Comprehensive Environmental Assessment and Synthetic Biology Applications Workshop
The Woodrow Wilson International Center for Scholars, Science, Technology & Innovation Program
Ronald Reagan Building and International Trade Center
1300 Pennsylvania Ave., NW, Washington, DC 20004-3027
Thursday, July 28th, 2011, 8:30am to 4:30pm

CONFERENCE NOTES VERSION 1.0 PREPARED AUGUST 2, 2011

TABLE OF CONTENTS

PARTIALLY ANONYMIZED NOTES

PREPARED PRESENTATIONS	NAMES ARE INCLUDED
FACILITATOR COMMENTS AND QUESTIONS	NAMES ARE NOT INCLUDED
PARTICIPANT COMMENTS AND QUESTIONS	NAMES ARE NOT INCLUDED

RULES FOR CIRCULATION

ALL CONFERENCE PARTICIPANTS WILL RECEIVE THIS VERSION OF THE NOTES (VERSION 1.0).
REFERENCES TO CONFERENCE PROCEEDINGS SHOULD BE BASED ON THESE NOTES

About the CEA workshop (Opening Remarks)
The goal of this workshop was largely to try and generate a set of questions and research areas that would enable a more comprehensive risk assessment of synthetic biology applications. To do this, participants are asked to consider a scenario where cyanobacteria (*Synechococcus elongatus*) engineered to produce sugars at an industrial scale have escaped into the surrounding environment. Participants were asked to follow a Comprehensive Environmental Assessment (CEA) framework currently used by the EPA's National Center for Assessment (NCEA) to evaluate risks associated with nanomaterials, and consider each component of the CEA and discuss potential knowledge gaps and research areas for synthetic biology applications.

It is important to note that this was NOT a risk assessment of *S. elongatus*, but rather a test drive of the CEA framework to gauge its usefulness AND to generate research priorities to enable comprehensive environmental risk assessments of synthetic biology applications.

Major questions that this workshop aimed to answer were:
- Is the CEA framework useful for understanding risk and supporting risk assessments?
- Can we create a frugal tool or decision tree that can guide researchers in dealing with the ecological issues of synthetic biology applications?
- Will current risk assessment methodologies work for synthetic biology applications?
- What information/data would we need to be able to accurately assess the environmental risks of synthetic biology applications?
- Is there a set of research areas or questions that would be important to support future environmental risk assessments?
- What may be early stage risk management opportunities for synthetic biology researchers?

Session I: Introductory Presentations 9:15 – 10:30

Presentations:
- *S. elongatus* and sucrose production – *Daniel C. Ducat & Patrick Boyle, Harvard*
- Important gene flow concepts – *Allison Snow, Ohio State (via phone conference)*
- CEA framework and approach – *Genya Dana*
- Case study scenario of *S. elongatus* in a photobioreactor system – *Genya Dana*

A. Introduction to *S. elongatus* (P. Boyle)

S. elongatus is a good candidate for genetic engineering. It is naturally transformable (taking up DNA from the surrounding environment) and uses homologous recombination. Additionally *S. elongatus* colonies grow at good rate, doubling once every 24 hours. It also has relatively few nutrient requirements, largely needing CO_2 and sunlight along with small amounts of trace metals and nitrogen.

Because of *S. elongatus* requirements for sunlight and CO_2, it must interface with the environment (unlike *E. coli*) and therefore has a high likelihood of environmental release.

These requirements also create issues from an economic point of view. The economics of *S. elongatus* production get better as you scale up (more volume for less cost), but containment costs get worse (more surface area and greater opportunities for escape).

Questions posed:
Can this geometric/economic issue be addressed early in the genetic engineering phase?

How do you regulate an organism that uses sunshine and CO_2 and produces other things?

B. On Sucrose Production (D. Ducat)

Sugars are a biotech energy "currency." For cells to make desirable end products, such as biofuels, they need sugar inputs. Cyanobacteria could potentially be grown in nutrient poor environments, creating a source of sucrose that would not compete with food production or interfere with environmentally rich areas.

Sucrose is a compatible solute whose concentration in the cell can be greatly increased without damaging or poisoning the cell. Sucrose can even protect cells in some stressful environments. Sucrose is a solute that *S. elongatus* builds up in the cell in response to osmotic stress to maintain an appropriate osmotic balance with the external environment, for instance, when a freshwater cyanobacterium is put into salt water.

