that lets experimenters record their own electroencephalograms; and perhaps most spookily (cue *Twilight Zone* theme), the Human/Human Interface Bundle, a package that empowers the experimenter to employ muscle signals to contract the muscles of another human being via a device known as a transcutaneous electrical nerve stimulator. All the devices employ open source hardware and software.

Not content with exploring the neural activity of the animal realm, the company has begun developing products that address plant electrophysiology. The Plant SpikerShield Bundle lets you record the electrical spikes of a Venus flytrap.

Backyard Brains is not yet in the black, but thanks to an unorthodox fiscal strategy, it does have sufficient operating revenue. Funding is balanced 50/50 between grants and sales, with 10% of the grant money coming from Chile and 90% from the NIH in the U.S. (The company also is "open source" about its finances, posting its books on its website.)

"We started receiving our Chilean support about three years ago," says Marzullo. "After the big Chilean earthquake of 2010, the government decided to diversify the country's economy, which was mostly resource-based—copper, agricultural products, fish. So they began funding some startup tech programs."

Marzullo and Gage made the cut, and the company is now bi-continental, operating in offices from both Santiago and Ann Arbor.

"The good thing about working in Chile is that it's easier to get government funding," says Marzullo, who's stationed in Santiago. "We did OK with NIH support, but the competition in the U.S. is incredibly fierce. In Chile, tech start-ups usually get 30%–40% of their seed funding, compared to 10%–15% in the United States. There are also a lot of very talented and motivated engineers in Chile."

And the downside? Marzullo laughs ruefully.

"The financial system is fairly unsophisticated, and basic materials can be hard to get. There's no Amazon here, no ordering something and receiving it the next day. There's only one online store. It was established in the 1990s, and it's not terribly efficient."

For the foreseeable future, Backyard Brains intends to keep doing what it does best, says Marzullo: providing affordable devices that make neurological processes compelling and understandable.

"Our primary goal is to provide the means for recording the interface of any biological electrical signal, whether it's generated by a cockroach, human muscle, or plant," says Marzullo. "A secondary goal is designing devices that allow you to exploit those signals to control other systems, such as we're doing with our Human/Human Interface product. We think that's a particularly interesting space."

Glen Martin covered science and the environment for the San Francisco Chronicle for 17 years, and has contributed to more than 50 magazines, including Discover, Science Digest, Wired, Audubon, and the Utne Reader. His latest book, Game Changer: Animal Rights and the Fate of Africa's Wildlife, was published by the University of California Press.

Molecular Diagnostics on the Smartphone: The Mobile Health Revolution

Marc DeJohn

When my cofounders and I launched Biomeme three years ago, we did so with an eye on a number of converging trends. Among them, the rapid adoption of mobile technology, a flourishing DIY culture, and the serious shortcomings of our centralized healthcare model. In particular, a conspicuous lack of diagnostic tools in both the home and the doctor's office. Sending samples out to labs for test results has been a staple of our healthcare paradigm. At Biomeme, we asked the question, "Why not build the lab around the smartphone so that diagnostics can be performed by anyone, anywhere?"

So we did just that, creating a battery-powered, real-time thermocycler that docks with an iPhone 5s and enables users to perform molecular diagnostics from sample to answer virtually anywhere. For anyone with an iPhone, it's essentially a lab in a box. Everything needed to complete a given test is in a kit, which consists of a small, inexpensive instrument and some disposables. Our vision is to have the system cost no more than an electric razor, toothbrush, or game console.

The technology is based on real-time PCR, or polymerase chain reaction, a biochemical process used to amplify target DNA sequences. PCR requires a machine that precisely heats and cools the DNA or RNA to achieve replication. For a tutorial on PCR, see the *Biocoder #5* article "PCR for Everyone, Everywhere" by Ezequiel Alvarez-Saavedra and Sebastian Kraves. For a tutorial on real-time PCR, see the *Biocoder #6* article "Open qPCR" by Josh Perfetto. In our case, the hardware that effectively replicates DNA or RNA is attached to a smartphone, which itself becomes part of the processing unit that conveys results. We're work-

ing with partners in human health, biothreat surveillance, education, and research; and we see possibilities in forensics, food safety, environmental monitoring, and more.

DIY Tests and Consumer Demand

It's been nearly 40 years since the home pregnancy test was introduced commercially, and few would argue the success of that particular "DIY" test[4] both in terms of popularity and as a moment of emancipation from a paternalistic healthcare system. Nevertheless, beyond the early pregnancy test (e.p.t), very few tests have become available for consumers to perform in the convenience and privacy of their home.

Some would argue there isn't a valid need or demand for self-performed molecular diagnostics. But the same was said for the e.p.t and many other technologies that have changed our lives for the better. A recent survey conducted by Ipsos and the National Council on Patient Information and Education (NCPIE) found that a majority of consumers are willing to take greater ownership of their health. Of 2,024 U.S. adults and 516 primary care physicians, "92 percent [of respondents] said they like being in control of their health and 80 percent said they feel expected to manage their health more now than ever before." In a statement, NCPIE Executive Vice President Ray Bullman said, "In addition to working with healthcare professionals, individuals are actively seeking ways to treat, diagnose, and manage their health conditions and want the information and tools to do it better."[3]

This response is certainly due, in part, to the challenges facing patients when accessing healthcare at the clinic. The fact remains that, resources there are in short supply, and the coming years will strain them further. However, with the right tools, patients can begin to fill some of the responsibilities once owned by the clinician. Those tools are emerging, and they're doing so around the smartphone.

Health Data at Your Fingertips

Nearly two-thirds of U.S. consumers own a smartphone, and about the same percentage of those smartphone owners have used their phone in the past year to look up information about a health condition.[1] Accessing data in this manner is nearly seamless. Indeed, some of you may be reading this article from a smartphone. If so, author Dr. Eric Topol might say, "The future of medicine is in your hands." In part, his message is this: much of our health will be administered around our smartphones, not at the clinic. Not just to look up information, but to

generate it by way of sensors and self-performed tests. In a TV interview with Jim Cramer, Apple's CEO Tim Cook referred to the mobile health market as the biggest frontier of all: "for years people have depended on, strictly, somebody else to determine their health and now these devices in essence empower people to manage and track their own health and fitness."[5]

For consumers today, there are a handful of health-related hardware add-ons that can be found on the shelf at the Apple store, among other retailers. This includes blood pressure cuffs, heart rate monitors, oximenters, scales, air quality meters, thermometers, sleep trackers, and activity trackers. However, a number of other medical devices made for the smartphone are available elsewhere or will be in the near future: electrocardiograms, spirometers, otoscopes, ultrasounds, eye exam devices, microscopes, glucometers, and even laryngoscopes. Beyond the glucometer, some products are addressing blood and urine chemistry by aggregating off-shelf colorimetric assays and using the phone camera to make a diagnostic call. And there are dozens of other projects that are currently under development.

It's clear that this "Internet of Medical Things"(IOMT)[1] is growing at a rapid pace. What this represents to DIY health, concierge medicine, and clinical practices around the world is exciting to ponder. Nevertheless, with all this mobile medtech, there are still a lack of devices that can perform the gold standard in molecular diagnostics: nucleic acid–based testing using real-time polymerase chain reaction (PCR).

The Central Lab Hurdle

PCR and real-time PCR have long been the standard testing methods for the identification of a spectrum of microbial, viral, and genomic targets. Whether they are disease causing or beneficial, the markers for these targets are segments of DNA or RNA unique to those targets.

When speaking of human diagnostics, these targets could be found in a variety of sample types. A pathogenic strep bacterium could be collected by swabbing the back of the throat, a cancer marker found in the blood through a blood draw, or a symbiotic gut microbe found in the stool. Just about any sample type carrying target DNA or RNA can be scrutinized with this method.

Currently, most diagnostic tests of this nature are run in central labs on large, high-throughput machines, with queues of samples collected from regional clinics. Health workers ship patient-derived samples to those sites for processing. It is an efficient process in terms of cost per sample, but there are trade-offs.

It is often days to weeks before results are returned to a doctor or patient. This delay often presents real problems, especially when results are actionable. When decisions concerning therapeutics or patient behavior are relevant, the wait for

results just isn't practical. For example, antibiotics are often prescribed with a presumptive diagnosis before a positive lab result verifies a suspect bacterial infection. Often the drugs are unnecessary or ineffective. This practice is responsible, in part, to the serious rise in antibiotic-resistant strains of pathogenic bacteria. Additionally, there is evidence now that these very medications can be harmful to the patient's microflora.[4] For this reason, conscientious physicians have begun to scale back on prescribing antibiotics and are relying instead on diagnostic tests. While this issue highlights the need for point-of-care solutions, there is something more that is commonly lost in this scenario: an ownership of the process and the data, and with it, patient engagement.

At Biomeme, we understand the value of tracking your own health history, not just punctuated diagnostic events. Understanding a multitude of correlated trends and events throughout our medical lives will bring valuable discoveries in personal health. This effect is more pronounced when we speak of pooling that data from populations and understanding the communal health picture. In short, to realize the power of the IOMT, an ethos of personal ownership and sharing of data must be supported. The path toward this vision might best be laid by DIYers.

A Smartphone Solution: Mobile MDx

It's important to know that a nucleic acid–based molecular diagnostic test typically requires a few steps and components:

1. The collection of a sample

2. The extraction of all nucleic acids (NAs) from that sample

3. An assay which contains the chemistry specific to the target(s)' NAs

4. An instrument to run those assays on

Our initial product lineup includes tools for assay developers as well as field users. We have:

- A mobile, real-time PCR thermal cycler that syncs up to your iPhone

- Software that runs the reaction, makes a diagnosis in real time on the device, and communicates the results wirelessly to our cloud-based data management system

- Our own sample prep to isolate nucleic acids from crude samples

- Various reagents that developers may need, such as master mix, buffers, and PCR tubes (see Biomeme's website for full descriptions)

SAMPLE COLLECTION AND PREPARATION

Isolate DNA and RNA targets from a variety of samples with our portable, disposable sample preparation system (see Figure 2-1). The kit is comprised of a syringe, extraction column, and buffers. The kit requires no incubation, alcohol precipitation, or extraction using phenol or chloroform. Instead, it utilizes a filtration-based method in which NAs selectively bind to the silica membrane inside Biomeme's Sample Prep Column. The column fits on the end of a syringe. This allows cycling through a sequence of specially formulated buffers to produce purified NAs upon elution in water. The procedure takes just a couple minutes. For a quick demo of the protocol, see my cofounders' 7-year-old daughter perform the procedure. See technical results in the appendix for results on human DNA and bacterial targets in urine.

For a list of available kits, check out the Biomeme Store.

THE ASSAY

Biomeme is offering a test portfolio of hundreds of developed assays for human health, including targets for infectious disease, pharmacogenomics, nutrigenomics, and microbiota. Those interested in assays for environmental quality, food safety, veterinary, agriculture and biothreat will also find a variety of choices.

We embrace an open architecture for our developers. So for the researcher looking to develop her own protocols on Biomeme's system, we offer empty bulk three-well strips (Figure 2-2). When ready, we help with filling and lyophilizing services to get quantities of your assays manufactured. Those interested in developing or porting over their own tests are encouraged to do so. We aspire to build a community of users exploring their biological world. An open platform puts the power in your hands. It also lowers our costs and streamlines our community building to create new assays for our platform.

Figure 2-1. A portable sample preparation system

One caveat to this long list of test offerings is that they have not all been validated with our sample preparation technology. We are currently validating a sexually transmitted infections (STI) panel in a study approved by an institutional review board with a women's clinic network at Drexel University-Hahnemann Hospital. The panel tests for *Chlamydia trachomatis*, *Neisseria gonorrhea*, and *Trichomonas vaginalis* in urine. Other organizations are also developing their own tests for use on the Biomeme platform in a variety of industries and applications ranging from biothreat to nutrigenomics to K–12 STEM education. As this validation continues, we will offer clearer guidance on a greater variety of assays and sample types.

For a full listing of our offerings, see Biomeme's website.

THE REAL-TIME PCR THERMAL CYCLER

The instrument is a portable, two-color, three-well, real-time PCR instrument. It performs much like any other real-time instrument, except that it runs on an iPhone 5S with a simple, intuitive app (Figure 2-3). With our developer app, you are able to program your own protocols in just a few minutes or run preset protocols with just a finger tap. The device runs on battery and will complete 8–10 com-

mon 40-cycle PCR tests per charge. We built it around a 5S for our government customers, but future versions will be stand-alone instruments utilizing low-energy Bluetooth allowing any of the iOS family of devices to interface with and control the thermocycler (we anticipate future Android capability as well).

Figure 2-2. A package of PCR wax strips

SECURING HEALTH DATA

The device analyzes test results in real time. Data is stored on the iPhone and synced with Biomeme's cloud-based data management system when the iPhone has a WiFi or cellular connection (Figure 2-3). Biomeme's web portal and accompanying dashboard enables users to track and manage test results seamlessly from anywhere on any device. Users can also manage and view test results within the iOS application and export raw data as a *.csv* file via email or text. You can also export results as a *.jpg* file of the amplification plot with a text description. Results can be securely sent to third-party databases on a case-by-case basis. We use 128-bit encryption on all data transfer. Human health test data is stored in a HIPAA-compliant database. Our data management system is cloud based and easily integrated with other cloud-based systems. We work with customers to identify an

electronic health record (EHR) of choice and then go to that EHR to request an API for that customer's case.

Figure 2-3. The real-time PCR thermal cycler app

What's Next?

We're looking to improve our products, and we need your help. Currently, beta testers in government and academia have purchased devices and kits for field study of genomic, infectious disease and biothreat targets. But there's a lot more work to do, so we're looking for those that may have existing test assays or want to develop assays for our platform. In addition, we're interested in getting those assays validated on our sample prep. So we want to hear from you:

- **Are you a clinician?** What's your burning need? Have access to samples?

- **A distributor?** Where and how can you help us get to market?

- **A research professional?** How can you use us in your existing research or use us to take your lab research into the field?

- **A student or educator?** How can you bake us into your STEM curriculum? How can you use us to deliver the latest in advanced lab technology to your students?

- **An assay developer?** Do you have a market need? Let's work together to make that happen.

References

1. Eric J. Topol, *The Patient Will See You Now; The Future of Medicine is in Your Hands* (New York: Basic Books, 2015).

2. Martin Blaser, *Missing Microbes: How the Overuse of Antibiotics Is Fueling Our Modern Plagues* (New York: Henry Holt, 2015).

3. Aditi Pai, "Survey: 64 Percent of Consumers Say They Can Make More Health-Related Decisions Themselves," April 30, 2015, *http://bit.ly/pai-survey*.

4. The Office of NIH History, A Thin Blue Line: *The History of the Pregnancy Test Kit*, *http://bit.ly/ept-history*.

5. MobiHealthNews, "In-Depth: Apple ResearchKit Concerns, Potential Analysis," March 3, 2015, *http://bit.ly/mobihealth*.

Appendix
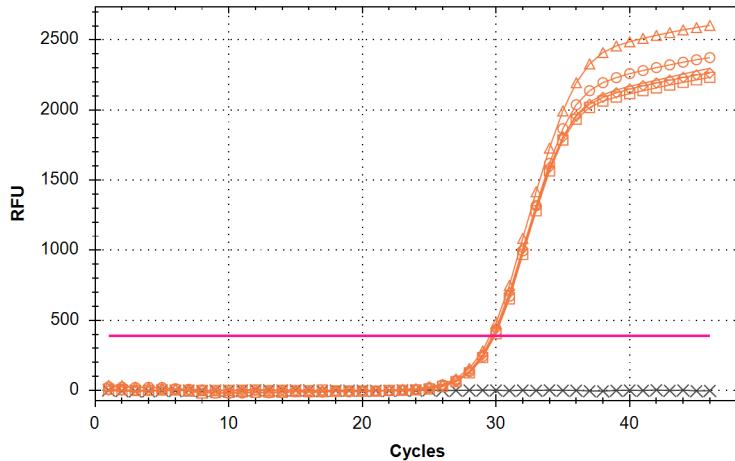
Figure 2-4. Amplification of a human gene target from urine shows high reproducibility
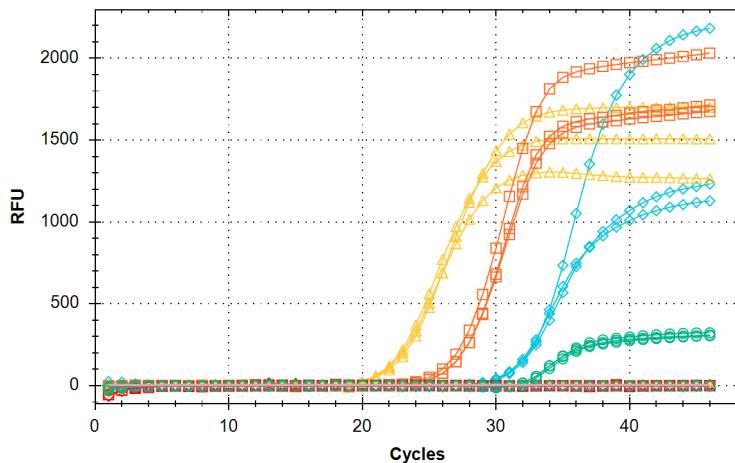
Figure 2-5. Three different bacterial genes (yellow triangles, blue boxes, green circles) and a human gene as an internal positive control were purified from human urine using Biomeme sample prep
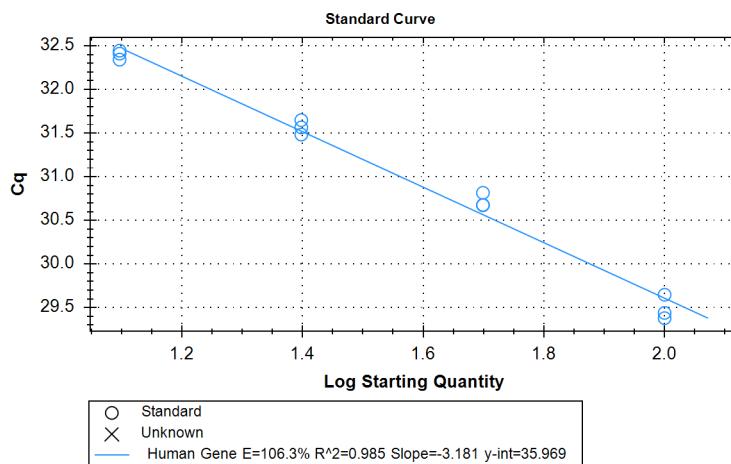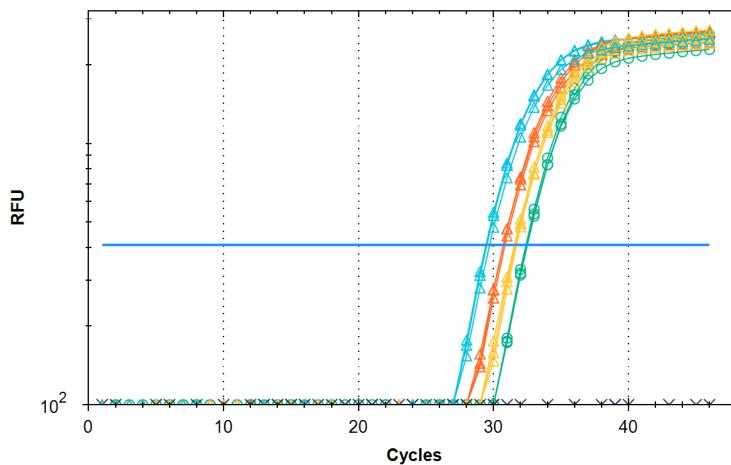
Figure 2-6. Purification of human genomic DNA with Biomeme Sample Prep is highly linear. Human genomic DNA was isolated from twofold dilutions of human urine and analyzed in triplicate using real-time PCR. The resulting standard curve shows strong linearity.

