Summary Report of the Meeting to Discuss Data Needs and Testing Methods for Assessing the Safety of Environmental Introduction of Synthetically Designed Algae for Biofuel Production

A Joint Workshop of the Woodrow Wilson Center, the MIT Program on Emerging Technologies, and the U.S. EPA

December 14, 2012

Report Date:
February 5, 2013

Report Authors:
Shlomiya Bar-Yam
Kelly Drinkwater
Todd Kuiken
Julie McNamara
Scott Mohr
Ralph Turlington
Kenneth Oye

NOTICE

This report was prepared by the Wilson Center and MIT PoET as a general record of discussion from the workshop *Data Needs and Testing Methods for Assessing the Safety of Environmental Introduction of Synthetically Designed Algae for Biofuel Production.* This report captures the main points and highlights of the meeting. It is not a complete record of all details discussed, and it does not interpret or enlarge upon statements made over the course of the meeting that were incomplete or unclear. All included points represent the individual views of meeting participants and should not be viewed as a consensus. Except where specifically noted, no statements in this report represent analyses by or positions of any of the meeting hosts or report authors.

MEETING SUMMARY

During 2011 and 2012, The Woodrow Wilson International Center for Scholars (Wilson Center) and the MIT Program on Emerging Technologies (PoET) co-hosted multiple workshops studying varying aspects of synthetic biology. Discussions repeatedly illuminated conspicuous data gaps throughout the field, with such uncertainties often significantly impeding forward progress on safety and security discussions. These findings, in combination with an interest by the U.S. Environmental Protection Agency (EPA) to more explicitly consider the ramifications of such gaps in the face of future Toxic Substances Control Act (TSCA) applications, led to the organization of this meeting, *Data Needs and Testing Methods for Assessing the Safety of Environmental Introduction of Synthetically Designed Algae for Biofuel Production*. Co-hosted by the Wilson Center with funding from the Alfred P. Sloan Foundation, MIT PoET, and EPA, the meeting was designed to bring together industry, academics, non-governmental organizations, and several agency stakeholders to discuss what the relevant data gaps are—and how they might be addressed—when considering the implications of environmental release of synthetically engineered organisms. For the purposes of considering a tangible concept, the meeting was conducted specifically through the lens of algae engineered to produce biofuels

Contents

Agenda

This workshop provides an opportunity for academia, industry, government, and non-governmental organizations to improve scientific understandings of ecological issues of relevance to evaluation of synthetically designed organisms, and to identify what methods exist or should be developed to assess the safety of a field release regardless of statutory or regulatory mandates. This workshop is one of a series dealing with scientific issues surrounding synthetic biology put on by the Wilson Center and the MIT Program on Emerging Technologies, with the support of the Sloan Foundation and the NSF Synthetic Biology Engineering Research Center. These workshops have assessed risks; identified scientific uncertainty associated with synthetic organisms and their interaction with the environment, and developed research agendas to address some sources of scientific uncertainty.

Session I: Introduction		8:00-8:30
Welcome	David Rejeski, Woodrow Wilson Center	
Overview of Schedule	Kenneth Oye, MIT and NSF SynBERC	
Self-Introductions	All Participants	

Session II: Overview of Current and Emerging Industrial Applications		8:30-10:15
Synthetic Genomics	David Hanselman	
Algenol	Pat Ahlm	
Sapphire	Yan Poon/Tim Zenk	
UCSD	Stephen Mayfield	
Agilent Technologies	Stephen Laderman	
Discussion:	What are the properties of algae optimized for biofuels production?	

Break 10:15 – 10:30

Session III: Overview of Current Review Process		10:30-11:00
EPA perspective	Mark Segal	
DOE perspective	Daniel Fishman, Kristen Johnson	

Session IV: Previous Workshops - Addressing Data Needs		11:00-11:45
EPA findings from 1990s workshops on biotechnology	Gwen McClung	
Wilson Center and MIT findings from previous workshops	Todd Kuiken	

Session V. Identification of Ecological Endpoints to be Assessed 11:45-12:00

Lunch 12:00-1:00

Session V. Identification of Ecological Endpoints to be Assessed 1:00 – 3:00
Defining potential receiving environments (terrestrial, freshwater, marine)
Defining endpoints within these environments
Defining immediate vs. long-term data needs to assess endpoints
Defining minimum data set needed prior to any environmental introduction vs. data set needed for large scale acreage
Panelists: Rex Lowe, Bowling Green University; Bruce Tonn, University of Tennessee; Kent Redford, Archipelago Consulting; Robert Stevenson, Michigan State University; and others.