To make the sucrose produced by *S. elongatus* accessible, we need to get the product out of the cell efficiently and cheaply. By engineering the cell to produce the protein CscB and placing the cell in a salt solution, CscB will export sucrose out of the cell. However, such salt water use and sucrose extraction will decrease the growth of the microbial communities.

Rough early estimates suggest that sucrose production via *S. elongatus* could potentially be competitive with sugar cane sucrose production.

Production Method	% Sucrose	%Supportive Biomass
Sugarcane	20%	80%
S. elongatus in 100mM of salt	~30%	~70%
S. elongatus in 150mM of salt	~50%	~50%
S. elongatus in 200mM of salt	~70%	~30%

This sucrose production would be used for downstream biotech products, not for sugar. However, the levels of sucrose produced by this method would be very dilute for these biotech activities. Additionally, other organisms may be needed to remove salt which would reduce economic efficiency of this system.

In summary, this production is not perfect, but has potential. Sucrose production increases with culture density. Production by *S. elongatus* is continuous and doesn't need complex removal schemes. Additionally, it is non-competitive with food production.

Audience questions raised by this product:
How confident can you be that you are not turning on a silent operator in the genome that makes toxins or something else? *S. elongatus* is not known to produce toxins or have appropriate functional pathways for needed genes such as polyketide synthases, but if function is limited by regulatory regions, not synthetic genes, then could one turn on a silent regulator?

There is a risk with monocultures- how susceptible is this process to contamination (natural or anthropogenic) and how do you deal with it?

Do the cultures act as single cells or communities and how will this change how they survive/thrive in the environment?

If *S. elongatus* is naturally transformable, how stable are they genetically? Will they stay the same or change overtime? What is their rate of evolution?

Introduction to Gene flow concepts (Allison Snow)

To understand the movement and persistence of genetically modified organisms (GMOs) and/or their DNA, we need to look at horizontal gene transfer.
There are 3 mechanisms of horizontal gene transfer:

1. Transformation of free DNA: Free DNA from other, possibly dead, organisms is taken up by the host organism
2. Conjugation: The transfer of DNA from one organism to another through a plasmid or bridge-like connection
3. Transduction: The transfer of DNA from one organism to another by a virus

The persistence of this engineered DNA will largely be determined by the fitness cost, neutrality or benefit it provides.

To talk about risk assessment in general, and the *S. elongatus* case in particular, we will make the following general assumptions:

- Physical containment is not practical at a large scale production system. We should assume the GMO will enter local environment and disperse widely.
- The GMO can survive in the environment into which is released.
- The GMO has reduced fitness compared to its wild-type. We assume there is some fitness cost from introducing the engineered DNA.

Some initial questions raised by gene flow involving GMOs are:

- Under what conditions do you see this reduced fitness?
- How much fitness reduction is needed to cause rapid extinction in all cases?
- Could DNA from dead GMOs be taken up and utilized by other microbes?
- Which other microbes can take up free DNA/exchange DNA with living GMOs?

For example, consider GM crops. Some plants can survive and become volunteers. Canola can easily start new populations (they are volunteers) where corn and soy do not likely to volunteer. Survival of these GMO volunteers plus rapid evolution can result in several outcomes:

1. GMOs can die out due to fitness cost that reduces their competitiveness.
2. GMOs can volunteer and form new population. If there is sufficient genetic diversity, would this population be able to evolve and create a high-fitness feral strain? (endo-ferality)
3. GMOs can volunteer and interact/combine with wild populations and create hybrid strains with high fitness (exo-ferality)

To understand the likely outcomes from these volunteers, we need to answer some general questions about rapid evolution:

1. If transgenes from GMO interact with wild populations, in other words there is gene flow, how important is level of exposure (consider low vs high exposure)?
2. Horizontal gene flow is very common in nature. Can we assume that populations are perfectly adapted at all times?
3. Or are the engineered genes stronger, enhancing fitness and increasing population growth rates? (e.g. antibacterial resistance)
4. Could the spread of the GMO and engineered DNA lead to unwanted consequences?