Marc DeJohn *is an engineering lead and cofounder at Biomeme, Inc. He earned his MS in electroanalytical chemistry at the University of Iowa and his BS in biology and chemistry from St. Ambrose University. Marc has spent more than 16 years working in research and development with an emphasis on small, portable technologies. His work spans the fields of electrochemistry, chemical actuators, acoustics, robotics, microfluidics, and more. Recently, Marc helped build the first acoustic cytometer, which was subsequently purchased by Invitrogen (now*

Thermofisher). Marc has coauthored a dozen patents, with the last three covering portable devices for nucleic acid amplification and detection.

Thank you to the team at Biomeme!

Open Source Biomaterials for Regenerative Medicine

Daniel Modulevsky,[a,b] Charles M. Cuerrier,[a,c] and Andrew E. Pelling[a, d]

In 1997, a striking image depicting a mouse with a human ear growing on its back rapidly became a viral sensation circulating the media around the world (Figure 3-1).[5,6] It was through the media that the "earmouse" quickly became associated with fears about Dr. Frankenstein-created laboratory monstrosities and fears about the limits of human experimentation and genetic engineering. In spite of the fears, the earmouse kindled many hopes about future possibilities in medicine. Unfortunately, the image of the earmouse has often been circulated without any context, which rapidly generates many false assumptions about the origin and intent of this animal experiment. In fact, the original research group led by Dr. Charles Vacanti at the University of Massachusetts Medical School, is often left unmentioned. What is significant about the earmouse model is that although it sparked much debate about genetic modification, it was not actually a genetically modified organism. Rather, the procedure to create the earmouse consisted of several key steps.[3] First, an alginate cast of an ear from a three-year-old child was created. Then, using the cast as a mold, a synthetic biodegradable polymer (polyglycolic acid, or PGA) was shaped into an ear-like structure. At this point, cells naturally found in cartilage (chondrocytes) were isolated from slaughtered calfs and implanted into the polymer ear. Over time, the chondrocytes degraded the polymer scaffold while gradually replacing it with a new biological scaffold (cartilage) in the form of a human ear. This resulting sample of an ear-shaped cartilage was then implanted into mice in order to examine its durability and biocompatibility.
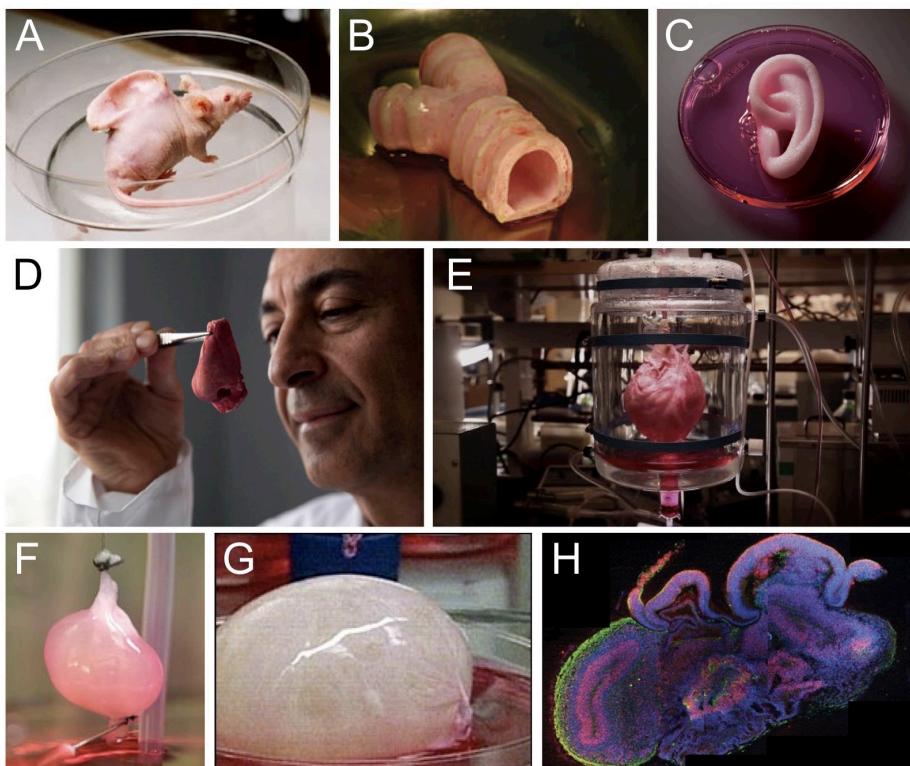
Figure 3-1. (A) The Vacanti earmouse.[1] Examples of other lab-grown organs and tissues using scaffolds of decellularized organs, including: (B) trachea[5] (C) ear[6] (D) nose,[7] (E) heart[8] (F) kidney[9] (G) bladder[10] and (H) model brain[11] Original image sources are provided in the relevant citations.[5–12]

Researchers in the fields of tissue engineering and regenerative medicine (RegenMed) often cite the promise held by the example of the earmouse. There is strong hope that interdisciplinary scientists will be able to harness advances in stem cell biology, polymer chemistry, material science, and 3D bioprinting to grow replacement body parts in laboratories to produce them at an industrial scale. Regardless of the fears and criticisms of the earmouse, the desire to regenerate defective body parts is expected to be at the center of the next-generation of medical treatments and therapies. In addition to creating molded polymer scaffolds,researchers have shown that it is also possible to remove all the cells from a donated organ, leaving behind the naturally occurring scaffold matrix. This process is known as decellularization and is a commonly employed technique to

produce "ghost organs." The ghost organs lack any of the cells from the donor and can be subsequently cultured with healthy cells derived from the patient or another source. Such regenerated organs have been shown to be functional and have even already been utilized in surgeries to repair defective tissues.[4] In the past several years, many body parts have been created using polymer-scaffold and decellularization approaches, including the trachea, ear, nose, heart, kidney, bladder, and brain (Figure 3-1).[5–11]

Commercially Available Biomaterials

The field of RegenMed is defined by a large spectrum of technical possibilities that share the same goal: the regeneration or replacement of defective organs. Approaches include the use of surgery, implants, and biomaterial scaffolds that will replace or regenerate the defective tissues or organs. This type of work is often complemented with the introduction of stem cells into the surgical sites. However, the development of novel biomaterials for RegenMed, which is the principal focus of this article, is undergoing exponential growth. Many commercial companies and universities around the world are driving this type of research. Advances in the field are being encouraged by many successful RegenMed treatments that utilize synthetic biomaterials and the significant potential for financial gain through lucrative intellectual properties and licensing opportunities. Some estimate the biomaterials market to reach a value between $30 billion and $90 billion in the next few years.[13,14]

   At the moment, there is an enormous array of biomaterial scaffolding products available commercially. These products vary considerably, from both biological and synthetic polymers, and are available in different forms such as powders, gels, membranes, and pastes (Table 3-1). Aside from creating replacement organs, such biomaterials are often employed to repair damage to various tissues and structures, such as skin, gum, cartilage, and bone. One of the defining characteristics of all these products is their high cost. The high prices are often due to several factors, such as the production procedures (which require specialized devices and sterile conditions), and the costs associated with research and development, licensing, and intellectual property.

Table 3-1. Commercial biomaterials and current pricing. Price ranges take into account any options of size format.

Product (company)	Material	Price[3] (USD/cm– USD/ml)
3D Insert™-PCL (3D Biotek)	Polycaprolactone	$30—$300
3D Insert™ (3D Biotek)	Poly(lactic-co-glycolic acid, or PLGA	$60
Optimaix 3D™ (Matricel)	Collagen type I/III	$40—$220
c-graft putty™ (Citagenix)	Demineralized bone matrix and carboxymethylcellulose (human and plant)	$150—$320
c-blast putty™ (Citagenix)	Demineralized bone matrix, cancellous bone, and carboxymethylcellulose (human and plant)	$150—$320
Neoderm (Citagenix)	Decellularized dermal grafts (human)	$260—$1,440
DynaMatrix™ (Citagenix)	Extracellular matrix (porcine)	$900—$1,980
DynaMatrix™Plus (Citagenix)	Extracellular matrix (porcine)	$623—$1,490
BioXclude™ (Citagenix)	Allograft amnion and chorion tissue (human)	$1,380—$3,170
Neomem® (Citagenix)	Resorbable collagen membrane (bovine)	$720—$1,660
DynaGraft D™ (Citagenix)	Demineralized bone matrix (human)	$200—$300
DynaBlast™ (Citagenix)	Demineralized bone matrix and cancellous bone (human)	$180—$365
Raptos® (Citagenix)	Cancellous particulate, cortical particulate, corticocancellous, and demineralized irradiated bone (human)	$565—$1,540
Raptos Flex® (Citagenix)	Demineralized cortical bone graft (human)	$1,370—$1,610
Bicon Resorbable Membrane (Bicon)	Resorbable collagen products	$555—$1,300
BioMend® (Zimmerdental)	Absorbable collagen membrane (bovine)	$810—$1,740
Puros® allograft (Zimmerdental)	Cancellous particulate, cortical particulate, and corticocancellous (human)	$145—$250

Product (company)	Material	Price[3] (USD/cm– USD/ml)
Puros® allograft (Zimmerdental)	Block allograft (human)	$290—$730
IngeniOs™ synthetic particles (Zimmerdental)	Hyaluronic acid (HA) bone (synthetic) and β-TCP bioactive bone (synthetic)	$60—$200
Puros® Dermis (Zimmerdental)	Dermis allograft tissue matrix (human)	$250—$1,420
Puros® pericardium (Zimmerdental)	Pericardium membrane allograft (human)	$500—$1,050
HydroMatrix™ (Sigma Aldrich)	Peptide nanofiber three-dimensional scaffold	$35—$60
HyStem® Cell Culture Scaffold Kit (Glycosan BioSystems)	HA-based matrix (synthetic)	$320
CellCeram™ insert (Scaffdex Oy)	Hydroxyapatite and ß-tricalciumphosphate	$240
BiostructureMatrix Scaffold Sheets (Synthecon Inc.)	PGA, PLLA, or PLGA 10:90	$35—$140
BiostructureMatrix Disc Scaffolds (Synthecon Inc.)	PGA, PLLA, or PLGA 10:90	$230
DirectGen™ Bone grafting-allograft (Implant Direct)	Cancellous particulate, cortical particulate, corticalcancellous blend, and demineralized cortical particulate (human)	$70—$120
DirectGen™ Putty (Implant Direct)	Demineralized bone (human)	$140—$220
BioResorb® Macro Pore (Implant Direct)	beta tricalcium phosphate (β-TCP; synthetic)	$80—$130
DirectGen™ Derm (Implant Direct)	Decellurized dermal allograft tissue (human)	$200—$800

Although some of the materials can be relatively inexpensive (~$30/cm), this can still represent a significant cost in some regions of the world.[3] According to the World Bank, about one-third of the world's population (~2.2 billion people) live on less than $2/day.[15] Therefore, for these people, a piece of the least expensive biomaterial in Table 3-1 (smaller than a sugar cube) would require two weeks of their salary. One must also keep in mind that this is only the cost of the raw

biomaterial and does not include shipping, processing by the hospital, overhead, consumables, pharmaceuticals, or other expenses. These expenses can, in some cases, lead to a significant increase in the cost associated with next generation RegenMed and healthcare. RegenMed biomaterials have the potential to effectively treat impaired tissues/organs due to birth defects or following catastrophic tissue damage. However, given the cost of these specialized biomaterials, it is unclear if they will be available to those in regions of the world where resources and accessibility are scarce. Will they be available in war-torn areas of the world? Will companies be willing to make such biomaterials available or affordable in regions of the world facing significant economic challenges? The prohibitive cost of these materials creates a situation that limits their widespread use and availability. In addition, these materials are closed source. Intellectual property, licensing agreements, and specialized production facilities make it impossible for local communities to produce these biomaterials directly on site; thus, scaffolds cannot be produced in the hospital in which they will be used).

Accessibility to these specialized biomaterials remains an important open question. Many of the biomaterials in Table 3-1 are of animal and human origin, raising concerns about the processes employed to obtain the source components, as well as human/animal welfare controls. In 2006, this issue became headline news in the US with respect to the origin of implants derived from human samples.[16] American body-harvesting labs illegally obtained tissue samples from cadavers from hospitals all over the country, without consent from donors or relatives. Although these crimes occurred in a wealthy country, it is not hard to imagine that economic pressures could drive a similar situation in the poorest parts of the world. The scandal in the United States has provoked many concerns about what oversight is applied to a billion-dollar industry that supplies essential biomaterials to hospitals worldwide so that they can perform millions of transplants and reconstructive surgeries per year.

Open Source Biomaterials

This situation begs the question, "Is it possible to create low-cost, DIY, open source biomaterials?" Biology and evolution have produced a nearly infinite number of natural structures and architectures in the plant world. Therefore, we hypothesize that it should be possible to find a naturally occurring scaffold candidate that possesses the qualities of a commercial RegenMed biomaterial scaffold. Indeed, we recently published work in the open-access journal, *PLOS ONE*, in which we describe the simple preparation of apple tissue to create a functional biomaterial (Figure 3-2).[17] In this study, we decellularized apple tissue, which created porous 3D scaffolds. This work proved that plant cellulose (specifically

apples) could act as a scaffold for the 3D growth of human and mouse cells *in vitro*. It is already known that cellulose is a good biomaterial candidate, as it has already been used for different purposes, such as permeable dialysis membranes and as diffusion-limiting membranes within biosensors.[18] In our study, we were able to demonstrate that mammalian cells were able to proliferate inside the apple cellulose scaffold *in vitro* and reach a very high density. While these results are promising, cellulose is not perfect. Mammalian cells do not naturally grow on cellulose structures and do appear to have a decreased proliferation rate compared to commercial scaffolds. Yet, in spite of a slower growth rate, we did observe that mouse and human cells still fully invaded and infiltrated the scaffold. Importantly, our intent here is not to extol the virtues of apple-derived cellulose scaffolds. Rather, our work is intended to provoke a discussion about the use of natural, renewable, and organic resources to create scaffolds, as opposed to proprietary (bio)chemicals and processes.
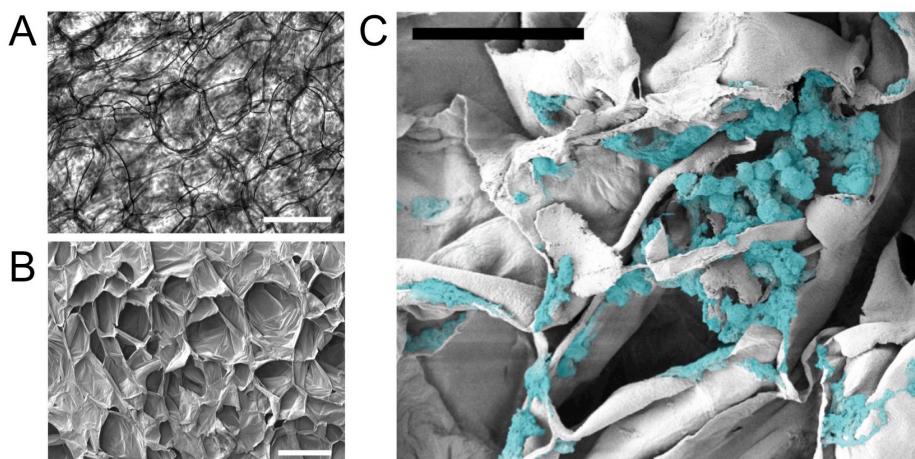
Figure 3-2. (A) Phase contrast microscopy and (B) scanning electron microscopy (SEM) of apple-derived decellularized cellulose scaffolds (scale bars = 200 µm). (C) SEM of mouse muscle cells attaching and growing on the cellulose scaffolds (scale bar = 50 µm). Figure is adapted from the Modulevsky et al. study[17].

Compared to commercial products, apples typically cost <$0.01/cm³, representing a >10,000 fold decrease in cost compared to the average price listed in Table 3-1. Moreover, to prepare apple tissue as a biomaterial scaffold, one only requires boiled water and liquid dish soap. For those interested, we present a DIY protocol for preparing apple-derived cellulose biomaterial scaffolds in Chapter 6.