Break 3:00-3:15

Session VI. Methodology & Protocols 3:15-4:15
Methods/Tools
Instrumentation

Session VII. Wrap-Up 4:15-5:00
Summary of data needs
Summary of instrumentation needs
Identification of areas of uncertainty
Identification of research paths to address uncertainty

List of Acronyms

ASU	Arizona State University
CBI	Confidential Business Information
CWA	Clean Water Act
DARPA	Defense Advanced Research Projects Agency
DNA	Deoxyribonucleic Acid
DOD	Department of Defense
DOE	Department of Energy
EISA	Energy Independence and Security Act
EPA	U.S. Environmental Protection Agency
FIFRA	Federal Insecticide, Fungicide, and Rodenticide Act
FISH	Fluorescence *In Situ* Hybridization
FOIA	Freedom of Information Act
GFP	Green Fluorescent Protein
GM	Genetically Modified
GMO	Genetically Modified Organism
HAB	Harmful Algal Bloom
HGT	Horizontal Gene Transfer
HPLC	High-Performance Liquid Chromatography
JBEI	Joint Bio-Energy Institute
MIT	Massachusetts Institute of Technology
NIH	National Institutes of Health
NREL	National Renewable Energy Laboratory
NSF	National Science Foundation
ORNL	Oak Ridge National Laboratory
OPPT	Office of Pollution Prevention and Toxics
PBR	Photo-bioreactor
PERF	Petroleum Environmental Research Forum
PNNL	Pacific Northwest National Laboratory
TERA	TSCA Experimental Release Application
TSCA	Toxic Substances Control Act
UCSD	University of California, San Diego
USDA	U.S. Department of Agriculture

Summary of Session II: Overview of Current and Emerging Industrial Applications

Representatives from Agilent, Algenol, Sapphire Energy, and Synthetic Genomics provided brief updates to inform workshop participants of the present state of the industry. Following, new research findings and possible future collaborations in the development of measurements and standards were presented. In closing, the floor was opened to all participants to discuss the characteristics of a hypothetical "ideal" organism.

Industry Update

Industry representatives covered three main areas in their presentations: methods of organism development, approaches to cultivation and containment, and mechanisms for hazard assessment. Several questions were also raised without immediate resolution.

Organism development

- Researchers aim to first identify organisms naturally displaying desired characteristics, including through conducting bio prospecting ventures in varied environments around the world. Wide-ranging screening has been made possible through metagenomics and high-throughput analysis of promising strains.
- Follow-on genetic manipulation includes support of natural and directed evolution processes. Active techniques include radiation and pathway engineering.
- Once organism performance has been tested within the laboratory, the top performing strains are moved on to larger-scale trials (e.g., pond screening). A field-validated strain is one shown to grow robustly in the field and proves capable of cultivation; a production strain is further treated to increase pest tolerance. For example, one company reported taking a production strain, subjecting it to several rounds of mutagenesis, and ultimately finding that the evolved line grew better (noting that the parasitic fungi originally in question continued to grow). Pond crashes were noted as occurring over as short a time as one to two days.

Cultivation and containment

- Cultivation and containment methods vary depending on a company's biofuel production technique. For open pond production, cultivation is akin to farming, as crops must hold up against threats such as pests and weather. For photo bioreactor (PBR) systems, most external cultivation threats are controllable.
- Containment methods can be broken out as biological and physical:
    - **Biological.** The hazard assessment process (described in the following section) precedes strain scale-up and aims to eliminate the most overt of biological threats, such as invasiveness and toxicity. One pilot facility is beginning to test the invasiveness of strains in each type of water an escaped organism would encounter prior to reaching the ocean.

    Preliminary findings have seen no evidence of invasiveness, although the screening has only newly begun and the adequacy of the mesocosms has not been verified.
  - **Physical.** With PBR systems, physical containment is focused on structural soundness of the PBRs to prevent leaks; concrete pads and earthen berms to prevent spreading should a leak occur; and comprehensive treatment of effluent. One company cited physical containment levels designed to meet 500-year storm threats. For open pond systems, operations proceed under a general assumption of field release. One company noted that while birds and other creatures have been observed along the pond edges, it is only now moving forward in a partnership to develop monitoring tools to better characterize their presence. Finally, one open pond company is currently trialing the use of unlined (soil only) ponds.
- One company posited that metagenomic analysis could serve as a useful tool for studying an environment prior to release. Citing a study finding significant reductions in species diversity around power plant effluent as compared to in a mangrove swamp, the company noted that receiving waters could be tested for species diversity before, during, and after release as part of the general monitoring process.