In this workshop we will ask: Can we answer these questions? What other information or research would we need to answer these? Do we need more discussion among experts from different disciplines? Do we need new research to inform risk assessments? Will consensus and research be sufficient for a thorough understanding of risks?

Audience questions raised by considering gene flow:

What constitutes extinction in the environment- how do you evaluate what that really is?

> Ideally, extinction would mean the death of all the GMOs, but this is unlikely and the DNA may still persist despite the death of the organism.

Some genes used in GMOs, such as resistance genes, already exist in nature. Do we regulate naturally occurring genes? Do we need to remove them from GM organisms? What about other genes?

> From a regulatory perspective, we do care about using naturally occurring genes. We do not want to add to an already existing problem of resistance development. Even though resistance genes occur in nature, they do not normally occur in such high concentrations. Using these genes in GMOs increases the likelihood or other organisms taking up the genes. So we are not only concerned just with artificial genes, but artificial concentrations of natural genes as well.

It is hard to maintain these cultures in monoculture production systems, so at some point someone may introduce pesticide production or herbicide resistance genes. Such genes would make engineered microorganisms hardier. Will these engineered genes be taken up by other organisms in nature? How likely is it that this will occur?

There are other methods of containment for GMOs other than physical containment, such as suicide genes and nutritional constraints. How confident are we in other methods of containment?

> These methods would contain the organism, but not the DNA. No matter how you contain the GMO, the DNA will be released into the environment which could be of big concern.

CEA framework and approach (Genya Dana)
Comprehensive Environmental Assessment framework- used for nanomaterials- tries to think very holistically about how materials or organisms move through the environment and change throughout the whole life cycle of a product.

Overarching Questions: What kind of info do we need to know to evaluate the organisms' movement and impacts? Where are the information gaps? What are our research priorities?

To help answer these questions, this workshop considered a cyanobacteria scenario:
A plastic bag photobioreactor system grows *S. elongatus* in an industrial area, near marine waters, and is subjected to a major environmental event (e.g., hurricane) that results in large scale release of the organisms.

Initial thoughts/questions raised by participants:

- How fast do these organisms change with time? We may end up with endless numbers of similar but genetically different versions. Are we worried about rapid evolution?
- We don't have comprehensive knowledge of genetic sequences of different strains already existing in nature. Do we need to compare them to our modified versions? What baseline would we use to compare GMOs to unmodified versions when evaluating risk?
- There is a gap in knowledge of the genomes found in nature that would potentially interact with modified genomes. How would they interact? How likely is functional DNA transfer?
- What happens to the animals that eat these GMOs?

Session II: Lifecycle of *S. elongatus* 10:30 – 12:30

Key Questions:

- What may be potential escape routes from the production system?
- What may happen to the organism upon escape: e.g., does it survive?
- What might be important research questions re: escape routes and organism survivability?

Facilitator: Initially we assumed the product is in the Product Manufacture/storage stage of its life cycle when released- The first major question is does the product survive or not? What information do we need to answer this?

Thoughts from the participants:

- We need a lot of information about the surrounding environment and the organism's compatibility with native flora and at different salinity levels to determine if it would survive.
- We would need to know the "fitness cost" of the engineered gene and what fitness cost would encourage rapid fall off or "extinction".
    - Models of evolution can predict the pervasiveness of specimen-based on fitness cost, population size, and competition with wild populations.
- We would need to be able to track these organisms when they're released. Otherwise we would not be able to separate them from wild populations to learn valuable information about the impacts of release. We would also need the instrumentation and methods to track this data.
    - Metagenomic analyses in general are improving rapidly, improving our prospects for tracking populations of microbes.
    - Sandia National Labs, working on bioterrorism, is able to break up DNA into small chunks for analysis. They are trying to develop genetic forensics, which could be useful for synthetic biology applications.
    - We can engineer in an innocuous marker (a barcode or watermark) so we can track GMOs among natural populations, enabling us to collect the information necessary to create mathematical models, conduct future risk assessments and facilitate remediation.
- Some cells can persist in dormant or resting states. Our definition of viability needs to be expanded to include temporal refuge. How long can they survive in a dormant state? What does survival mean?
    - What kind of experiments or lab studies could predict survival through dormancy or resting? How long would they take?
    - FDA requires that the genetic construct is stable and stays within the organism, and that it can be tested.
    - On a moral plane, do we object to the idea that something we've modified or "created" may persist for a million years even if it is relatively innocuous?