Although one can enhance the production of these materials with the use of antibiotics, biological buffers, and specialized surfactants, they are not necessary. Therefore, DIY open source biomaterials can be easily prepared with readily available plants and supplies from the grocery store. Moreover, there is no need for specialized equipment, and the biomaterials can be kept sterile with boiling water and proper training. An interesting implication of the apple-based biomaterials is that the apple tissue can be easily shaped (e.g., through CNC milling, laser cutting, or simply with wood carving tools) into almost any desired structure. Moreover, plants naturally possess many intriguing structures (e.g., leaf vasculature) that might be exploited in some way in order to elicit additional functionality.

Open source biomaterials also present an opportunity to explore physical biohacking. In contrast to the more well-known practice of biohacking through the manipulation of DNA, physical biohacking involves the re-purposing of the components of living matter (in this case, cellulose and animal cells). In essence, the objects prepared in our lab are a hybrid between plant and animal (in the spirit of the movie *Little Shop of Horrors*). The samples were created without the use of genetic engineering; instead, researchers relied on standard cell culture techniques. This work highlights the ability to create living, functional, biological composites that do not naturally exist in nature, without resorting to the manipulation of DNA. Recently, we presented this work in two bioart installations, *Re-Purposed* (TOXICITY, Plug In Gallery, Winnipeg, Ontario, Canada, 2014) and *Re-Purposed 46* (BioArt | Collabroating with Life, Karsh-Masson Gallery, Ottawa, Canada, 2015; see Figure 3-3). These works provoke a reflection on the hopes, fears, and possibilities associated with engineering biological objects that blur the lines between science fiction and reality.

Open Source Biomaterials: Too Good to Be True?

As is true for all good things, there may be potential limitations to the widespread usage of open source biomaterials. Although the academic and DIYbio communities may be able to develop these novel materials, bringing them into the clinic in a meaningful and global way represents a significant, but not insurmountable, challenge. As with all medical technologies and therapeutics, open source biomaterials will need to be tested in animals, followed by human clinical trials. Preclinical studies will require both *in vitro* and *in vivo* testing with different cell cultures and animal models. If the results from such studies are promising, human clinical trials will then need to be initiated and carried out. Such studies are remarkably expensive and time consuming and will require close collaboration and knowledge sharing between the DIYbio, academic, and clinical communities. Although open source biomaterials should be relatively cheap and easy to pro-

duce, significant financial support will be required to ensure that they are safe for human use. Companies, or national granting agencies, often support clinical biomedical research when they foresee a potential economic return on their investment. In the case of open source biomaterials that lack intellectual property protections, such clinical studies may only be possible through philanthropic investments.

Figure 3-3. Re-Purposed 46 by D. Modulevsky and A.E. Pelling. Apples, Human HeLa Cells, Cell Culture Plastic, Acrylic. Shown at BioArt | Collabroating with Life, Karsh-Masson Gallery, Ottawa, 2015. Forty-six slices of decellularized apples were impregnated with human cells derived from Henrietta Lacks and preserved (HeLa cells). Photo by Luc Lalande (@LucLalande).

Another challenge to the widespread use of open source biomaterials will be to develop standardized production protocols to ensure quality control and reproducibility. It remains unclear how one would ensure quality control over a material that could be easily and cheaply produced anywhere in the world. Moreover, there are potential dangers due to contaminants in local water sources and environments. These challenges are just some of the issues that will need to be

addressed in the future in order to pursue real-world clinical applications of open source and DIY biomaterials.

Conclusion

In this article, we have highlighted the potential benefits and challenges of using different materials, from artificial to natural origins, in the development of therapeutic strategies for RegenMed. Although there are challenges to bringing open source biomaterials into the clinical setting, they are certainly not impossible to solve. Open source biomaterials have the potential to immediately impact the costs of RegenMed healthcare, not just in terms of production but also in costs associated with licensing or buying the intellectual property. We want to challenge others to begin thinking beyond developing yet another closed source, proprietary biomaterial and begin developing simple, cheap, and open biomaterials. Our lab is entirely funded through public tax dollars, and it is a priority for us to ensure that the science being done is not only open, but will ultimately be accessible to anybody, regardless of state, geographic, or economic resources. By releasing the general protocols and intellectual property into the public domain, we hope that a larger body of researchers in the DIYbio, academic, and industrial spaces will be able to propel such work forward. We are certain that others can improve and continue to develop the protocol published alongside this article and look forward to seeing the type of work it disrupts, provokes, and inspires.

References

1. Famous Picture Collection, "Ear Mouse," 2013, *http://www.famouspictures.org/ear-mouse/*.

2. BBC News, "Artificial liver 'could be grown,'" April 25, 2002, *http://bit.ly/bbc-liver*.

3. Y. Cao, J. P. Vacanti, K. T. Paige, J. Upton, and C. A. Vacanti, "Transplantation of Chondrocytes Utilizing a Polymer-Cell Construct to Produce Tissue-Engineered Cartilage in the Shape of a Human Ear," *Plast. Reconstr. Surg.* 100, (1997): 297–302; discussion 303–4.

4. S. Locke, "An Organ Shortage is Killing People. Are Lab-Grown Organs the Answer?" *Vox,* November 20, 2014, *http://bit.ly/lab-organs*.

5. K. Vezina, "First Fully Synthetic Organ Transplant Saves Cancer Patient," *MIT Technology Review,* July 8, 2011, *http://bit.ly/synth-organ*.

6. J. Glausiuasz, "The Big Idea: Organ Regeneration," *National Geographic Magazine,* March 2011, *http://bit.ly/ng-organ*.

7. E. Zolfagharifard, "The Hospital Growing Noses, Tear Ducts and Blood Vessels: Scientists Make Custom-Made Body Parts Using Stem Cells," *The Daily Mail,* April 8, 2014, *http://bit.ly/dailymail-lab*.

8. C. Johnson, "Seeing Local Scientists at Heart of Regenerative Medicine," *Boston.com*, February 7, 2011,*http://bit.ly/boston-sci*.

9. A. Coghlan, "Kidney Breakthrough: Complete Lab-Grown Organ Works in Rats,"*NewScientist*, April 14, 2013,*http://bit.ly/rat-kidney*.

10. BBC News, "Lab-Grown Bladders 'a milestone,'" April 3, 2006,*http://bit.ly/bbc-bladder*.

11. J. Gallagher, "Miniature 'Human Brain' Grown in Lab,"BBC News, August 28, 2013, *http://bit.ly/bbc-brain*.

12. A. Sheela, "Global Biomaterials Market to Become Worth USD 33.6 Billion in 2019: Transparency Market Research," PR Newswire, October 9, 2014, *http://bit.ly/biomat-market*.

13. "Biomaterials Market Worth $88.4 Billion - 2017," *MarketsandMarkets, http://bit.ly/biomats-mkt*.

14. "Poverty Overview," 2015, *http://bit.ly/poverty-ov*.

15. M. Powell, and D. Segal, "In New York, a Grisly Traffic in Body Parts," *Washington Post*, January 28, 2006, *http://bit.ly/nyc-traffic*.

16. D.J. Modulevsky, C. Lefebvre, K. Haase, Z. Al-Rekabi, and A.E. Pelling, "Apple Derived Cellulose Scaffolds for 3D Mammalian Cell Culture."*PLOS One* 9 (2014): e97835.

17. K. Ishihara, H. Miyazaki, T. Kurosaki, and N. Nakabayashi. "Improvement of Blood Compatibility on Cellulose Dialysis Membrane. III. Synthesis and Performance of Water-Soluble Cellulose Grafted with Phospholipid Polymer as Coating Material on Cellulose Dialysis Membrane." *J. Biomed. Mater. Res.* 29 (1995) 181–188.

Affiliations

[a]Centre for Interdisciplinary NanoPhysics, [b]Department of Biology, [c]Department of Physics, [4]Institute for Science, Society and Policy, University of Ottawa, Ottawa,

Ontario, Canada, and ^dSymbioticA, School of Anatomy, Physiology and Human Biology, University of Western Australia

Acknowledgments

Daniel Modulevsky (@doc_blood) received his undergraduate degree in biochemistry at the University of Ottawa, Canada. He is currently a PhD student there under the supervision of Dr. Andrew E. Pelling. He is the recipient of PhD fellowship from the Fonds de Recherche du Québec—Santé and was awarded the Presidential Scholar Award from the Microscopy and Microanalysis department in 2013.

Charles M. Cuerrier (@ChuckCuer) received a PhD in pharmacology at the Université de Sherbrooke, Canada. He is currently a postdoctoral fellow at the Centre for Interdisciplinary Nanophysics, University of Ottawa, Canada. He is the recipient of a postdoctoral fellowship from the Fonds de Recherche du Québec—Santé and of the Ernest and Margaret Ford Research Fellowship in Cardiology from the University of Ottawa Heart Institute.

Andrew E. Pelling (@pellinglab) received a PhD in physical chemistry at the University of California, Los Angeles, and did his postdoctoral research as a senior research fellow at the London Centre for Nanotechnology, University College London, England. He is currently an associate professor and Canada Research Chair in Experimental Cell Mechanics at the University of Ottawa, Canada. He leads the Pelling Lab for Biophysical Manipulation at the Centre for Interdisciplinary Nanophysics, University of Ottawa.

Flexible Robotic Platforms That Allow Scientists to Remain Scientists

Raghu Machiraju

Biotechnology has not delivered on its promise of breakthrough discoveries of immense benefit to humankind. Laboratory automation that was meant to accelerate the pace of discovery is mostly inaccessible and unusable for most experimental biologists. Current solutions are expensive and only offer improvements in throughput of specific steps in a given experimental protocol.

In this article, we argue that progress in biotechnology will come from the use of open user interfaces and open-specification middleware to drive and operate flexible robotic platforms. Such middleware will accelerate and scale the *plan-execute-analyze* cycle of development in the biosciences. It will also allow one to integrate multiple kinds of robotic platforms, 3D printers, and sensors, and drive these systems through biologically meaningful user interfaces. Further, the availability of such middleware will provide reasonable-cost solutions rather than the current expensive offerings. Both tinkerers and mission-oriented researchers can then create or acquire affordable and custom robotic platforms and explore the marvels of all things biological or complete specific tasks, respectively. Most importantly, such platforms will allow scientists remain scientists and not computer technicians.

In this article, we first describe the state of affairs in biotechnology and then discuss aBioBot's solution, which provides flexible, low-cost, ergonomic experimentation while shortening the time to discovery.

The Unfulfilled Promise of Biotechnology

Although the potential of biotechnology is often touted, the industry only recently saw a scintilla of exponential growth scales that consumer electronics and personal computing have enjoyed in decades past. The emergence of genome sequencing that blows past Moore's law spurred some of this growth. High-throughput sequencing and phenotypic screening generated a prodigious amount of data, which is not often well interpreted. There are urgent needs to test and utilize the data produced. Although this technology and its promise led to strong sales within the biopharma industry, the technology has not yet led to the expected mass-market adoption or the creation of new industries. Consequently, there is very little room for the average biotech tinkerer or the serious molecular biologist to create novel consumer technologies or services.

Although some processes in biotechnology, such as manufacturing and instrumentation, have matured, the innovation and discovery process remains archaically manual for all but the most sophisticated and resource-rich companies. It is the hearth of biotechnology, the wet laboratory, that needs and deserves a close second look.

Impediments to Wet-Lab Productivity

For molecular biology and medicinal chemistry, the vast majority of innovation is still a largely manual process performed at the bench in a wet laboratory. Typical metabolic engineering and systems biology problems have high degrees of freedom (e.g., many genes in a synthetic pathway), leading to heterogeneous responses to perturbations. Biological science is inherently iterative—composed of several rounds of plan-execute-analyze steps. The plan-execute-analyze cycle is the norm in science and engineering. The laboratory scientist plans by choosing a protocol, then executes the protocol in the wet laboratory, and analyzes the resulting data from experiments. These steps are repeated many times until the desired results are obtained and adequate insights are gleaned. Therefore, in the absence of automation, a significant biotech project is often measured in person-decades. There are additional costs; for many mundane and cumbersome procedures, the resulting errors lead to expensive downstream protocols. Most importantly, irreproducible research is often the result.

Crucial obstacles to the widespread adoption of biotechnology are the elimination of manual processes, commoditization of methods and machines to perform them, and an open structure of the processes and assembly across the industry. Consider this: the personal computing industry arose from the commoditization of the von Neumann architecture in the form of a usable, albeit adaptable, hard-

ware layout and an operating interface. This early, open architecture and other contextual changes (e.g., the growth of the Internet, the use of tangible user interfaces) allowed for the growth of the personal and iComputing industry. The open community of innovators and the ecosystem it spawned also ushered in the age of personal computing. Similarly, it can be argued that widespread adoption of bioinformatics also arose from the growth of open-access tools and environments including BLAST and the R language. Further, the open community of users and data analysts who shared their trials and tribulations of processing a certain sequencing data. The need for open interfaces and architecture and thriving open community is a requirement in the biotech industry.

Let's now examine the state of the automation in the wet lab; we will undoubtedly learn that the industry offers a closed and proprietary approach.

Automation for the Wet Lab

Technologies and processes exist that significantly multiply a bench scientist's productivity, but they are not accessible. As they exist, lab robots produce highly reproducible work and can potentially improve productivity by several factors for encodable tasks. There are many examples where automation has positively contributed to expedient and robust treatments and procedures. For instance, through shallow sequencing and highly optimized automation, the painful amniocentesis procedure can be replaced by noninvasive prenatal tests (NIPT), molecular profiling of the mother's blood.

Commercially available laboratory robotics platforms at best occupy a small niche in the market. Some of them deliver very high throughput for expedited execution. However, high throughput does not translate into expedited discoveries, because the other two stages of planning and analyzing are still left to manual means.

Existing solutions are often unsafe to operate, overtly expensive (some in excess of $300,000), operate in limited orthogonal Cartesian coordinate systems and are encased in rigid steel truss, thus proportionally increasing the cost of ancillary equipment such as well plates and temperature sensors. Further, they require very proprietary equipment for manufacturing, such as trademarked pipette tips. Lastly, the flexibility in a typical lab cannot be replicated on the robotic bench. A variety of tubes and well plates cannot be used on a typical robot and can be placed in any orientation. A change in protocol requires extensive reprogramming of the various material handling procedures. It is not common to have automation engineers in well-heeled institutions that are dedicated to these tasks.

To reiterate, we stridently believe that this state of affairs exists because of the closed nature of existing biotech enterprise. Clever innovators cannot easily incorporate their ideas into existing robotic frameworks. There is no common and tangible interface that is accessible to one and all. Even more importantly, the apparatus has very little self-awareness and cannot be taught to be aware of its configuration and capabilities. Given the plethora of degrees of freedom in space, time, and function, the latter ability is very important for biotech to be a viable and tangible possibility. Still, is the situation really that dismal?

Emerging Solutions

New efforts have emerged to mitigate the situation. Emerald Cloud Lab, Transcriptic, SyntheGo, and Arcturus BioCloud are companies scaling laboratory automation through services in the cloud. They automated experimental design as a service, which has been well received. However, they still do not offer the flexibility and direct touch and feel of the bench as desired and needed by the typical investigator. To their credit, the workflows offered by these companies improve the planning and execution stages of the iterative scientific development cycle as it pertains to selected protocols.

Another set of companies and start-ups (Modular Science, iorobotics, and OpenTrons, to name a few) are experimenting with lab-automation paradigms that share our vision. They adopt simple, inexpensive, open-design and extensible hardware platforms. Still, they offer limited capabilities to improve the iterative plan-execute-analyze development cycle. The software infrastructure to plan the protocol and execute it efficiently on the robotic platform is not available. There are novel academic solutions that have been recently proposed. The Riedel-Kruse laboratory at Stanford University offers an open source biotec processing unit (BPU) with a tangible user interface and highways to the computing cloud.

To chart a useful path, you must understand the essential computational tasks of wet lab procedures and choose those that are encodable and can be automated to accelerate the pace of discovery.

Wet Lab Work Is Encodable

Each protocol can then be represented either with a collection of natural language statements or as a dataflow process graph. In either case, protocols are a set of computations and translators that can be built between the two representations. There are at least two sets of nodes; one set represents materials, and the other set describes operations (e.g., mixing) resulting in creation of new or intermediate materials (see Figure 4-1). The edges indicate the transport of material, and in this case, describe the material mixing for preparing the master mix for polymerase

chain reaction, or PCR. The dataflow graph can be translated, in turn, into G-code or CNC-code that can drive a robotic head. However, for this to occur, various receptacles or wells have to be bound to different reagents. The dataflow process graph, when annotated with spatial locations can be examined for optimal path planning and can include constraints of containment. We now describe the visual sensing module and the role it can play.
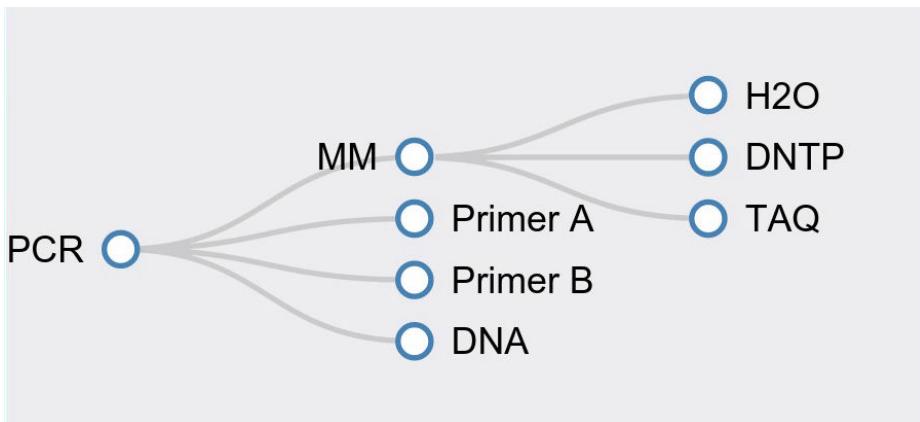
Figure 4-1. A data flow graph for preparing master mix for PCR; each node is material and can also implicitly define an operator (mixing in this case)

Visual Sensing and Feedback

The breakthrough we seek initially is to equip the robot with visual sensing capabilities and thus achieve visual servo control of the robots. The rewards are many. Because the sensors can identify all objects on the bench, it is easy to plan robotic operations there. The binding of bench objects to material nodes in the dataflow graph is achieved through the culmination of analysis of acquired data. An additional operational benefit is that the machine can be trained to compensate for small and large distortions in the work area. Visual confirmation of the work surface will allow the robot to be safer (i.e., it will be able to sense hands and obstructions) as well as less error-prone (i.e., it will automatically be able to sense if a well has liquid in it, whether a tip is missing from the box, or whether colonies or contamination exist on a plate). Further, when it rains pipette tips, it will be useful for the operator to be informed. We firmly believe that our approach of using sensing in a middleware parlance will allow for the creation of automated workflows that will be safe, reliable, accurate, and cost effective.