Hazard assessment

- Once a species has been identified as of interest, a hazard assessment is conducted to ascertain its practicality as a commercial starting point. Several properties are instant disqualifications, including one company citing risk level rankings above Biosafety Level (BL) 1 and another screening through bioinformatics analysis to evaluate the presence of enzymes required to produce known toxins.
  - Multiple presenters noted that the hazard assessment process is regularly stymied by a lack of available information. While much information is available on a select few strains (i.e., those responsible for repeated harmful algal blooms and those already employed in commercial processes), little is available for others. Further, general taxonomy has become increasingly complex as actors have repeatedly shifted between “good” and “bad” groups.
- One company reported that of approximately 40 high level hazard analyses conducted at the genus level, only a handful have subsequently resulted in strain abandonment.
- In terms of hazard assessment, one company suggested that there was little value in identifying a strain as “native” or “non-native” based on state-level communications.
- Post initial strain selection, companies noted performing various types of horizontal gene transfer (HGT) studies prior to advancing strains further. Additionally, beyond the initial bioinformatics analyses, high-performance liquid chromatography (HPLC) is routinely performed to assure toxins are not being produced.

Other

- During the presentations, multiple companies highlighted data and/or knowledge gaps requiring further attention:
    - Additional information on taxonomy is needed, as the data are highly valuable for understanding risks yet are lacking in multiple areas.
    - No good exposure assessment model exists, though it is important and would be useful.
    - Does a well-characterized “bad bug” list exist? For example, a compilation of organisms and parts that should be avoided? One company stressed the importance of companies facilitating a collegial sharing of information so as to better advance the industry as a whole.
    - If a strain is completely non-toxic, should zero release be expected?
    - If an organism displays a three-fold increase in growth rate, what would be the implications upon escape into the wild?
    - If escape results in the replacement of one species by another, would that be considered harming the population in a substantial way?

New Research Developments

An algal researcher and product developer presented a brief summary of recent advances in the field. In particular, the participant emphasized the following key areas:

- Whereas much of the original algal biofuels research focused on freshwater species, such advances are now being tackled with saltwater organisms.
- Regulators will need to consider a wide range of products from synthetically engineered algae in the future, as many applications beyond biofuels are advancing toward commercial production. For example, the researcher explained a successful trial algal production of nutrients traditionally found in colostrum.
- In developing countries, it is unlikely that products will be able to bear the additional costs associated with production in PBRs; therefore, it should be assumed that applications will be produced in open ponds in such locations.
- Presented data displaying the successful incorporation of a synthetically engineered gene into another organism (in this instance, involving sensitivity to high- versus low-light).
- Cited research by Susan Golden identifying four traits that successfully decrease grazers.

Tools for Developing Methods and Standards

A representative from a company specializing in measurement methodologies provided an overview of their technology development process alongside emerging technologies, and highlighted some possible areas for collaboration in the algal field.

- The representative emphasized the importance of identification of current unknowns and data needs within the industry so as to allow for targeted product development, and noted the utility of public-private partnerships in such endeavors.

- As examples of current applicable technologies, microarrays—enabling the development of large libraries of sequences as well as for genome partitioning products—and oligo fluorescence in situ hybridization (FISH)—allowing high sequence specificity for targeting microbial applications—were described.

Characteristics of an Ideal Organism

In a discussion of ideal organism traits from a *production* perspective, workshop participants built off an initially prepared slide highlighting eight broad areas likely to be of focus. Following discussions, clarifications, and some points of debate, the following list was derived (with qualifiers noted):

- Enhanced photosynthesis
- Enhanced lipid production
- Rapid growth
- Enhanced nutrient uptake, production, or utilization
- Enhanced survival in pond monoculture (e.g., resistant to herbicides, pests, and pathogens)
- Increased tolerance to adverse environments
- Allelopathic [*allelopathy:* the inhibition of growth of one species of plants by chemicals produced by another species]
- Geared toward more cost-effective sections of supply chain
- Genetic malleability, or increased ease of modifying genomes (at least initially to facilitate further strain modification)

The topic of biological containment mechanisms was discussed, though few specifics arose owing to significant knowledge gaps remaining in the area. Additionally, participants debated the merits of ease of organism traceability, though no resolution was reached. Finally, many participants expressed concern that the list was not ideal when considered from the perspective of environmental concerns. However, the participants were reminded that these traits were only being gathered so as to be able to better focus discussions later in the day regarding understanding possible ecological endpoints of modified traits.