Facilitator: Now assume that *S. elongatus* dies and the engineered DNA persists. What sort of information will we need to know to evaluate risks related to DNA persistence?

Thoughts from the participants:

- Will the wild population take up the DNA just because it's there?
    - Organisms in aquatic environments take up DNA fairly readily. DNA can persist in some environments for a long time.
- How do we handle disposal of the organism or its product?
- What biological adjustments may reduce the likelihood of these events?

- How do you evaluate the probability of these "rare events" (mass escape)?
- What sort of tests would regulators require?
- How do you allocate funding to research for developing these tools and test?
- Do we know the prevalence of the natural gene in nature?
    - We'd need to redo DNA viability studies with newer methods.
- Are there benchmark levels of natural genes? Can we identify genes that are so prevalent in nature that we would not be concerned about re-introducing them? Is there a threshold where we no longer consider the DNA a great threat?
    - It depends on the product of the gene- does it makes sucrose or does it produce something more sinister or toxic
    - Scale matters
- We choose organisms because they are easy to modify (i.e., for their genetic handle). Doesn't that also make us more concerned about their ability to transfer DNA to nature? By choosing these organisms, how much are we increasing the probability that engineered genes will be picked up by other organisms?
    - Can we do a study to understand the genetic transfer in different environments? Can we measure gene transfer?
- Are there characteristics that can be checked in the host organisms to ensure that they do not have toxic or blooming abilities?
    - There are countless products/by-products that these cells produce or can produce that are harmful. We need to have a profile of how the genome and products of the cell are changed by the addition of engineered genes (through functional genomics, transciptomic and proteomic analysis).
- It is fairly cheap to run a sequence and gauge what the cells are doing on daily basis. Have they changed? Are they the same cells you started with? Are they behaving as desired?

Session III: Environmental Compartments and Organisms 1:00 – 2:30

Key Questions:

- Which environmental compartments may be exposed?
- What external factors might influence escape and movement into the compartments?
- What external factors might influence exposure/uptake of the organism or DNA?
- What kinds of research questions would be important to answering the previous questions?

Facilitator: Much of the discussion has already centered on water. Are there further things to consider in this environmental compartment?

Thoughts from the participants:

- Based on the cyanobacteria example, we are worried about creating (purposefully or inadvertently) organisms that can cross the salinity barrier from

freshwater to brackish water, resulting in toxin producing populations in new environments; this sort of problem does exist in nature.

- There is a limited amount of resources available for product research and risk assessment. How do we allocate these resources when considering the different environmental compartments? Are there priorities? Are some environmental compartments more important or in need of funding than others?
    - From a regulatory perspective, the EPA deals with companies and products on a case by case basis. There is no prescribed format for information required for approval. No checklist. Should there be? What would this include?

Facilitator: What about other environmental compartments? What about air?

Thoughts from the participants:

- Based on what naturally occurs we can determine if these organisms are transferable by air. Because information on air transport already exists, a literature search may be sufficient. We may not necessarily require lots of money and years of research for air transport. Air transport can also be evaluated by mathematical models- significant work already done has already been done in this area.
- On the other hand, air plays a huge role in transport. This could make site-specificity irrelevant. An escape in one place can easily be a global escape. So how does this impact risk assessment?
- Borrowing from the human health perspective, we should ask what informs the risk assessment to the greatest advantage. What are we most afraid of and how do we gauge this? In the human health sphere there is a list of seven deadly sins that raise red flags- changing the host environment, etc. Can we generate a similar list for synthetic biology? What are the major changes or possible things that we might change that should raise red flags? What should we be focused on?
- The similarities of the GMO to the wild type are both good and bad. It's reassuring because the GMO is not that different from what already exists, but it's bad because this enables easy gene transfer. If we make organisms increasingly artificial, on what criteria should we evaluate? In 10 years, we may not be talking about this wimpy sucrose producing organism sitting a bag, but a more robust organism in an open pond making something other than sucrose. What then should we be considering? On what criteria would we evaluate these?
- Expanding our view from sucrose producing cyanobacteria, let's consider the possibility of modularity. Joule may be looking at creating a cyanobacteria chassis with modularity and plug-in components. We are talking about large scale production of organisms that can produce different high value items with

just small tweaks in DNA. If we're not overly worried about sucrose, we are definitely worried about other materials that could be produced using these microbes.