Middleware Is Essential

We believe that solutions lie in the availability of middleware interfaces that will allow for the planning and completion of a large number of encodable experiments. The middleware layer will then communicate with devices (robots, sensors, etc.) and consummate precisely routine and encodable tasks that would be otherwise completed manually and erroneously.

Consider this: a user interface can serve as the command and control of the experiment. Protocols in the form of annotations will be listed and annotated for materials and processes. Typical protocols include the preparation of master mix for PCR (as described earlier), DNA extraction, ELISA assay, and serial dilutions. Using either automated or manual (through annotation) means, a dataflow process graph can be extracted from the text-like description and displayed. The video from a camera will be captured and analyzed for the presence and location of well plates and other equipment. This information will be delivered to the user interface, which will be further used to plan experiments; the user can assign various reagents to well plates and even more specifically list the various sources and destinations for each reagent of the protocol. Once all assignments have been done, the user can execute the protocol. Appropriate commands (G-code) will be generated and communicated to the printer. As the robot head completes each of the protocol, it can communicate with the executive middleware, which then logs the progress and also updates the process dataflow graph.

The plug-and-play approach fostered by the middleware will allow the precise completion of ordinary tasks at reasonably lower costs with a simple and accessible user interface. Our reliance on middleware and scripting languages also allows expansion of the platform by users and other companies and the customized generation of special control work surfaces for tablets and customized workflows. New pieces such as plate loaders, refrigeration blocks, and colony pickers/counters can be designed by third parties or users and integrated for use with the middleware. We now describe our platform.

The aBioBot Solution

At aBioBot, we wish to re-create the flexibility of the familiar wet laboratory bench on a robotic platform. Our lab assistant is a robotic platform combining state-of-the-art hardware and software:

LabBench
  A web browser user interface to facilitate protocol authoring, observe the layout on the bench, monitor the progress of the experiment, and log the proto-

col on the cloud. The user interface of LabBench is realized in HTML5 and Javascript, and leverages toolkits, including tornado and bootstrap.

Yan = Eyes

In our platform, Yan is machine vision that watches over the experiment for you. Yan's software module learns the layout of wells and provides surveillance of the bench for untoward accidents. Yan is implemented in Python and uses the OpenCV library.

The Bot

Our robotic platform is modified from a 3D printer. We also offer a series of adapters that allow The Bot to use standard pipettes and equipment. As of now, there exists an adapter that allows the use of a standard Eppendorf pipette. The Bot is derived from open source 3D printer hardware and rapid prototyping machinery.

Although aBioBot is built on open source hardware and software, both LabBench and Yan will be accessible through open APIs and extended as necessary. Lab Bench has two functional components accessible to users:

- Every lab staff member has a book of her favorite protocols, which have been refined and improved over time. LabSmith is a protocol-authoring tool with the capability to import lab procedures and notebooks from various repositories, including OpenWetWare and protocols.io. Most importantly, it will also allow for protocols to be changed and adopted as required.

- Every log, video, and status report for your experiment is automatically uploaded to the cloud through LabCloud.

The Path Forward

What we have accomplished over the last three months is the construction of an efficient middleware and a functional desktop prototype. We are currently testing the platform on preparing a master mix to prepare for PCR. Also, we plan to release an open API that will allow for sensors and any 3D printer platform, which will be treated as a device. We are considering the implementation of our software on both desktop and lab bench-sized hardware platforms. We believe that this is the way forward for biotech and the creation of this middleware and robotic platform will be the first steps in creating unexpected multibillion dollar new industries that harness the power of industrialized and roboticized biotechnology.

Keep watching out for announcements at the aBioBot website.
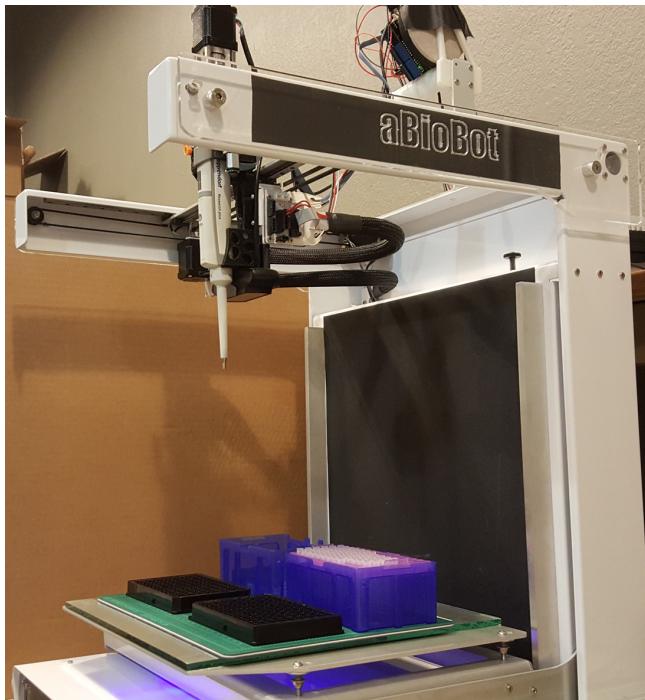
Figure 4-2. The aBioBiot prototype repurposed from a 3D printer. The black-colored trays hold required input and output material, while the blue-colored boxes hold unused and use pipette tips. Finally, the pipette is engineered to replace the printhead.

Raghu Machiraju is a cofounder and CEO of aBioBot and a professor of computer science and engineering at Ohio State University. The mission of aBioBot is to build flexible robotic plat-forms that will help scientists make breakthroughs by shortening the time to discovery. Raghu was aptly helped in the writing of this article by Ron Shigeta of IndieBio and his co-founders Chaitanya Kulkarni, Kun Huang, Ken Rinaldo, and Trademark Gunderson.

The Development of the Personal Genetic Kit

Joshua Elkington

Scientific advancements in synthetic biology, coupled with entrepreneurial projects, are creating the future of biology. The ability to create, manipulate, and program life is in the hands of a limited number of central institutions. However, maturation of personal biotech tools will empower more people to invent new tools and products. In this article, I will highlight multiple personal biotechnology tools developed to bring biotechnology to the masses.

Bento Bioworks is developing a personal biotechnology tool called Bento Lab that is marketed as the "Arduino for Biotechnology" (Figure 5-1). This all-in-one lab device is a personal biotech product with a PCR machine, centrifuge, and gel electrophoresis, along with a blue LED transluminator. It is a platform to copy and detect DNA. Although Bento's product is functionally simple, the emergence of tools that allow an individual to conduct biotech experiments promises to change the course of humanity. The advent of tools like Bento signals the coming age of personal biotech. The Bento team is led by Bethan Wolfenden and Philipp Boeing, and they are currently testing the Bento Lab device prior to selling it to the public. However, the device will be available to beta testers this summer. By creating a device to help democratize bioengineering, Philipp and Bethan are pioneers in personal biotechnology.

OpenTrons is another platform that allows biologists to rapidly prototype biotech products. It is selling an open source, liquid-handling robot called OT.One (Figure 5-2). As advertised on its Kickstarter page, the robot's price starts at $2,000, and it comes with support software, Mix.Bio, to help users design protocols for the robot. The robot is modular, so it has the ability to add new features. For example, plasmid preparation and electroporation ability can be added onto

the OT.One. In collaboration with Synbiota, OpenTrons developed an introductory kit where designed DNA is electroporated into bacteria to make them glow red under blacklight.
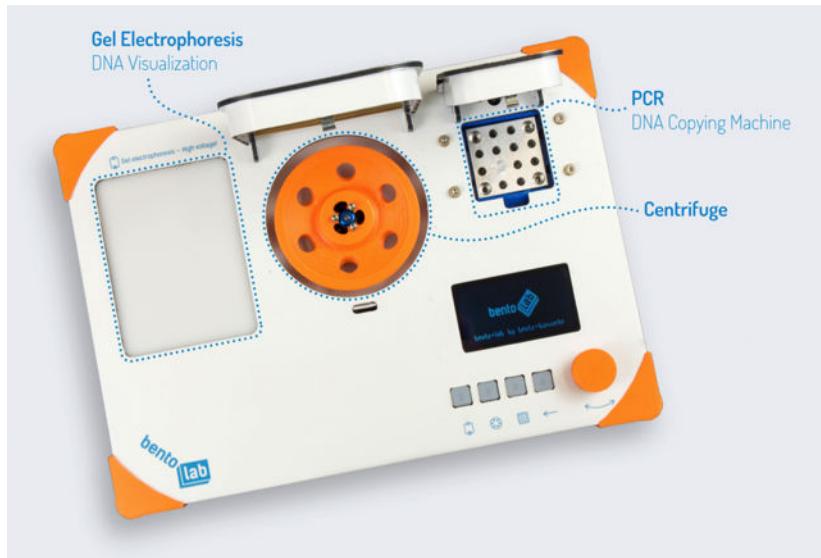
Figure 5-1. Bento Lab

OpenTrons is automating many of the laboratory processes that slow down the pace of research. Through automation and easy-to-use software, they are making biology accessible to people without formal lab training. OpenTrons grew up at Genspace, a DIYbio space in New York City, and is led by Will Canine. In an Open Electronics interview, OpenTrons cofounder Will Canine stated that the company's mission is "to empower a new generation of biodevelopers with powerful, open, and affordable biotechnology tools." The software tool used along with OT.one will accelerate creativity and discovery in biology by providing a platform for automation collaboration via protocol sharing. OpenTrons is developing tools to help individuals innovate in biotech by empowering them, allowing them to easily share ideas and making experiments more reproducible.

ROBOTICS
- 100 Micron XYZ accuracy with the Smoothieboard open-source hardware 5x motor controller.
- Raspberry Pi on board with wifi, BLE.
- OpenTrons open-source NodeJs software.
- Industrial grade chassis.
- 2ft x 2ft x 2ft, approx 60lbs

LIQUID HANDLING
- A single motor mount to attach a standard micropipette, single or multi-channel.
- Microliter accuracy, every time.

WORKBENCH
- U15 SPE well-plate capacity.
- 100% aluminum for easy sterilization.
- Quick snap mounting system.

Figure 5-2. OT.One Robot

Another device that is within the personal biotech movement is the Open qPCR tool developed by Chai (see *BioCoder #6* for a full description). It is an affordable, real-time polymerase chain reaction (RT-PCR) thermocycler first announced on Kickstarter. The device costs $1,999 and is meant to make real-time PCR accessible to more people. On the Kickstarter page, it mentions the project is targeting "doctors in developing countries, students in high school and university labs, companies in the food supply chain, and biohackers who are developing some of the most innovative synthetic biology applications." A similar device is the miniPCR (see *BioCoder #5* for a full description), another funded Kickstarter project developing a DNA extraction, PCR, and visualization kit (Figure 5-2). The kit now costs $749. Both the Open qPCR and the miniPCR are opening up access by lowering the costs of commonly used biotech methods.

Other projects doing similar work for the personal biotech movement are Smart-DART by Diagenetix and Biomeme. A project successfully funded on Indiegogo, Smart-DART is a portable, real-time DNA amplification tool (Figure 5-4). The tool supports any isothermal amplification method. A single-well device currently costs $200. BioRanger is the newer version of the tool based on the previous SmartDart technology, which is currently used by the food industry and others to detect microbes. Another similar product, Biomeme, is a real-time PCR thermocycler that attaches to a smartphone to detect DNA on site. The user isolates DNA from the environmental sample of choice and adds it to a cartridge specific for that target DNA, and puts the cartridge into the thermocycler. Both Diage-

netix and Biomeme are developing low-cost RT-PCR tools that can be used on site and make DNA detection more accessible without the need for bulky laboratory equipment and reagents.
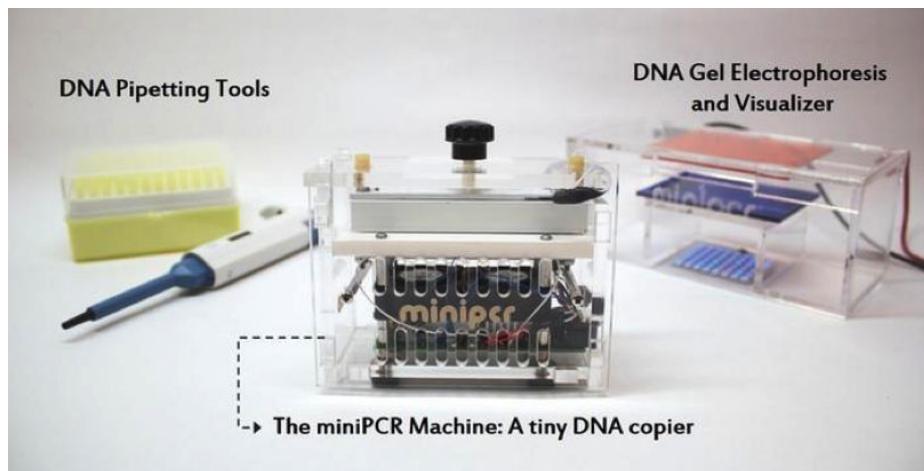

Figure 5-3. miniPCR

In support of the development of personal biotech, an infrastructure is being created. Around the world, DIYbio or community laboratory spaces have cropped up, lowering the barrier for entry (some of which have been discussed in previous articles). They provide paying members access to basic molecular biology tools. Just as OpenTrons came out of Genspace, the Y Combinator-backed Glowing Plant Project grew out of Biocurious. Similarly, Counter Culture Labs is a DIYbio group in the East Bay with the purpose of generating curiosity for science and empowering people to explore. The open space gives people a place to learn about and tinker with biology. An exciting project at Counter Culture Labs in collaboration with people at BioCurious is genetically engineering yeast to produce vegan cheese protein.

In addition to laboratory spaces to grow an idea, many services have been created to allow people to develop personal biotech tools. Benchling is software that makes it easy to manipulate and share DNA sequence information for cloning projects. Transcriptic is like the AWS for biotechnology, providing automated solutions for biology. Through Transcriptic, a researcher can outsource experiments such as cloning, protein assays, and cell growth assays. And, if you have a very specific need, such as running LCMS samples or next generation sequencing, you can find a provider on Science Exchange.

Figure 5-4. Smart-DART

Gene design and biological engineering by anyone with an idea will spur the development of classes of products that have yet to be thought up. In a garage, with a tool like OpenTrons, people can create new products with biology, which is already coming to fruition in the iGEM (International Genetically Engineered Machines) competition, now accepting DIYbio participants. People with no affiliation to a university or corporation are now beginning to use tools and methods developed over the last 50 years to hack biology.

Tools that were once only available to trained scientists are finding their way into the hands of creative individuals who may not have formal training. Innovation in biotech will no longer be limited by technology and accessibility. Nature will grow as a function of human imagination.

Joshua Elkington is a student at Harvard College; he is interested in biotechnology, politics, and startups. Joshua is the founder of Biodevices, a startup at the Harvard iLab focused on building inexpensive science tools. He also does research on small RNAs and aging in the Hunter and Ruvkun Lab, respectively. Joshua is the cofounder of the Harvard Biotech Entrepreneurs Club; in his spare time, Joshua supports BOSS Lab and mentors the iGem team at Codman Academy. He will be entering the Molecular and Cellular Biology graduate program at the University of California, Berkeley, this fall. You can contact Joshua at jmelkington@yahoo.com.

DIY Open Source Biomaterials

Daniel Modulevsky[a,b] and Andrew E. Pelling[a,e]

The regeneration of diseased or damaged tissues and organs is one of the major goals of regenerative medicine (RegenMed). Currently, there are a variety of approaches under intense investigation to achieve the repair of diseased or damaged tissues and organs. In addition to STEM and/or gene therapy, physical (bio)polymer scaffolds are being developed in 3D shapes that mimic both the physical and biochemical properties of the original organ, which may aid in the overall healing process.[1-6] Such scaffolds are generally very expensive and costly to produce, as discussed in our other article in this issue (Chapter 3). One promising approach has been the use of decellularization techniques to produce "ghost organs" that can be repopulated with living cells derived from donors or possibly even the patient.[7] One drawback is that donor tissues/organs are still required in order to produce the replacement. Also in development are synthetic or naturally derived polymers which can be molded or 3D printed into specific shapes. This removes the need for donor organs; however, the synthetic chemical or biological scaffolds can be very costly due to the processes required to produce them and/or the intellectual property rights associated with licensing of such materials. The downstream effect is that the widespread use of such scaffolds in RegenMed will likely be limited to nations that can afford the significant increase in healthcare costs.