Summary of Session III: Overview of Current Review Process

Regulators from the Environmental Protection Agency (EPA) discussed how existing rules, laws, and mandates might define the agency's role in regulating synthetically bioengineered algal biofuels. While a number of different regulatory mandates give EPA potential jurisdiction in this area, none clearly defines EPA's role or establishes a set of activities and criteria to guide such a role.

The specific laws that were discussed as potential mandates for EPA regulation of synthetic algal biofuels include the following:

- Energy Independence and Security Act (EISA)
- Toxic Substance Control Act (TSCA)
    - TSCA may apply because "new" microorganisms (depending on how this is defined) can fall under the rubric of "new chemicals," which TSCA grants EPA jurisdiction over in the context of manufacturing, importation, and research and development for commercial purposes. TSCA would also require the submission of an Experimental Release Application (TERA) 60 days prior to the introduction of microorganisms to an uncontained commercial facility.
- Clean Water Act (CWA)
    - CWA might apply because engineered organisms could be considered "pollutants."
- Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA)
    - FIFRA would apply to any disinfectants or pesticides used in the commercial growth process. At present, however, despite the existence of these several potential regulatory mandates, EPA's role is relatively undefined.

EPA's regulatory role in this area will depend substantially upon the novelty of a bioengineered organism. This, in turn, will hinge on what definitions and standards are put in place for synthetic biology and the criteria for determining whether and how genetic modifications lead to an organism possessing "new" characteristics.

EPA would likely have to expand its assessment capabilities, develop new areas of expertise, and develop new standards in order to keep pace with the current and expected pace of innovation in algal biofuel research, development, and production.

Representatives from the Department of Energy (DOE) approached the issue of synthetically engineered algae from a very different perspective than EPA. DOE's mission in this area is driven by the government's priorities under EISA and other acts in promoting the development of new sources of fuel that can effectively substitute existing fossil fuels. DOE's primary focus has been to support industry initiatives to promote the development of new technologies by offering sources of funding, and by partnering with industry to promote more effective research and development approaches.

In the regulatory area, this consists primarily of DOE offering guidance and assistance to private firms on maneuvering through government regulatory requirements, establishing research and development protocols that best minimize and address regulatory barriers to technological innovation, and cooperating with other U.S. government agencies to facilitate regulatory transparency and compliance.

DOE has also sponsored more limited work at the national laboratories on developing criteria for the assessment of new synthetic biology applications in the production of algal biofuels. This has consisted primarily of the development of a set of indicators focused on environmental and human safety.

Summary of Session IV: Previous Workshops – Addressing Data Needs

The December 2012 workshop is not the first time these topics have been discussed by EPA. In 1994, a three-day workshop was run by EPA that focused on bacteria and fungi, but not algae. As explained Gwen McClung, EPA Office of Pollution Prevention and Toxics (OPPT), Risk Assessment Division, the goal was to work toward developing different testing schemes. Some of the main conclusions from the 1994 report include the fact that TSCA does not have any specific testing requirements. It has specific informational needs, but there is no hard and fast set of rules to acquire these data. This means that regulating a GMO under TSCA is different from regulating it under the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), which has specialized, required tests. The 1994 workshop also differentiated between environmental and ecological effects of an organism. Risk assessments should be based on an organism and its respective modifications that could alter its behavior in the environment (exposure x hazard = risk). These tests would be performed in microcosm tests in tier 2. When evaluating a GMO, EPA is limited to focusing on the ecological issues associated with the immediate environment in which the organism is introduced.

1994 Workshop Notes

In 1994, the U.S. Environmental Protection Agency’s Office of Pollution Prevention and Toxics, Health and Environmental Review Division, Office of Research and Development, and Environment Canada, Commercial Chemicals Evaluation Branch, sponsored a workshop to develop ecological tier testing schemes for genetically engineered microorganisms. There was general agreement that the “potential ecological impacts of microorganisms released into the environment have not been well characterized.” Ecological effects endpoints identified as potential areas of evaluation included:

1. Effects on Primary production
2. Effects on cycling of limiting nutrients
3. Effects on community structure and diversity
4. Effects on community function
5. Trophic level changes / effects on grazers
6. Effects on sensitive species

Based on the discussion, the group of participants developed a tier testing scheme (0-3) to evaluate genetically engineered microorganisms in closed, semi-closed, and open applications. Tier 0 contains preliminary information, taxonomic identification, proposed use, and site characterization. Tier 1 contains initial exposure and hazard assessment components such as persistence, dispersal, pathogenicity, toxicity, and basic ecological effects. Tier 2 addresses additional questions about exposure and hazard from Tier 1 and contained longer term and more complex ecological effects testing. And Tier 3 contains open or limited field tests in the selected environment.