- Need consider the industrial parks of the future with multiple GMO's producing multiple products in close proximity. Are there special risks in this case?
- If we engineer algae to produce compounds that are different from what's in nature, aside from gene transfer, we also a have the very real possibility of accidental and large production of such compounds.
- Is there a template from the pharmaceutical companies that we can consider? They largely regulate the product, but also regulate the process.
  - The microorganisms are highly contained by the use of HEPA (high efficiency particle air) filters and they incinerate waste and spills. However, scale and containment issues would prevent us from adopting their template. Production outdoors on an industrial scale is very different from limited indoor production by pharmaceuticals.
- The debate has largely centered on 2 different issues. First, what is the impact of genetically engineered material being present in the environment? Second, how do scale, production methods and end products influence the risk assessment needed?

Session IV: Research Priorities	2:45 – 4:00

Research Topics:

I. Stability of DNA- Investigating rates of evolution and changes of functionality
II. Persistency- Compatibility and survivability of organisms, dormancy and resting stages
III. Fate and transport of genetic material.- Does the target gene remain functional in other hosts
IV. The physiological differences and functionality between the wild and novel organism
V. Probabilistic modeling of gene transfer

This portion of the workshop aimed to generalize the questions generated by the cyanobacteria scenario.

Facilitator: Knowing that risk assessment can vary greatly depending on the product, can we generate any foundational research needed that would apply to issues across the board, so that when the organisms become more robust, and we are doing large scale, outdoor production we can do a thorough environmental risk assessment? Is there a set of investigations we can do in order to know more about how to assess the risks? Is there anything we can generalize?

Thoughts from the participants:

- Ultimately we need to investigate rate of evolution for changes in functionality. In terms of robustness, you want the rate of evolution to be as low as possible, and the organisms to only divide when necessary. Cell reproduction can be slowed, but evolution is definitely sped up by scale. We can study genetic diversity vs time. However, we may not be measuring what we want. We can easily measure changes in

base pairs, but it is harder to measure changes in functions. We can also do studies in comparative genomics, but the original culture changes over time as well.
    - We need to use functional genomics to look at the whole physiology of the new cell. (General omics)
- Investigate survivability and compatibility (genetic and organismal) with native flora and fauna under different environmental factors.
    - To test organism and DNA compatibility, we can "create a big soup, dump the stuff in and see what happens." We can also introduce DNA fragments and measure what is picked up. However, we do not have standards for how long to stir the soup, what type of soup. We would also need to study what is "natural", to put the results in context.
    - Regarding survival, we need to be concerned with competitiveness. How many survival competition tests do we need? It should be in a whole community analysis. We need to consider everyone (the grazers), not just the competitors.
    - Dormancy and resting stages are important. How long can an organism rest?

- Are we concerned about DNA transfer, the added risk of putting the DNA in a new location, or is scale the major issue?
- From the regulatory perspective, we need to remember that people can come in with an application for a desert location and once it's approved for commercial process, they can use that product anywhere unless the EPA has put restrictions on the product. In order to put these restrictions in place, the EPA would need to know beforehand what to consider and what they should restrict.

- How can we track GMOs in the environment? Genome sequencing is becoming cheap enough that we can barcoding/watermarking novel organisms. This helps manage risk, but also provides tools to trace and monitor GMOs in case something happens.
- Investigate the activity of assimilated DNA; does the target gene become functional in other hosts? This can be answered with "soup" experiments and functional genomics tests. Is this level of research on gene transfer practical for all genes or just for ones where the trait is particularly harmful?
- Can we develop better modeling? Can the modelers guide the parameters and data needed to predict gene uptake? The entire gene may not be adapted. What happens if new organisms pick up only parts of the new gene? It is doubtful that we will ever have the capability to model gene uptake and transfer. However, models can be used to predict the physical spread of the organism.
    - Would it be useful to have a model that separates naturally occurring genes into ones that are prevalent enough that we can assume they have been thoroughly sampled throughout evolution and ones that are rare enough that we cannot be confident of how they will react in nature? Can we create a threshold of exoticism for genes to guide us?
- Understand secondary metabolites. How many should we look at? At what concentrations are we concerned? Some secondary metabolites, though possibly harmful can also be useful.