In our recently published work,[8] we have demonstrated that the cellulose-based scaffolding structures found in plants display many of the same characteristics of the commercial scaffolding products. We demonstrated that apple flesh tissue can be easily carved into many 3D shapes, decellularized (removing all the plant cells while leaving the cellulose intact) and that the resulting cell wall scaffold can be repopulated with human and mouse cells (in fact, virtually all mammalian cells; see Figure 6-1). Most importantly, at a minimum, the process to

produce such scaffolds only requires basic kitchen supplies (dish soap and boiling water). Therefore, the entire process does not require any proprietary materials. We refer to such materials as "open source biomaterials" because anyone can easily grow or purchase an apple and prepare a scaffold with a shape and size of their choosing using off-the-shelf supplies. In the following paragraphs, we have outlined a protocol using common household goods that will allow any DIYbio enthusiast to create her own open source biomaterials.
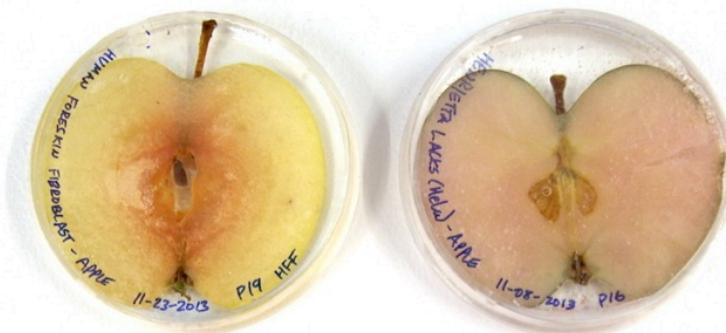
Figure 6-1. Male and female cells cultured on apple cellulose scaffolds. Human Forskin Fibroblast (HFF) and Henrietta Lacks Epithelial cells (HeLa) are seen here growing on decellularized apple slices. From the installation Re-Purposed (Bryan, Modulevsky and Pelling, 2014 at TOXICITY, Plug In Gallery, Winnipeg, Canada[9])

Methods

If you want to create your own decellularized apples as open source biomaterials, we have summarized a suggested protocol in the following sections. This protocol has been adapted from our work previously published in the open access journal PLOS ONE.[8] In this adaptation, the DIYbio enthusiast can complete apple decellularization with readily available supplies found in any grocery store or corner shop.

STEP 1: CUTTING THE DESIRED SHAPE

For best results, obtain thin apple tissue. Cut the apple tissue into the desired shape (this protocol is for ≈ 1 cm³ piece). When cutting the tissue, remember that a larger surface-area-to-volume ratio will result in the quickest and best quality scaffold. However, you can cut any desired shape and size, as long as you keep in

mind that total decellularization time will depend on the diffusion of the detergent solution into the apple tissue (Figure 6-2). Some experimentation may be required.
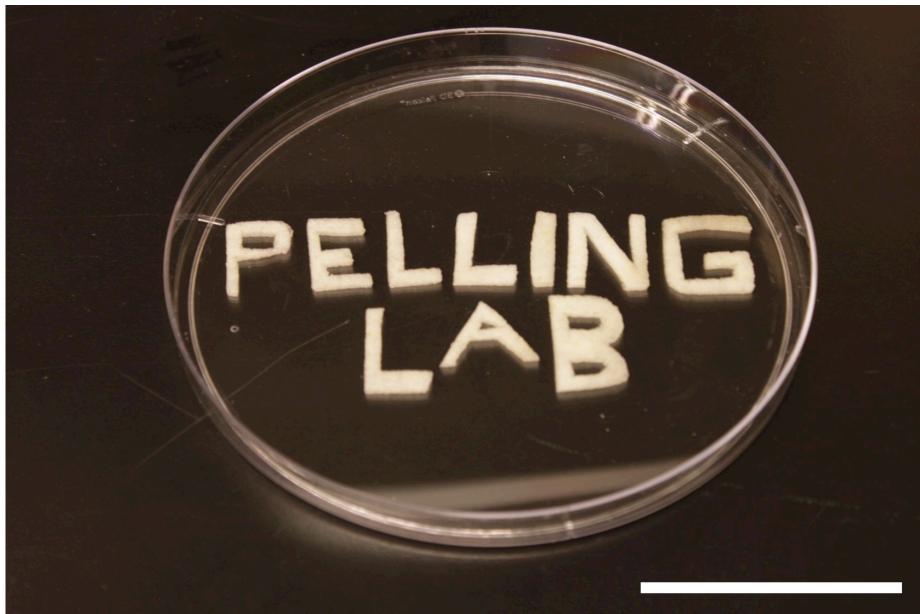
Figure 6-2. Letters hand cut from a single Macintosh apple, highlighting the variability of scaffold shapes that can be produced. The scaffolds were cut using a mandolin slicer and a scalpel. Scale bar = 5 cm.

STEP 2: APPLE PRE-SOAK

After cutting the apple tissue into the desired shape, the tissue needs to be rinsed to remove smaller debris accumulated during the cutting process. The apple scaffold can be rinsed repeatedly with previously boiled water. *Make sure you bring the water to a boil for at least 10 minutes and let the water reach room temperature before using it.* In order to sterilize it, all water used in this protocol should be boiled and allowed to cool to room temperature.

> **NOTE** Many dishwashers have a sterilization cycle that can sterilize the glassware that you may use throughout the decellularization protocol (e.g., jars used to store sterilized water). If your dishwasher does not have that setting (or if you do not have access to a dishwasher), you can boil your glassware in a large pot for 10 minutes and allow it to cool.

STEP 3: WASHING

While the apple tissue is submerged in sterilized water, prepare the decellularization detergent solution. This solution is required to break up and remove any cells or cell debris from the scaffold. To make the detergent solution, 1 teaspoon of liquid soap can be poured into a large, empty, sterilized jar. Then fill the jar with 2 cups (500 mL) of water to produce a ~1% solution. To make a larger volume, measure 4 teaspoons of Sunlight soap (approximately 20 mL) and pour it into an empty 2 L bottle (Figure 6-3). Then fill the bottle with 2 L of water and mix the solution so that the soap is fully dissolved.

> **NOTE** You do not have to use Sunlight liquid detergent as shown here. If you purchase another liquid detergent, make sure that it is tough against grease. If you have access to sodium dodecyl sulphate (SDS), this is an ideal reagent for decellularization, and a 1% solution can be made up as an alternative to dish soap. SDS is best to use for decellularizing animal tissues.

Figure 6-3. Commercial Sunlight soap and a 2 L plastic soda bottle. A 1% soap solution can be made by mixing 4 teaspoons of soap and 2 L of water.

STEP 4: DECELLULARIZATION

Place the tissue sample within a container, and gently pour the soap solution so that it fully covers the tissue. To speed up the decellularization process, slowly agitate the mixture either through stirring or shaking. If you have time, a DIY shaker can be made from an old CD-ROM drive. There are several plans already on the Web (Figure 6-4).[9-11]
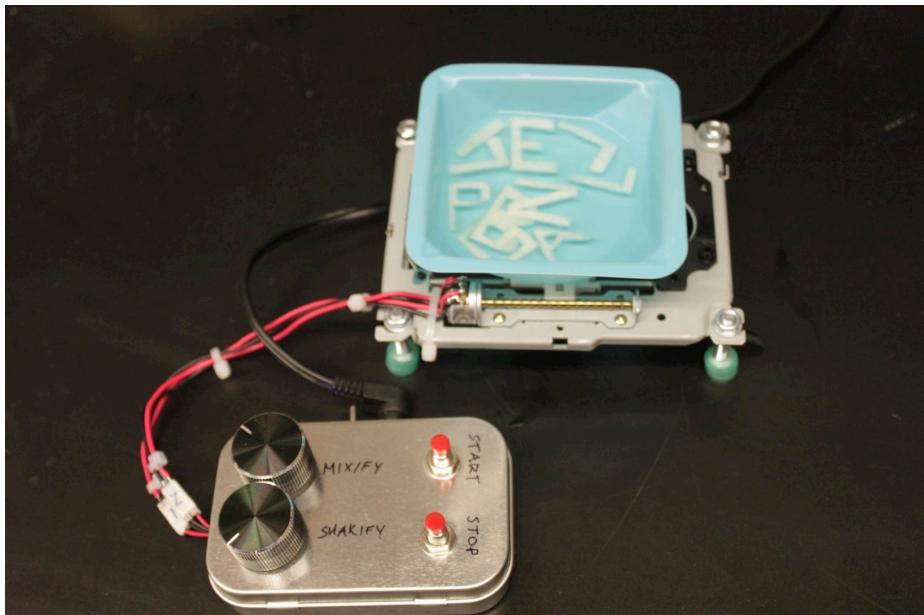
Figure 6-4. The apple scaffolds are submerged in a 1% soap solution and agitated with a DIY shaker. See references 9-11 for instructions on how to make your own.

Depending on the size of the tissue and the concentration of the soap solution, the decellularization time can vary. Make sure to refresh the soap solution when it appears to be cloudy or begins to smell bad. Theoretically, bacteria should not be able to live in your soap solution. However, if it has become unsterile for any reason and begins to smell bad, replace it immediately with fresh soap solution. It is important to maintain sterility if the goal is to grow mammalian cells in the scaffold. However, we recognize that some of you may want to grow bacteria or fungi in the scaffold, and an infection may be beneficial for that purpose. If this is the case, be as dirty as you like!

STEP 5: THE GOOD, THE BAD, AND THE TRANSPARENT APPLE

Take the apple tissue out of the soap solution four hours after it appears translucent (Figure 6-5). The timing and degree of agitation requires some trial and error, as the tissue can appear to be decellularized at the surface but still contain cells in the interior of the sample. It is normal for larger pieces of apple tissue to begin to appear brown as decellularization proceeds.
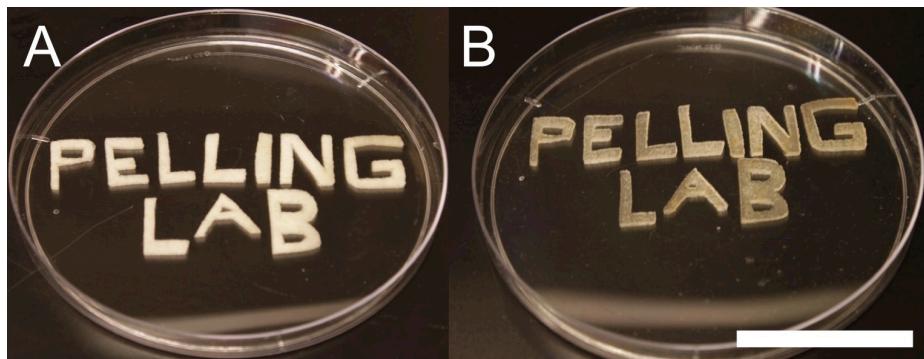

Figure 6-5. Before and after images of decellularized apple scaffolds. (A) Macintosh apples before decellularization. (B) Macintosh apples after 72 hours in 1% soap solution. The decellularized scaffolds become translucent, revealing the colour of the bench top. Larger samples can sometimes appear brown.

STEP 6: THE CLEANUP

After decellularization, rinse the apple scaffold with water at least 10 times and place into 0.5 L of water for three hours. This step is critical, as all the soap solution must be washed out of the apple tissue. The soap solution is required to remove any cells from the apple tissue. If you would like to grow new cells into the scaffold, it is important to make sure all the soap solution has been thoroughly rinsed away.

STEP 7: SO YOU WANT TO GROW MAMMALIAN CELLS IN YOUR DECELLULARIZED APPLE?

We have shown that it is possible to culture established secondary cell lines within the apple scaffolds.[8] Although DIYbio techniques are rapidly advancing, it is still difficult to culture mammalian cells at home. There are a number of reasons for this. First, unlike bacteria, mammalian cell lines are significantly more expensive and difficult to acquire. Secondly, mammalian cells require incubators that not

only control temperature, but also control the CO_2 content of the atmosphere. These incubators can be expensive to acquire and maintain, although we have recently published plans for a DIY CO_2 incubator that was successfully used to culture mouse and human cells (Figure 6-6).[13]
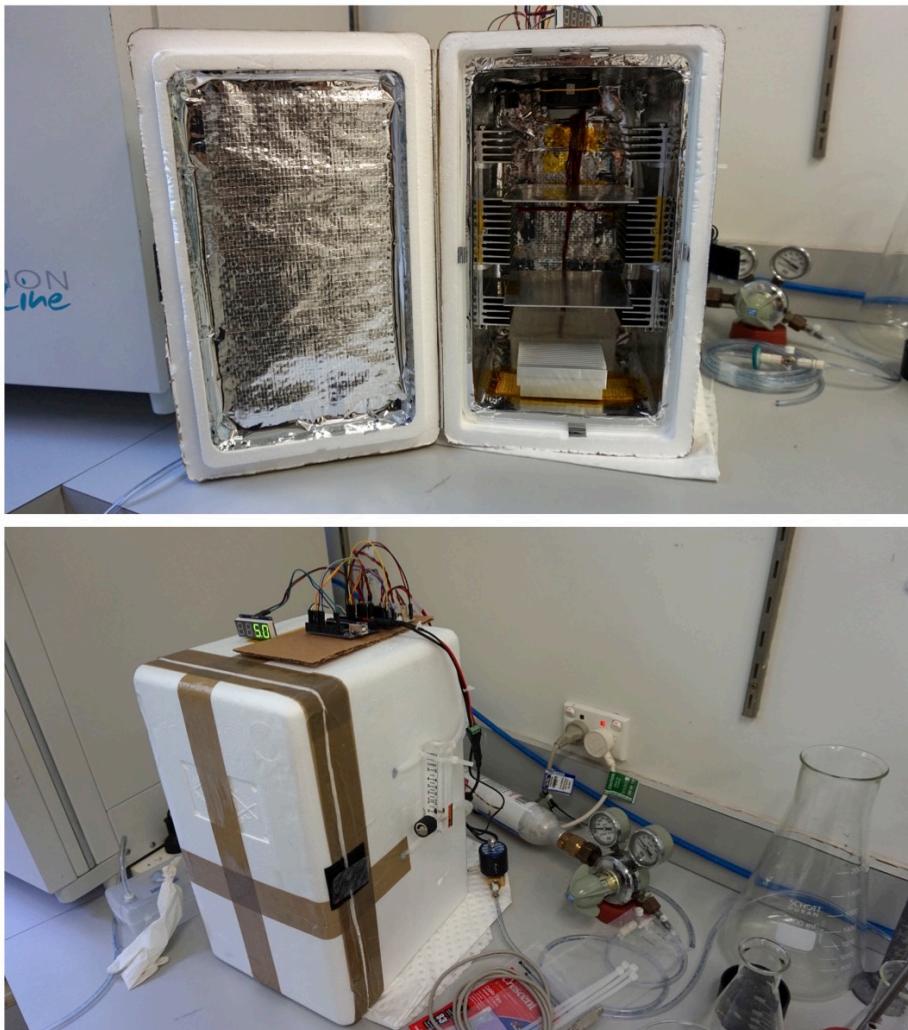

Figure 6-6. Example of a DIY CO_2 Incubator for mammalian cells. This incubator was built almost entirely from garbage and reclaimed parts. See reference 13 for instructions.

If you have access to a conventional lab equipped for mammalian cell culture (or want to create your own DIY lab), we have a few general comments about using the decellularized apples as scaffolds for 3D cell culture. Most importantly, the growth of mammalian cells in the apple scaffolds does take time. The amount of time depends on several factors, such as the cell type, how large the scaffold is, how many cells are implanted, and the conditions of the media. Implanted cells rarely attach immediately to the bare scaffold given the high porosity of the cell wall and the fact that mammalian cells have not evolved to specifically attach to cellulose. Other plants may have cellulose scaffolds that are higher density and less porous compared to the apple. These plants may provide a better platform for growing mammalian cells, and the decellularization protocol should be a good place to start. Regardless, with some patience, we have found that both mouse and human cell lines (NIH3T3, C2C12, and HeLa) will grow and invade the apple scaffold over the course of several weeks. If you want to give this a try, here are a few comments and suggestions on getting started.

IMPORTANT SAFETY NOTE: Biosafety Protocols for Working with Immortalized Mammalian Cell Lines

For any work involving the mammalian cell lines described in this article, appropriate biosafety laboratory procedures should be rigorously followed. In the United States and Canada, this includes safety standards up to Biosafety Level 2 (BSL2) or Containment Level 2 (CL2), respectively. Readers are strongly advised to consult their own local/national standards before commencing work on mammalian cell lines. However, we offer a few important procedures (based on Canadian CL1 and CL2 standards) as a starting point. These items represent some of the key physical and operational biosafety best practices:

- Work in a well-designed and functional space.
- Use cleanable work surfaces.
- Use good microbiological practices.
- Conduct local risk assessments on activities to identify risks and to develop safe work practices.
- Provide training.
- Use personal protective equipment (PPE) appropriate to the work being done.

- Maintain an effective rodent and insect control program.

- Decontaminate work surfaces appropriately, in accordance with biological material in use.

CL2 standards build upon the preceding basic best practices. Biosafety and biosecurity at CL2 are achieved through operational and physical containment practices that are proportional to the risks associated with the agents handled therein. Operational practices for CL2 include administrative controls (e.g., biosafety program management, training) and procedures (e.g., work practices, PPE use, decontamination) that mitigate the risks associated with the activities conducted within the zone. Physical containment features include facility design (e.g., location, surface finishes, access control) and biosafety equipment, such as primary/secondary containment devices (e.g., biological safety cabinets, sealed transport containers) for certain activities.

The Public Health Agency of Canada (PHAC) has developed a well-described set of guidelines for appropriate biosafety practices. The information provided here is sourced from its website. The PHAC has developed a freely available book (physical copy or as a PDF), website, and an app that contain all relevant information available to the user under any circumstance. Visit the Canadian Biosafety Standards and Guidelines website for more information.

The apple scaffolds need to be sterilized and kept sterile from this point on. Incubators are ideal places for bacteria growth. If any bacteria are in the apple when you start incubating, you are going to have issues. To sterilize the apples, you can submerge the apples into 70% ethanol for no longer than 15 minutes. After the sterilization, the apple scaffold will need to be repeatedly rinsed with sterilized (boiled) water.