The 1994 workshop was followed in 1996 by a workshop on testing methods. The goals of this workshop were to identify protocols to test GMOs prior to release and to develop standard operating procedures to guide data collection. Some resulting major questions included:

1. Is this organism temporary or will it persist in the environment?
2. Will there be a recovery in the organism's population after an unknown amount of time?

These questions were asked nearly 20 years ago. What is still relevant? What additional information is still needed for us to be comfortable releasing microorganisms into the environment?

More recently, the Wilson Center has hosted workshops on both rE. Coli and cyanobacteria. Some of the major ideas raised during these discussions included considerations of fate and transport of DNA, modeling gene transfer, and understanding the respective time lags involved.

Findings from the previous workshops highlight the still-existing broad areas of uncertainty in the field. Some of the topics discussed more recently are very similar to topics discussed in the 1990s. Overall, stakeholders must be aware that history will repeat itself if it is not sufficiently studied. It is important to now come up with a clear set of issues that need to be worked out in order to properly test these organisms.

The system must be thought of in totality as opposed to simply comparing organisms' interactions with one another. There must be objectives and ecological endpoints that are considered. Overall, how can these organisms be tested for the long term? How many "cycles" are needed?

Arizona State University (ASU) is developing and validating methods that deal with these questions. At least one annual cycle is needed, possibly more. However, these are all context-dependent and if there are more unanswered questions that arise, more data will be needed.

What about the by-products of these facilities? One industry representative noted that at present, all by-products are returned to the productions process. Another stakeholder noted that there is also an algae interagency working group that is currently looking at the use of algae in animal feeds.

Summary of Session V: Identification of Ecological Endpoints to be Assessed

This discussion considered immediate and long-term data needs for algae synthetically engineered for biofuels production.

Broad receiving environments: terrestrial, freshwater, marine

- Relevant questions: What are the options? What do we expect to be coming? Where will facilities be located and what is the effect of location?
- There are applications in bags, raceways lined and unlined in the desert, and potential applications near waterways.
- For companies based in deserts, which a number are, early applications will potentially be in the desert.
- There is a need to take things stepwise to gather information to fill in blanks before proceeding to environments that are more difficult from a knowledge or risk angle; i.e., not jumping right into deploying by the ocean.

Identification of locations and environments that participants viewed as "good" for deploying the technologies

- An ideal location may have: water, abundant sunlight, CO_2 from power plants or other sources, nutrients, land.
    - Resource availability is important: Lands must meet certain criteria, such as availability of saline water in sustainable supply. Availability of CO_2 is important: CO_2 is one of the most expensive inputs today. Nutrients are also important (N, P, K).
    - Co-location may be an option. Nutrients and other inputs can be sourced from farming.
    - Pacific Northwest National Lab (PNNL)/National Renewable Energy Lab (NREL)/Argonne have techno economic analysis work on availability of resources, standards, fresh and salt water, specifically for algae. It may provide a framework for making good decisions.
    - An industry member noted that his company is focused on inland solutions, and that coastal applications would involve very different technologies.
- One good option is land that cannot be used because of former pollution:
    - People at DOE have tried to develop those lands for other renewable energy (e.g. wind, solar).
    - What are issues there compared with a pristine (e.g. desert) site? The polluted site is already polluted.
    - There are also lands that have previously had agricultural activity but are now degraded to the point of no longer being useable for such purposes

    - (e.g. salted too much). The environmental effects have already happened. The USDA does not have a program for treating such land.
    - Use of polluted/degraded lands would not satisfy environmental groups, due to concerns about the organisms getting out. The organism would need to be shown to be safe.
- Identification of other examples of sites where algae facilities could be beneficial or less harmful:
    - In Florida, citrus groves are dying near a CO_2 source. There are thousands of unused acres.
    - Some participants noted that water is being removed from the Salton Sea and argued that as the water recedes, dried material will blow to populations and become a medical liability, and thus an algae pond on top of it would be a benefit to mitigate environmental disaster. Some participants representing environmental concerns argued otherwise, particularly if there are endangered species nearby (e.g., pup fish in nearby hot springs).