- We could run toxicity tests as you would for other substances.
- From a business model point of view, this knowledge is likely to be generated in the process of product development/optimization and therefore easily available.

These generalized research questions and areas where then further condensed into 5 broader categories. Participants were asked to vote on which they felt were the most important and briefly say why. Below is a summary of the research categories, the smaller initial questions they encompassed, the number of votes they earned and why.

Research category	Specific questions	Vote tally	Reasons given by participants
Rates of evolution and changes in functionality	• Investigate the rate of evolution for changes in functionality.	0	NA (no votes)
Survival and persistence of the organism	• Is the organism compatible with the environment and other populations? • Can the organism survive in a dormant or resting state? • What is the "fitness cost" of the engineered gene and how much of a fitness cost would encourage rapid fall off or "extinction" of the organism in the wild? • How many survival competition tests do we need? Studies should include a whole community analysis, under a variety of env conditions. Consider everyone (the grazers), not just the competitors.	5	• Encapsulates the genetic history of organism and be useful in understanding its evolution. • Companies are not expected to do a lot of work in this area; this information is difficult to come by, but important.
Fate and transport of functional genetic material	• Ability of DNA to persist after death? • What may acquire the gene? • Does the target gene remain functional in other hosts? • In what ways can the target gene alter existing genomes? • Introduce fragments of the introduced cassette and measure what is picked up by other microorganisms.	11*	• A non-scientist would be very interested in this. A risk assessment would certainly need to cover this to satisfy the public. • Fills in gaps, leads to useful information for both regulation and development of organisms. • Least understood of what we talked about today and therefore most interesting. • Most relevant from the policy perspective. • A risk we don't understand • * limiting fate of genetic material

Physiological differences and functionality between the wild and novel organism	• What is the natural risk of these wild organisms? • How do we compare the additional risk due to novel genes? • Investigate secondary metabolites. How many should we look at and at what concentrations? • What are cells doing on a daily basis? Have they changed? Are they the same cells you started with? Are they behaving as desired? • Generate a profile of how the genome and the products of the cell are changed by the addition of engineered genes.	9	• Captures a broad understanding of the organism before it is modified and allows you to compare your modified organism to something. • By focusing on this category, we would be addressing issues contained in research categories 1 and 2 • Need to know before we can say if the new organism will change ecosystems • This category has the least amount of available data • This represents the hazard part of the risk assessment which is important • This will be the trigger of regulation • This information is important for the 1st step for the risk assessment and will temper what questions you ask in other areas
Probabilistic modeling of gene transfer	▪ Can modelers guide the parameters and data needed to predict gene uptake? ▪ Would a model separating naturally occurring genes prevalent enough to assume that they have been thoroughly sampled throughout evolution from ones that are rare be useful? Can we create a threshold of exoticism for genes to guide us?	0	NA (no votes)

Other issues raised by participants

- There needs to be an increase in the accessibility of data. Both research communities and the public need better access to research to make better decisions.
- On the CEA framework:
  - There is no component on economics, which is the controller, so a feasibility check should be early on.
  - It was useful to remind researchers of areas they may have forgotten or neglected.
  - EPA's Office of Pollution Prevention and Toxics looks at all these components during MCAN applications, though not necessarily in this organized fashion.
  - Provides an orderly way of thinking about, but the EPA sometimes like to combine compartments because they are not isolated in reality.

Closing Remarks (Ken Oye, David Rejeski, Genya Dana)

What are the next steps to be addressed in the future?

We need to talk about steps to mitigate risks in engineering such organisms. Another exercise could be to do a risk assessment on something that is closer to commercialization. How do we mitigate these risks, how do we test this risks?

We need to discuss bioremediation. Organisms for this purpose cannot be engineered to be weak, and they are not contained. They must be more artificial in order to eat more synthetic materials. What are the risks from such organisms, and how do we deal with them?

We need to discuss more novel, orthogonally contained organisms. Re-engineered *E. coli* from another Boston lab is an obvious case study organism that we should look at.

In general, when do we intervene and how in the engineering process? How do we balance IP concerns and product development?

What would be the "7 Deadly Sins" for creating genetically modified organisms (and what should we call them)?

End of discussion##