The next step is to make the apple scaffolds a good home for mammalian cells. This includes incubating the apples within the culture media that is used to grow the respective cell line (Figure 6-7).
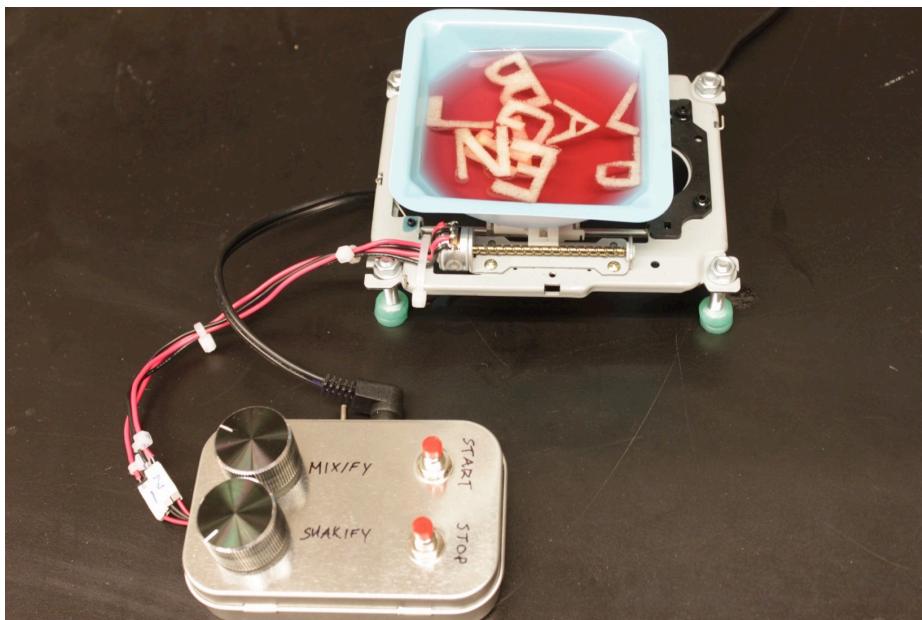
Figure 6-7. After decellularization and washing, the scaffolds are submerged in standard DMEM culture media (including 10% serum and 1% antibiotics). Agitation allows for perfusion of DMEM into the apple scaffolds and adsorption of serum proteins and growth factors to the cellulose surface.

Before "seeding" the apple scaffolds, you will need a substantial amount of cells. We recommend at least six million cells for a scaffold size of 0.5 cm x 0.5cm x 0.1 cm. Place the apple scaffold onto a preferably hydrophobic surface and add enough cell solution to fully cover the apple scaffold surface, forming a droplet on top of the apple (Figure 6-8). Any surface will do, but a hydrophobic will help produce a nice droplet. Carefully return the apple scaffold to the incubator for approximately two hours, so as to allow the cells to invade the scaffold and adhere to the cellulose.
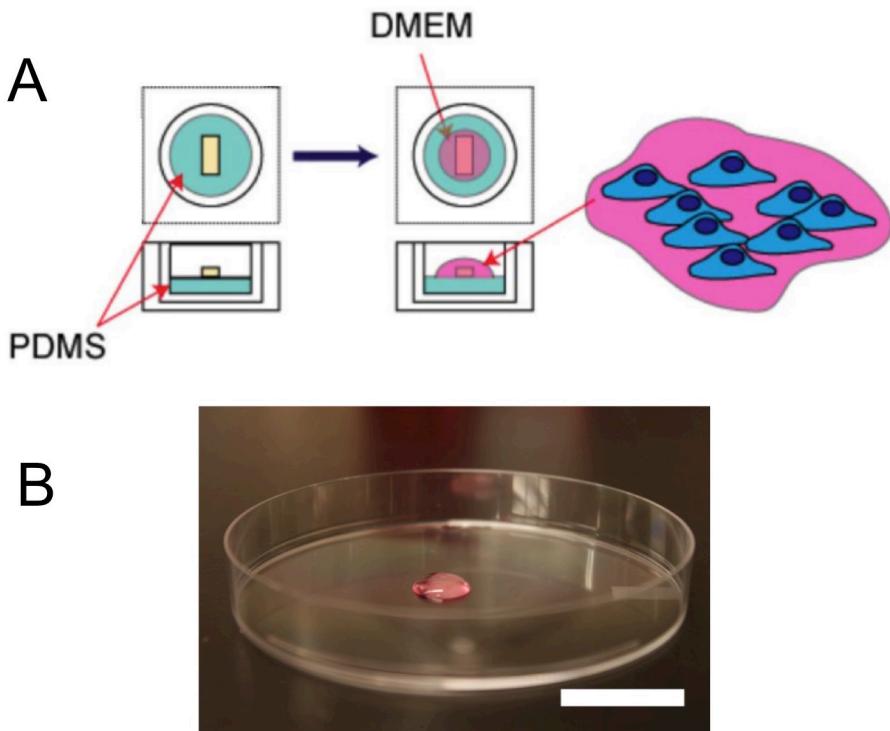
Figure 6-8. The mammalian cell "seeding" protocol for decellularized apple scaffolds. (A) Cellulose scaffolds are placed on hydrophobic surfaces, and a droplet of culture media containing a high-density cell suspension is carefully placed on top. Hydrophobic substrates were created by coating the surface of a plastic chamber with Polydimethylsiloxane (PDMS). (B) If you do not have access to PDMS, a plastic surface should work if one is careful to create a well-defined droplet of cell suspension over the cellulose. Figure is adapted from reference 8. Scale bar = 3 cm.

After seeding the apple, submerge the entire piece into culture media and leave in the incubator. If done correctly, there will be a large amount of cells within the apple scaffold, and the culture media will need to be replaced daily (Figure 6-9). Now sit back and watch your apple hybrid come to life.
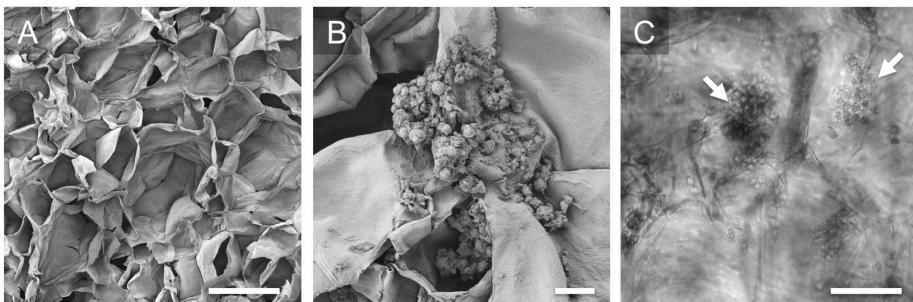
Figure 6-9. (A) Scanning electron microscopy (SEM) image of a decellularized apple cellulose scaffold (scale bar = 100 um). (B) SEM image of mouse C2C12 cells growing on the surface of the cellulose scaffold (scale bar = 20 um). (C) Transmitted light image of mouse C2C12 cells growing in the scaffold after one week in culture (scale bar = 100 um).

Conclusion

Using the preceding protocol, it is possible to easily prepare scaffolds that can act as 3D environments that support the growth of living cells. In our work, we have primarily focused on the growth of human cells in apple-derived cellulose scaffolds. However, we have observed that they will also support the growth of bacteria and fungi. And, although our focus has been on apple tissue, these protocols can be applied to other plants, and we encourage experimentation. In the long run, it remains to be seen whether such plant-derived scaffolds will have an impact in RegenMed. However, we feel this work challenges scientists to more actively consider the development of biomaterials that can be treated as open source. Typically, one of the major motivations for developing advanced biomaterials is that there are many opportunities to protect their intellectual property in order to generate economic wealth. While this closed source approach benefits a small number of people, open source biomaterials will decrease the broader cost of RegenMed healthcare, make such treatments more widely available, allow local hospitals or communities to develop the biomaterials in-house, and potentially lead to more rapid development and discoveries through community-based science. Numerous challenges remain, especially in regard to the production and safety standards of such open source biomaterials if they were to find clinical applications. However, we believe that a community-based approach to the development of open source biomaterials will be an effective way to engage stakeholders at all levels in order to develop affordable and universally accessible next-generation therapeutics.

References

1. J.J. Song, et al."Regeneration and Experimental Orthotopic Transplantation of a Bioengineered Kidney." *Nature Medicine* 19 (2013): 646–51.

2. S. Sabetkish, et al. "Whole-Organ Tissue Engineering: Decellularization and Recellularization of Three-Dimensional Matrix Liver Scaffolds." *J. Biomed. Mater. Res. - Part A* (2014): 1498–1508 (doi:10.1002/jbm.a.35291).

3. P.M. Baptista, et al. "The Use of Whole Organ Decellularization for the Generation of a Vascularized Liver Organoid." *Hepatology* 53 (2011): 604–617.

4. W. Liu, S. Thomopoulos, and Y. Xia, "Electrospun Nanofibers for Regenerative Medicine." *Advanced Healthcare Materials* 1 (2012): 10–25 .

5. S.A. Hurd, N.M. Bhatti, A.M. Walker, B.M. Kasukonis, and J.C. Wolchok, "Development of a Biological Scaffold Engineered Using the Extracellular Matrix Secreted by Skeletal Muscle Cells." *Biomaterials* 49 (2015): 9–17 .

6. J.M. Singelyn, et al. "Naturally derived myocardial matrix as an injectable scaffold for cardiac tissue engineering." *Biomaterials* 30 (2009): 5409–5416.

7. H.C. Ott, et al. "Perfusion-Decellularized Matrix: Using Nature's Platform to Engineer a Bioartificial Heart." *Nature Medicine* 14 (2008): 213–21.

8. D.J. Modulevsky, C. Lefebvre, K. Haase, Z. Al-Rekabi, and A.E. Pelling, "Apple Derived Cellulose Scaffolds for 3D Mammalian Cell Culture." *PLOS One* 9 (2014): e97835.

9. "Toxicity," December 6, 2013. *http://www.toxicitywinnipeg.com/*.

10. dannewoo, "CD-ROM Agitator," *http://www.instructables.com/id/CD-ROM-Agitator/*.

11. Jgroner, "Cheap and Easy Lab Agitator/Shaker (Great for PCBs Too!)," *http://www.instructables.com/id/Cheap-and-Easy-Lab-Agitator-Shaker-Great-for-PCBs-/*.

12. mzsolt, "PCB Agitator from an old CD ROM," at *http://www.instructables.com/id/PCB-agitator-from-an-old-CR-ROM/*.

13. Pelling Lab, "DIY CO2 Incubator Bioreactor for Mammalian Cell Culture," at *http://www.pellinglab.net/diy/diyco2incubator/*.

Affiliations

[a]Centre for Interdisciplinary NanoPhysics, [b]Department of Biology, [c]Department of Physics, [d]Institute for Science, Society and Policy, University of Ottawa, Ottawa, Ontario, Canada, and [e]SymbioticA, School of Anatomy, Physiology and Human Biology, University of Western Australia.

Acknowledgments

Daniel Modulevsky (@doc_blood) received his undergraduate degree in biochemistry at the University of Ottawa. He is currently a PhD student there under the supervision of Dr. Andrew Pelling. He is the recipient of PhD fellowship from the Fonds de Recherche du Québec–Santé and was awarded the Presidential Scholar Award from the journal Microscopy and Microanalysis in 2013.

Andrew E. Pelling (@pellinglab) received a PhD in physical chemistry at the University of California, Los Angeles, and did his postdoctoral research as a senior research fellow at the London Centre for Nanotechnology, University College London. He is an associate professor and Canada research chair in experimental cell mechanics at the University of Ottawa, Canada. He leads the Pelling Lab for Biophysical Manipulation at the Centre for Interdisciplinary Nanophysics at the University of Ottawa.

Genome Editing Systems

Vikram Dhillon

Researchers are often interested in placing specific genes into bacterial or eukaryotic cell lines for further research into cell physiology and function. A common set of techniques used in molecular biology research include the manipulation of genetic sequences (genotype) and observing the impact on the organism (phenotype). In the past few years, genetic manipulation has been supplemented by huge advances in sequencing and the rise of cheap, efficient, and accurate sequencing technologies. The two technologies combined allow for an unprecedented amount of precision in locating individual base pairs. An enormous amount of knowledge can be obtained about a potential sequence and its function in normal conditions versus in pathological states. This vast obtained information allows for the construction of a vehicle for editing the genome inexpensively and accurately. Such a vehicle would allow us to locate a specific sequence in the DNA and perform a simple manipulation. Being able to replace DNA sequences with engineered sequences using a vehicle would encompass these separate parts into a larger whole, as a comprehensive system: a genome editing system.

Naturally occurring genome editing systems serve to provide immunity to their hosts (bacteria or archaea) against viral predators. More generally, the genome editing systems provide protection against foreign DNA molecules and mobile genetic elements by neutralizing them. This neutralizing response primarily constitutes the immune response of the bacterial or archaea. Viruses present a persistent threat to their microbial hosts because they outnumber their hosts in the natural environment or ecological settings. The bacterial response to viral invasion was historically understood as a restriction modification of foreign DNA where un-methylated DNA would be recognized as foreign, or a mutation of the host receptors so that the viruses are unable to attach to the cell membrane. These defense mechanisms were mostly classified as a type of innate defense system.

More recently, genetic and bioinformatics analysis of the immune components in the bacteria have revealed that they are not just innate, but they also have a level of specificity to them. In addition, it was discovered that prokaryotic immunity relies on an RNA-based adaptive system to target and neutralize foreign DNA.

The CRISPR-Cas system is one such example of RNA-mediated adaptive immunity which evolved in bacterial and archaea to cope with viral attacks. This system in the host incorporates short sequences from foreign DNA or genetic elements into a region of the genome that is characterized by clustered, regularly interspaced short palindromic repeats (CRISPRs). To create a complete system, these incorporated sequences are transcribed and processed into small noncoding RNAs called Cas proteins that recognize and cleave foreign genetic material. The CRISPR system contains repetitive sequences that serve as memory of pervious encounter with pathogenic DNA, and this system is contained within the NA of the host so this record is heritable. What makes CRISPR a comprehensive system is that CRISPR-derived RNAs (crRNAs) get packaged into a surveillance complex that patrols the intracellular environment to detect foreign genetic elements.

Discovery

The acquisition of a novel genetic trait can be helpful or detrimental to the host. The new strain can confer a selective evolutionary advantage, such as antibiotic resistance. Such transfer/exchange of genetic material is not limited within the species. Bacteria can acquire new genetic material from unrelated species in a process called horizontal gene transfer (HGT). HGT allows for the uptake of external DNA through three distinctive processes: uptake of environmental DNA directly though transformation, incorporation of mobile gene elements such as plasmids through conjugation, and insertion of the DNA through bacteriophages through transduction. The acquirable traits are not all beneficial; therefore, bacteria and archaea have developed a genetic interference pathway to block HGT for conjugation and transduction. The arrays of clustered, regularly interspaced palindromic repeats (CRISPRs) have emerged as a novel interference mechanism that also provides an adaptive immune system for prokaryotes. The sequence structure of the array consists of a leader sequence followed in succession by repeat sequences. These repeats are interspersed by space sequences at regular intervals along the length. A visual description of CRISPR locus is provided in Figure 7-1.
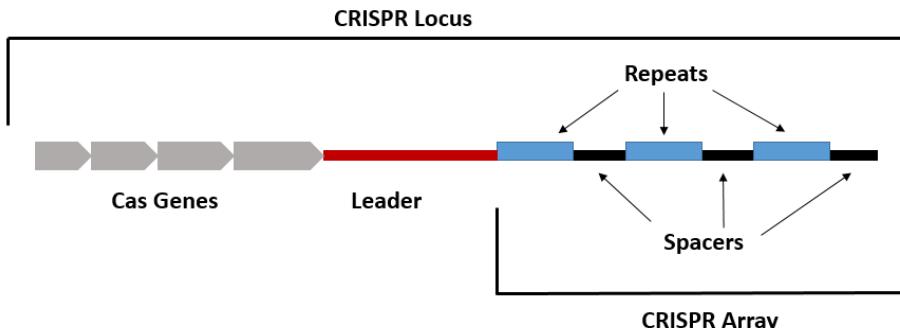
Figure 7-1. Schematic of the CRISPR Locus

The CRISPR system was discovered in 1987 when Ishino et al. were sequencing the iap gene, which functions as a conversion for the alkaline isoenzyme and noticed downstream of the gene, the presence of spacers. Incidentally, the spacers present in a CRISPR array all result from previous exposure to the viral DNA or other invading genetic elements. A few years later, several research groups also noticed the presence of spaces in archaea and classified them. The function of these spacers remained unknown until they were sequenced and recognized as being identical to the plasmid or phage DNA. The acquisition of spaces is not completely conserved. In some studies, the appearance of unique spaces was accompanied by the loss of other spacers. It is interesting to note that the disappearance of spaces correlates with a common ancestor and diverge of species all with different space units. The diversity of the spacers plays a big role in the adaptation of the CRISPR systems to the ever-mutating threat of viruses.

In eukaryotes, similar adaptive analogues exist such as the RNA interference (RNAi) pathway. There are similarities between CRISPR systems and RNAi in their mechanism of action to silence specific nucleic acids. Also, both systems use small noncoding RNAs. In addition, there exist significant similarities in the silencing complexes: RNA-induced silencing complex (RISC) in eukaryotes and CRISPR-associated complex (Cas) for antiviral defense. One interesting difference between the two systems is that of target DNA. CRISPR systems target incoming dsDNA; however, some mRNAs can be targeted as well. A more accurate comparison would be between CRISPR and Piwi-interacting RNA (piRNA), which guard the DNA from mobile genetic elements. These mobile elements are generally parasitic to their host genome; therefore, neutralizing an attack preserves the integrity of the genome as well.

Mechanism of Action

A comprehensive mechanism of action for a genome editing system like CRISPR is beyond the scope of this article. However, a generalized outline on the adaptive defensive system is outlined here. There are three general features of a prokaryotic immune system in action: adaptation, incorporation, and interference. The acquisition of invading DNA is the first step in the immune system. In this stage, short segments of DNA are integrated at the leader end of the CRISPR sequence (see Figure 7-1), and the CRISPR-repeat sequence is added for each new spacer acquired. The detailed mechanism of new spacer acquisitions process still remains unclear. However, Cas1 protein has been implicated to influence it in a few species. Some research groups have extensively studied the Cas proteins and classified them based on their abundance across species and their functional conservation.

Recall that we earlier mentioned that CRISPR was an RNA-guided DNA targeting mechanism. The acquisition of a new spacer is only the first step. Successful neutralization response against viruses requires incorporation of the spacer and transcription into these CRISPR-derived RNAs (crRNA), which are short RNA units that become integrated into large DNA-cleaving complexes. The crRNA units associate with Cas proteins, forming effector complexes where the crRNA guides the complex to the invading nucleic acid. The crRNA sequences from the complex then form base-pairing interactions with the invading sequences and cleave the invading genetic element. This process very nicely illustrates the feedback loop in the genome editing system: crRNA helps in the cleavage of invading sequences, and a portion of the invading DNA is incorporated into the spacers and finally used as memory in the case of future attacks. Figure 7-2 represents a graphical illustration of the crRNA mechanism of the CRISPR-editing system.
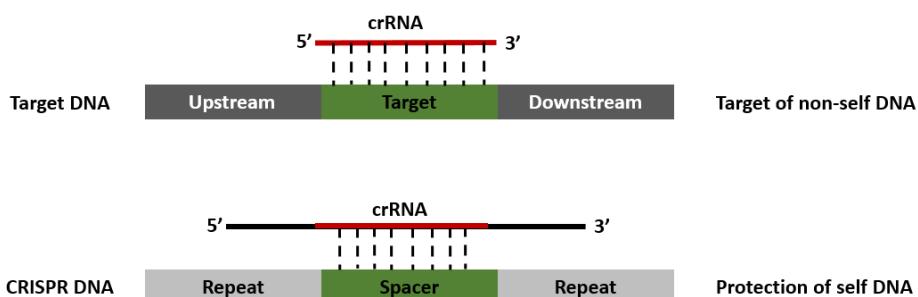

Figure 7-2. crRNAs base-pairing with self and neutralizing nonself targets

Ethical Concerns

CRISPR and Cas have become a crucial tool for genetic manipulation in biotechnology. If CRISPR systems can recognize specific sequences in the DNA and cleave them, it is not too far-fetched that such systems can be generalized to other sequences, too. One such peculiar system, the CRISPR-Cas 9, has become the standard for genetic editing, as it can be tailored to target any set of DNA sequences. The CRISPR-Cas 9 complex contains a guide RNA (gRNA) that can be manipulated to target and cleave any set of sequences. The ease and efficiency of CRIPR-Cas 9 systems has resulted in variety of investigative studies of gene function and gene therapy in animal cell lines. Recently, work led by Huang et al. tackled the issue of genomic editing in non-viable human embryos and raised a heated debate on the issue of genome editing systems. Their work revealed the serious practical and clinical limitations of the applications of CRISPR-Cas 9 systems, along with the potential dangers. Embryos contain special types of cells called germline cells, which can differentiate into a target cell type during the embryonic developmental phase. A crucial advantage of editing embryonic germline cells is that the systemic genetic diseases that affect all body systems can be eliminated. The mutation is removed from the germline itself, so when it differentiates into the cells for a mature fetus, none of the developing or dividing cells contain the mutation anymore. The work done by Huang et al. work demonstrated the flip side of the coin: The edits in the germline cells can be passed on and potentially cause unknown changes. Additionally, CRISPR-Cas 9 editing is not as targeted in embryos as expected and many unknown off-target edits may occur leading to further complications. After their work was published, a very high profile debate on the issue of genome editing with CRISPR-Cas 9 systems ensued. Many of the key figures in the biotechnology field asked for a call for moratorium on germline editing. The purpose of this request is so that the research in this area can be stopped until we further understand the implications. The result of this call was that the International society of stem cell research supported the public discourse and broader discussion of the unexpected consequences of such gene editing.

Charo et al., along with Doudona's group, highlight two different schools of thought regarding the editing of the human germline. The first one embraces balance between benefit and risk toward a risk new endeavor such as germline editing. The second one embraces limitations of what we ought to edit in nature. Both outlooks have supporters. Nevertheless, it is generally agreed by both sides to halt research that may have unintended consequences. Another study by Gantz et al. demonstrates the importance and risks associated with those unintended conse-

quences in germline editing using CRISPR-Cas 9 system for *Drosophila*. Their study focused on creating homozygous mutants in half the time as compared to traditional techniques, making it an incredibly lucrative tool for research. The authors highlight several future applications of being able to create such mutants for pest control and more importantly, for controlling malaria and eventually gene therapy. Off-target editing is a major concern in this study, too, and the nature of the induced mutations in their study is the rapidly propagating mutations that could affect wild populations. The authors propose absolute restrictions on any laboratory doing this research to have 100% containment. The National Institutes of Health issued a statement on germline editing, agreeing with the experts to not continue germline editing until the mechanisms are better understood. In addition, its stand was that it would not fund any such research either.

This is arguably the most important debate for our generation. A new system for successful gene therapy is on the horizon. Still, stretching this technique further and further into the future is turning out to be dangerous. Just like the burning wings of Icarus, the experts warn us about using caution until the system is perfected in animal models. This brings us to interesting crossroads where we finally have a mechanism to deal with genetic disorders such as cystic fibrosis and replace the faulty gene with a normal one. However, using the current version of the genome editing system is not safe and may have several unintended consequences that might be just as harmful. How much control can we exert over natural genetic reprogramming? Or put differently, how much control should we have over editing any germline cells? Science begets progress. Therefore, halting a new technique right in its tracks is counterproductive. Moreover, it is very difficult to give clear and binary answers to such difficult ethical issues. That's why in this article, we propose an outline to shape a frame of thinking, instead of seeking clear-cut answers to difficult questions. With more discourse, the momentum will shift in one direction and CRISPR research will achieve breakthroughs in precision, minimizing the off-target editing. Eventually, this research will be applied to human germline cells, but with new protocols, different resources, and new aspirations.

> *Man is an animal with primary instincts of survival. Consequently his ingenuity has developed first and his soul afterwards. The progress of science is far ahead of man's ethical behavior.*
>
> **CHARLIE CHAPLIN**

References

1. B.D. Baltimore, et al., "A Prudent Path Forward for Genomic Engineering and Germline Gene Modification," *Science* 348.6230 (2015): 36–38.

2. Luciano A. Marraffini and Erik J. Sontheimer, "CRISPR Interference: RNA-Directed Adaptive Immunity in Bacteria and Archaea," Nature Reviews Genetics 11.3 (2010): 181–190.

3. Rotem Sorek, Victor Kunin, and Philip Hugenholtz. "CRISPR—A Widespread System That Provides Acquired Resistance Against Phages in Bacteria and Archaea," *Nature Reviews Microbiology* 6.3 (2008): 181–186.

4. Puping Liang, et al., "CRISPR/Cas9-mediated gene editing in human tripronuclear zygotes," *Protein & Cell* 6.5 (2015): 1–10.

Vikram Dhillon (@opsbug), recently graduated from the Burnett School of Biomedical Science at the University of Central Florida. You can email him at dhillonv10@ucf.edu.

Analyze Your Own Microbiome

Richard Sprague

As a long-time software developer, I used to think that the machine architecture of a person is based on human genes, the genetic instructions crafted in our DNA, but over the past 10 years, an explosion in the use of low-cost genetic sequencing has revealed that *human* DNA makes up only a tiny portion of the DNA in our bodies. In fact, if we put all human hardware into a counting machine, it would find that we are roughly:

- 99% human by weight
- 10% human by cells
- 1% human by DNA

With such a tiny amount of our bodies' DNA residing in human genes, whose values are fixed at birth, it's no wonder that getting the DNA in our genome analyzed (e.g., by sending a sample and $99 to 23andme) often turns out to be disappointingly *unactionable*. With rare exceptions, few of the 23andme results can suggest behavior improvements that you didn't already know: eat plenty of fruits and vegetables and get some exercise.

The rest of you—the 90% of your cells, the 99% of your DNA—comes from non-human organisms, mostly bacteria, living on and within you. Because bacterial cells are so much smaller than human cells, they live invisibly all over your body: on the skin, in the nose and mouth, inside your intestines. There are so many that collectively they weigh several pounds, about the same as your brain, and they change over time in response to their environment. In fact, this is the most intriguing thing about your microbiome: you can alter it through diet and environmental changes.

For the past year, I've been studying my bacterial genes using low-cost 16S rRNA technology to analyze an entire ecosystem of microbes—my microbiome—in one swoop. I take a body sample (a swab from the inside of my nose, for example, or from some soiled toilet paper), and send it to a laboratory that can loosen all the genes from the millions of organisms inside, spin it in a centrifuge to skim off a precisely weighted subset that is common to bacteria alone, and then feed it to a high-powered gene sequencing machine that digitizes whatever genes it finds and compares it to large databases of known bacteria.

The most popular laboratory is a San Francisco startup called uBiome, that originally began as an Indiegogo campaign, but has since handled tens of thousands of samples from people like me who are curious to know what a 16S analysis will reveal about their own microbiome. For less than $100, you send the company a sample, and a few weeks later, it gives you access to a private web page that will show you a breakdown of the types of bacteria they found (Figure 8-1).
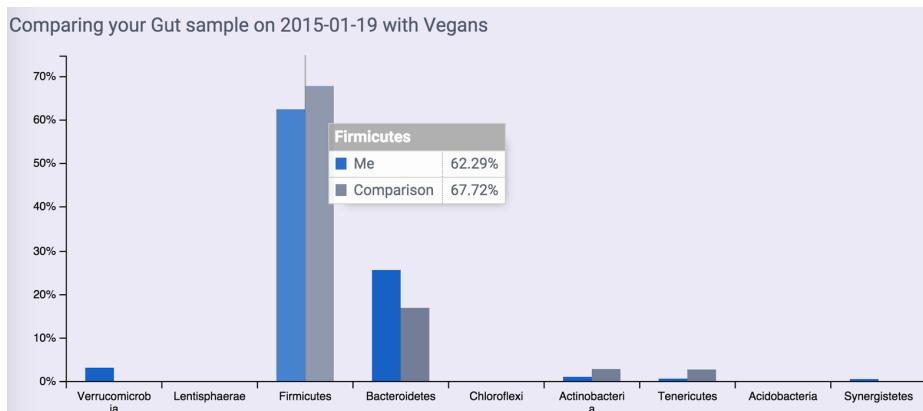
Figure 8-1. Sample uBiome results from my gut sample

The best part about uBiome is that they also give you complete access to the raw data, so you can analyze it in much more detail yourself. The website lets you download either a JSON-formatted version of the overall taxonomy, or the raw "read" data collected directly from the expensive, high-end, next-generation sequencing machine in the company's lab.

This article will show you how to study the uBiome JSON taxonomy data through easy filtering functions in a spreadsheet like Excel, as well as some open source tools using the free R statistical programming language.

The raw data, often more than 50 MB of the familiar ACGT letters of the genetic code for all those bacteria, is difficult to analyze without some bioinfor-

matics knowledge, but I'll get you started with the basics and show you some references for how to learn more.

Working with uBiome Taxonomy Data

Before analyzing the JSON data, it helps to understand something about botany. Since Carl Linnaeus in the 1700s, the science of taxonomy divides all life into seven major categories (called *ranks* in uBiome's JSON schema): kingdom, phylum, class, order, family, genus, and species (which I was taught in sixth grade to remember by the mnemonic "King Philip came over for girl scouts").

Bacteria make up their own kingdom. Just as the animal kingdom includes everything from humans to jellyfish to beetles, the diversity of bacterial life is enormous, a point which can't be emphasized too much. This is true at every rank in the taxonomy. Even two organisms that are the same at a lower rank, like genus, might have radically different effects on the human body, just as a member of the animal genus *canis* could be anything from a wolf or coyote to a Chihuahua.

You cannot mix and match these ranks. If you know something about the number of organisms in one genus, for example, this is meaningful only in comparison to the numbers of another genus. Keep that in mind during our analysis.

The uBiome JSON taxonomy data includes the following fields:

count
> An absolute measure of number of organisms found in the sample. Without knowing the size of the sample, or how many times the DNA inside was processed through PCR amplification, this number doesn't mean much except in relation to other counts *at the same taxonomical rank*.

count_norm
> A "normalized" version of the count, based on some uBiome number crunching, but you can think of it as parts per million: each unit is 1/10,000th of a percent.

tax_name
> This is the classification of the organism based on the level of its taxonomy. If you were looking at a human being, for example, you would see *Homo sapiens* if you selected tax_rank = species, but you'd see *mammalia* if you selected tax_rank = class.

tax_rank
> This indicates the level of the taxonomy. In daily conversation about animals or plants, we usually refer to the species (e.g., *Homo sapiens*), but sometimes

it's more useful to talk about bigger groupings of related organisms. For example, humans are members of the class mammalia, along with tigers and horses. If this spreadsheet were counting organisms at the level of class mammalia, the `count_norm` would almost certainly be bigger than the `count_norm` for humans alone, unless humans were the only type of mammal found in the sample.

`taxon` *and* `parent`

These help identify the ranking in a more precise way by pointing out which `tax_ranks` are subsets of which. For example, bacteroidia above has a parent = 976, meaning that it is a subset of the taxon 976, bacteroidetes. When you follow the various taxons and parents up the chain, you'll see they all end in the superkingdom bacteria, which has a taxon of 2. The values for these numbers, incidentally, are taxonomical numbers from the curated database at NCBI, the national bioinformatics center run by the U.S. government. Enter the number into the taxonomy browser at the NCBI Taxonomy Browser and you can learn as much as you want about that organism.

`tax_color`

This doesn't matter for this anaysis, but uBiome software uses this to colorize its pretty graphs to make them more readable.

Get the Raw Data from uBiome

After your uBiome sample has been submitted and processed, you will be given access to a private web page with some basic tools to help you understand your results. For the merely curious, this will be plenty of information; but if you want to do more sophisticated analysis, you'll need access to the raw data.

Fortunately, uBiome makes this very easy. Look around the results page you received after your sample has been processed, and you'll find two options: "Download taxonomy" and "Download raw data." The taxonomy is the easiest, so let's start there. Click that link to go to a page with a bunch of apparently garbled noise like this:

[screenshot raw JSON]

```
`{"ubiome_bacteriacounts": [{"count": 79951, "count_norm":
1010733,"parent": 0, "tax_rank": "root", "tax_color": null, "avg": null,
"taxon": 1, "tax_name": "root"}, {"count": 79102, "count_norm": 1000000,
"parent": 131567, "tax_rank": "superkingdom", "tax_color": null, "avg":
null, "taxon": 2, "tax_name": "Bacteria"}, {"count": 43, "count_norm":
544, "parent": 203692, "tax_rank": "order", "tax_color": null, "avg":
null, "taxon": 136, "tax_name": "Spirochaetales"}, {"count": 2,
"count_norm": 25, "parent": 135621, "tax_rank": "genus", "tax_color":
```

null, "avg": null, "taxon": 286, "tax_name": "Pseudomonas"}, {"count": 7, "count_norm": 88, "parent": 91347, "tax_rank": "family", "tax_color": null, "avg": null, "taxon": 543, "tax_name": "Enterobacteriaceae"}, {"count": 17195, "count_norm": 217378, "parent": 171549, "tax_rank": "family", "tax_color": null, "avg": null, "taxon": 815, "tax_name": "Bacteroidaceae"}, {"count": 17013, "count_norm": 215077, "parent": 815, "tax_rank": "genus", "tax_color": null, "avg": null, "taxon": 816, "tax_name": "Bacteroides"}, {"count": 1043, "count_norm": 13186, "parent": 816, "tax_rank": "species", "tax_color": null, "avg": null, "taxon": 820, "tax_name": "Bacteroides uniformis"}, {"count": 459, "count_norm": 5803, "parent": 375288,`

That's the raw data in uBiome's special JSON format of just the highlights found by the Ilumina gene sequencing machine in its lab.

Analyze uBiome Results in Excel

You'll want to convert the JSON into CSV, which can then be read into Excel for much easier analysis.

If you are analyzing a single JSON file, the simplest way to convert is to Select All the information on that page, and then copy/paste it into *https://json-csv.com* or any similar free site. If you have several samples, this can get tedious, so alternatively you can use the tool *"convert_json_files_to-csv"* on my open source GitHub site (more details to follow).

With that out of the way, the next step is to tell Excel to apply a filter to the whole sheet. In Mac Excel, I just select the filter icon, as shown in Figure 8-2:

Figure 8-2. Excel filter

The first row is transformed into a nifty filtering device. Note how the right side of each cell has a little upside-down triangle. Select that and a new pop-up menu will appear that lets you sort and filter the column however you like. Figure 8-3 shows how it appears when we look just at the tax_rank = phylum and

sort the count_norm in descending order (I also hid the avg column to simplify the image).

E	G	H
count_norm ▾	tax_name ▾	tax_rank ▾
584498	Firmicutes	phylum
274848	Bacteroidetes	phylum
65624	Actinobacteria	phylum
7699	Verrucomicrobia	phylum
3590	Proteobacteria	phylum
1176	Lentisphaerae	phylum
695	Synergistetes	phylum
544	Spirochaetes	phylum
430	Tenericutes	phylum
240	Acidobacteria	phylum
88	Chloroflexi	phylum
38	Ignavibacteriae	phylum
25	Cyanobacteria	phylum

Figure 8-3. Excel view filtered by phylum

The count_norm column now corresponds exactly to the percentage breakdown in the fancy charts on the uBiome site!

The chart from the uBiome site is shown in Figure 8-4. Compare it to the Excel chart shown in Figure 8-5.
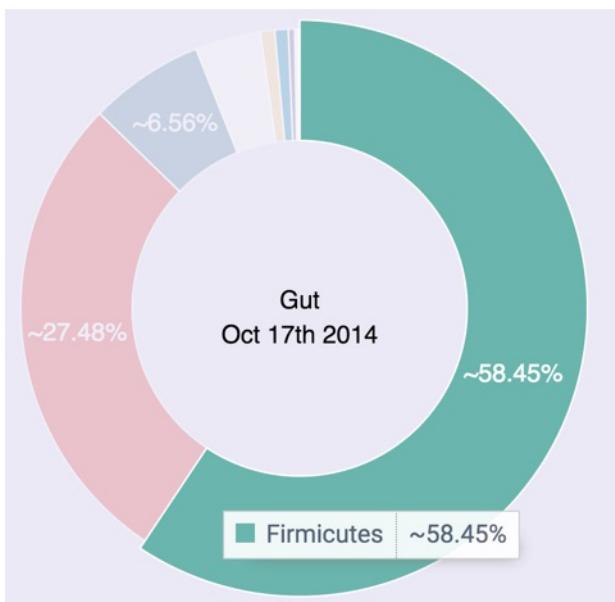
Figure 8-4. An Excel donut chart made with the count_norm data by filtering for tax_rank=species
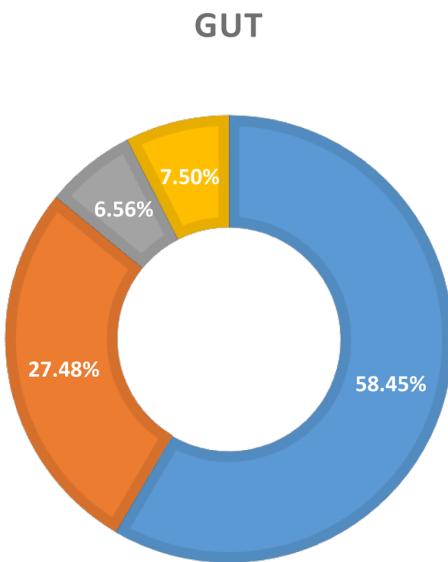
Figure 8-5. Excel chart

See? Same info!

But now that it's in Excel, you can do much easier analysis than what you get currently on the uBiome site. For example, try selecting `tax_rank` = species and then sort `tax_name` alphabetically. Now I can look alphabetically at all the different Bifido species found in my sample (Figure 8-6):

| E	G	H	J	K
count_norm	tax_name	tax_rank		
19001	Bifidobacterium adolescentis	species		
37534	Bifidobacterium animalis	species		
76	Bifidobacterium catenulatum	species		
1858	Bifidobacterium longum	species		
25	Bifidobacterium ruminantium	species		

Figure 8-6. Filtering for species that begin with "bifido"

If you're already comfortable with Excel as a way to review data, you'll find it an easy way to explore your uBiome results as well. This works great for analyzing a single sample, and it's my first step upon obtaining a new data set.

Incidentally, note that the JSON data includes species information, which is not provided by the standard web tools. 16S technology is not good enough to reliably determine information below the genus level, so uBiome refers to this JSON data as "experimental," and for that reason, makes it available only to those with the technical skills to find it. One study, for example, finds that 16S is > 90% reliable for genus identification, but only 65%–83% reliable for detecting species. As a hobbyist, that's good enough for me, but it's a reminder that uBiome should not be used for serious medical purposes. If you're sick, see a doctor.

Compare and Study Multiple uBiome Results

Once you have obtained more than one sample, either from yourself, or from another individual you'd like to compare to yours, you'll soon want to do more sophisticated comparisons. To make that easier, I've written a collection of simple utilities in the R programming language that is popular among statisticians and

scientists. My utilities, as well as all the data files from my own uBiome samples, are freely downloadable from GitHub.

You won't need to know much R to run these examples yourself, but it may be helpful to learn a few basics. I suggest you start by downloading the excellent RStudio package, and then skim the FAQ files at the homepage for the R project.

I recently submitted two samples to uBiome, each taken a week apart. For several weeks before and during this period, I carefully tracked all the food I ate using the free app MyFitnessPal, hoping to get some insights into how diet affects my own microbiome. Here's how I used these utilities to conduct the analysis.

Begin by setting my working directory using the R command `setwd` or from within the R environment. Then load the utility files using the R `source` command `setwd("uBiome/sources")`:

```
source("uBiomeCompare.R")
A<-read.csv("sprague-ubiome-150421.csv") # you may need to include the entire
# path if your CSV files are kept in a different directory
B<-read.csv("sprague-ubiome-150428.csv")
```

How do the samples compare? I'll create a new variable `A_vs_B` and asssign it to the results of running the `uBiome_compare_samples` function:

```
A_vs_B<-uBiome_compare_samples(A,B, rank="species")
head(A_vs_B)

##                       tax_name count_change
## 1        [Ruminococcus] obeum        -6949
## 2  Acidaminococcus fermentans          112
## 3 Adlercreutzia equolifaciens         -637
## 4          Alistipes finegoldii        1421
## 5        Alistipes indistinctus          19
## 6         Alistipes onderdonkii        1017
```

By default, `uBiome_compare_samples` returns a result sorted alphabetically. Here's how to sort by species count. For brevity, I'll show only the top and bottom 10 values:

```
sortAB <- A_vs_B[order(A_vs_B$count),]
head(sortAB,10)

##                                  tax_name count_change
## 54              Coprococcus sp. DJF_CR49       -25077
## 81 Peptostreptococcaceae bacterium TM5       -23030
## 71            Methanobrevibacter smithii       -22461
## 43            Clostridium clostridioforme       -19006
## 8                    Anaerostipes hadrus       -18478
```

```
## 88              Ruminococcus bromii      -17735
## 31               Blautia faecis          -13980
## 41             Clostridium baratii       -13565
## 1             [Ruminococcus] obeum        -6949
## 29         Bifidobacterium animalis       -6455

tail(sortAB,10)

##                          tax_name count_change
## 33        Brachyspira sp. NSH-25          2952
## 27    Barnesiella intestinihominis       5577
## 79 Parasutterella excrementihominis     5641
## 30          Bilophila wadsworthia        5819
## 85         Roseburia inulinivorans       5898
## 69           Lactobacillus rogosae       9910
## 26          Bacteroides uniformis       12303
## 86            Roseburia sp. 11SE38       25807
## 61    Faecalibacterium prausnitzii       59182
## 14           bacterium NLAE-zl-P430      66457
```

Negative numbers indicate a drop from the first sample; positive numbers mean there are more in the second sample than in the first. The numbers on the left are references to the original R data frame row before the sort.

As noted earlier, all references to count or count_change in these examples are using uBiome's normalized count, which you can think of as units of parts per million. Dividing by 10,000 gives the percentage.

In this example, I see a large increase in *Faecalibacterium prausnitzii*, often considered a marker for health. I'll want to look at my meals over the test period to see what might have driven this positive change. Using the macronutrient data I collected daily with MyFitnessPal (exported to CSV with the handy DesignByVH exporter, I produced the simple chart shown in Figure 8-7 to see if my eating habits had an effect on my microbiome.

	Calories	Carbs	Fat	Protein	Cholesterol	Sodium	Sugars	Fiber
Average (month)	1841.7	192.2	102.7	94.9	268.0	2298.5	64.1	15.2
Average (Week)	2242.6	241.9	124.7	108.7	262.3	2814.3	78.3	16.9
Difference from Ave	400.9	49.7	22.0	13.9	-5.7	515.7	14.2	1.6
% Diff from Ave	122%	126%	121%	115%	98%	122%	122%	111%

Figure 8-7. Macronutrient summary from MyFitnessPal

As you can see, though I apparently ate about 20% more overall during the week of my test, my dietary cholesterol was conspicuously lower than normal. Did this affect my microbiome? My next step will be to search the medical and biology

literature to see if there are any known relationships between dietary cholesterol and Faecalibacterium. Will I find something new? That's the exciting part about this type of citizen science: armed with my own data and imagination, new discoveries are everywhere.

Analyze a uBiome Sample Yourself

The uBiome JSON taxonomy information will be more than enough for most people new to biology. But for BioCoders and anyone interested in expanding their bioinformatics skills, the raw data files are the place to look.

FASTQ files are the text-based results of a run at your sample using one of Illumina's high-end gene sequencers.

For example, the first line in my data is:

```
@NS500457:29:H2HLGAFXX:1:11101:16760:1033
1:N:0:ACTTACA+NTGCGAA
```

This contains information about the gene sequencing hardware uBiome used for this sample. The other lines in the raw text file contain the actual base pairs detected by the sequencer, along with quality information to indicate the certainty of a particular read. In my case, the entire file is over 170,000 lines long, and that's just one of the eight files in the FASTQ archive. That's a lot of data!

Although it's possible to analyze this data on my own in a tool like BioPython, it's much easier to submit the raw data to a public gene processing server, such as the Metagenomics Analysis Server (MG RAST) hosted by Argonne National Labs at this site:

> **NOTE** You must use the Firefox browser (not Chrome, IE, or Safari), and before submitting any jobs, you must apply for a (free) login ID, which you receive by email in a day or two.

After logging in with your new account ID, click the "upload" icon on the home page and follow the process to prepare and submit your data:

PREPARE DATA

| 1. prepare your metadata |
| 2. upload files |
| 3. manage inbox |

DATA SUBMISSION

| 1. select metadata file |
| 2. select project |
| 3. select sequence file(s) |
| 4. choose pipeline options |
| 5. submit |

Figure 8-8. MG-RAST submission home page

1. Ignore the part about preparing the metadata; you can add it later.

2. Upload files. The site offers to let you upload from the web browser, but I've had better luck using the API and the following `curl` command:

   ```
   curl -H "auth: <your web key>" -X POST -F
   "upload=@ssr_13603__R1__L001.fastq" "http://api.metagenomics.anl.gov/1/
   inbox/" > curl_output.txt
   ```

   I just type this from Terminal on my Mac. Note that you'll need to get your own webkey (hit the "view webkey" button on the submission page), which you will substitute into the preceding `curl` command. Also change the filename part of the "upload=@" section to point to the correct file. Note that uBiome fastq files are compressed, but it's OK to upload them as is. You'll uncompress them on the server after the upload.

3. Manage your inbox. Once the file has successfully uploaded, you'll see it in your inbox. At this point, you should "unpack selected" if you uploaded a compressed file. Depending on how busy the server is, it may be minutes or even

hours of waiting and pressing "update inbox" before you see the file information shown in Figure 8-9:

**File Processing Operations**

| unpack selected | Unpacks selected zip, gzip, bzip2, tar gzip, or tar bzip2 files. | demultiplex | Demultiplexes selected files. |
| convert sff to fastq | Converts selected sff files to fastq format. | join paired-ends | Joins FASTQ-formatted overlapping paired-end reads. |

[OctGutSample] - 0 files
ssr_13603__R1__L001.fastq

File Information

ambig char count	844
ambig sequence count	744
average ambig chars	0.019
average gc content	49.921
average gc ratio	1.030
average length	148.634
bp count	6477024
creation date	2015 May 18 17:46:26
file checksum	28815210c80bb754fb9465994cf93ff8
file name	ssr_13603__R1__L001.fastq
file size	15922154
file type	fastq
length max	151
length min	32
sequence content	DNA
sequence count	43577
sequence type	Amplicon
sequencing method guess	illumina
standard deviation gc content	5.433
standard deviation gc ratio	0.238
standard deviation length	15.433
suffix	fastq
type	ASCII text
unique id count	43577

Directory Management Operations

update inbox	Refreshes the contents of your inbox.
move selected	Moves the selected files into or out of a directory.
delete selected	Deletes the selected files.
create directory	Creates a new directory in your inbox.
delete directory	Allows you to select and delete an empty directory.

Figure 8-9. Managing your MG-RAST file submission inbox

4. Next, you're ready to submit the data (Figure 8-10).

DATA SUBMISSION

1. select metadata file

2. select project

3. select sequence file(s)

4. choose pipeline options

5. submit

Figure 8-10. Submitting your data to MG-RAST

Here are my suggestions for this section:

- Check the box to say you won't supply metadata.

- Create a new project and select it.

- Select the files you uploaded.

- Use the default values to choose pipeline options.

- Submit the job, being sure to make the data publicly accessible immediately, to increase the priority.

There you go! Now a job has been submitted, and you wait for it to complete. It usually takes a day or two, after which you'll receive an email notification. After that, click the bar chart icon in the upper-right of the page (Figure 8-11).
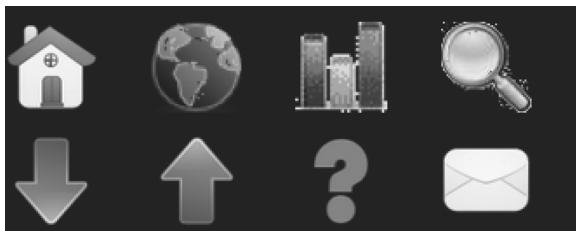
Figure 8-11. Click the bar chart icon to go to the analysis page

This brings you to an analysis page where you can study your data in much more detail.

When I select my data under "Metagenomes" on the analysis page, I'm offered several options for annotation sources. Because the uBiome fastq data comes from 16S ribosomal RNA, I choose to compare my data against one of the large databases of known RNA. The popular Greengenes database is one of them, so I select it, and then ask to generate a table of the results (Figure 8-12).
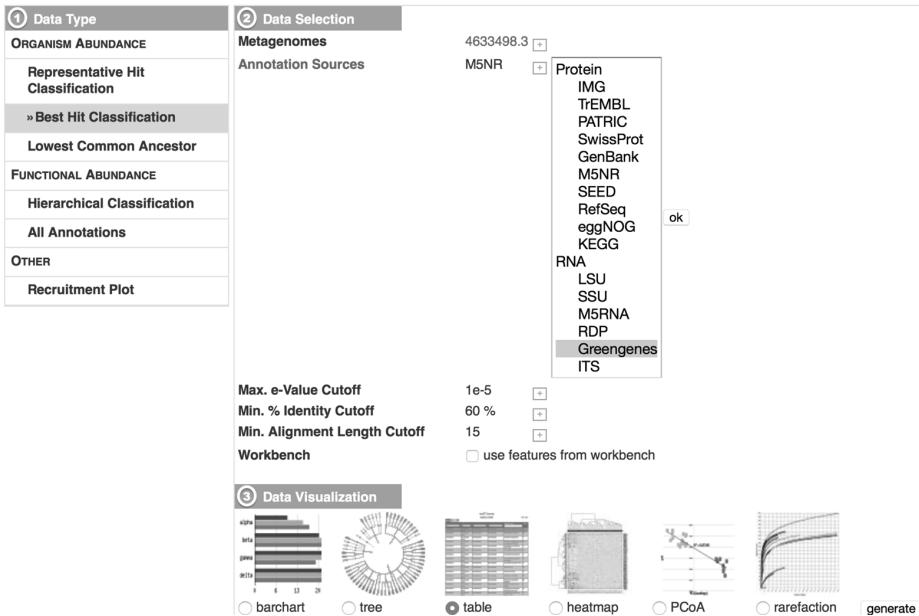
Figure 8-12. Select one of the RNA databasesand hit the "generate" button

The table, sorted by abundance and filtered to phylum, looks like Figure 8-13.

Figure 8-13. Table view of organisms found in this sample, as matched to the Greengenes database (try RDP or M5RNA for slightly different annotations)

The results roughly conform to what we see from uBiome: the most common organisms are the same in each case, and although the scales are different, the abundance computed corresponds roughly to what we see from the uBiome count field. A comprehensive analysis, which requires processing and uploading all the FASTQ files, is beyond the scope of this short article. Experiment for yourself!

Other MG RAST tools let you compare your data against other databases of bacteria and proteins, graph the results, and perform sophisticated statistical analysis to determine other unique features of your sample. You can also compare your sample to microbiomes that have been uploaded by others, enabling you to study your results in much more detail.

Limitations

There is so much analysis possible that it's easy to overestimate what the technology can do, so it's worth pointing out some of the limitations. The swab that I submit represents only a tiny portion of my overall gut microbiome, and because the bacteria are unlikely to be evenly distributed throughout the sample, this is at best an approximation to the actual ratios present in my body. Furthermore, I'm measuring only the organisms that *exited*; it's possible that many of the most bioactive species are more integrated into the lining of the colon and this test may not see them at all.

That said, my experience over many samples is that the uBiome results are consistent enough that, yes, many conclusions are actionable as long as we keep the limitations in mind. My results are broadly similar between samples and always consistent with the ranges reported by other "healthy omnivore" submissions that uBiome recieves from others with diets and health histories similar to mine. By contrast, my results do not much resemble uBiome customers on a vegan diet or those with self-reported obesity.

Fortunately, 16S technology has a well-proven track record of accurately identifying the organisms that it finds, so I can probably trust information about the overall level of diversity. Because diversity tends to correlate with health—and is manipulable based on what I eat—this is perhaps the most interesting aspect anyway.

Conclusion

There is much more to analyze. The human body is many orders of magnitude more complex than anything I've encountered in computer software. Access to the raw data will expose you to many of the uncertainties and contradictions surrounding much of what science knows and doesn't know about the microbiome. But the field is very young, and armed with your own data plus a little coding knowledge, you can find interesting, actionable ways to improve your own health while contributing at the cutting edge of science.

References and Further Information

Buy the 16S rRNA gut microbe test kit (under $100) from uBiome. For $400, you can purchase a five-site kit to test more of your microbiomes: skin, genitals, mouth, nose.

Although I've not tried it myself, another similar test is available from The American Gut Project. If you are looking for medical-grade testing and you (or your insurance company) are willing to pay a few hundred dollars more, ask your doctor about Genova Diagnostics.

To use the MG RAST Server (MG RAST):

http://bit.ly/mcs-rast
http://bit.ly/anl-rast

Track your eating and nutrition with MyFitnessPal and then export your data using the DesignByVH exporter.

To read more about the microbiome and its effect on health, I highly recommend the book *Missing Microbes* by Martin Blazer, as well as *An Epidemic of Absence* by Moises Velasquez-Manoff. For online information about the care and feeding of specific gut organisms, see Dr. Grace Liu's site, the Gut Institute.

All the raw data discussed here, as well as the sources to all the utilities I use to analyze my own microbiome, are on GitHub.

Richard Sprague (@sprague), co-founder of Ensembio, Inc., has been building and marketing consumer software internationally since the 1980s at numerous technology companies including Apple, Microsoft, and several Silicon Valley startups. With more than a decade working in Japan and China, he leads development of personal self-tracking products for the Chinese market, with clients including market-leading calorie-tracking app MyFitnessPal. He has an undergraduate degree from Stanford University, and an MBA/MA from the University of Pennsylvania, Wharton School. He lives on Mercer Island, Washington with his wife and their three teenagers.