Identification of locations and environments that participants viewed as "bad" for deploying the technologies, and limitations

- Identification of types of locations to be avoided:
    - Algae facilities should avoid collocating with another algae facility, e.g. cultivating algae for fuel production near algae for food or dietary supplement uses. Even if the genes do not integrate, if the food facility tests its pond and finds algae for fuel production, it loses all of its sales for a period of time, until it can prove that there is no more contamination. This is a tort issue and a location question.
    - A "bad" location is near where there is an endangered species, one that is water dependent, or worse, algae dependent. Environmental groups would want to know how the producer would verify no harm done to the endangered species.
- Identification of other examples of poor sites:
    - A participant identified an "unqualified entrepreneur's garage" as a worst possible location.
    - A participant identified Minnesota, as an area with many lakes and lower light, as a worst possible location.
- Identification of location limitations:
    - Land use rules in general act as limitations: Where can one put a hundred thousand square foot facility?
    - Open or semi-open ponds today are limited to areas without extreme cold. Look at where everyone is locating facilities: Arizona, Florida, warm areas. They have to be able to run year round to make it economically feasible. If can make it yield twice as much, only half as much land is needed. NREL and PNNL are conducting a study on locations, overlaying many local conditions.

    - There may be international treaty issues, for example: There are ducks regulated by treaty with Canada and Mexico that eat algae. They may want to know whether those ponds are replacing ponds that ducks normally eat.

Risks of engineered/synthetic biology vs. natural strains

- If natural strains used for biofuels production are located in areas identified as “problematic,” are the natural and engineered strains so different that alternative conclusions on siting would be reached? Are there things about engineered algae that have stakeholders worried in comparison with natural algae?
    - Natural species are only regulated through land use laws. Now added regulatory/safety constraints are being introduced due to synthetically engineered traits.
    - The natural ones have been reproducing and living there for thousands of years.
- Most current-generation algal biofuel strains are generated using directed evolution. Are there concerns about directed evolution strains vs. natural strains? Should directed evolution be added to methods of creation to be concerned about?
    - Some participants argued that yes it should be added; it is not clear how concerned to be about it, but concerns exist. There are concerns about the effects of speeding up the rate of evolution.
    - A participant argued that directed evolution means speeding up one organism in relation to others, so directed evolution could result in a mismatch of competition, by creating an ecological imbalance in how systems react to change.
    - However, another participant argued that directed evolution is a means of finding “the needle in the proverbial haystack”: Making the organism through directed evolution is just easier than finding it naturally.
- A participant argued that the idea that humans can make an organism that is better than what nature can produce is false. There are millions of organisms all competing with one another. Humans are the only organism that has outcompeted everything. Whatever humans could do is to benefit people, not provide the organism with a competitive advantage. There is no organism that can outcompete everything. The ones that produce damage are already there, producing harmful algal blooms. One wants to avoid making more of those, and must consider whether an algae that had never been in an environment before will cause perturbances when introduced.

Invasiveness, escape, and effects of natural dispersal on risk

- Some participants argued that algae have already been transported everywhere (e.g. by wind) so there is no need for concern about invasion.
    - There are examples of where algae are not as cosmopolitan, but as a rule they tend to be well distributed.

    - Some participants noted the uniqueness of the case of using a very localized/endemic algal strain for such applications. Does the potential uniqueness of localized algal strains factor into risk characterizations? Regulators may base a history of being able to use an organism safely on the scientific literature. However, a company's one submitted paper may be the only scientific literature on that organism. It would have to pass other tests: Could it be grown in a suitable environment, could the molecules of interest be made with it? A localized organism is probably too unique; it could not be grown in different places. Industry would want to find something easier to use. But lacking information, companies would need to study about it. This relates to discussions of the idea that the first-in-class gets more scrutiny.
    - Rock snot is an example of a plant exhibiting invasiveness in a new environment. There are more examples of that. A participant noted that rock snot is a naturally occurring species, unlike the engineered algae being discussed. Some participants argued that animals may be less widely distributed, so concern about animal invasives such as cane toads could be different from concerns about widespread organisms such as algal strains.
- Some participants argued that invasion is a concern.
    - There are whole algal genera that are common in the southern hemisphere but not found here. Common general ones are everywhere, but more specifically evolved ones are not as widespread. There are some very restricted environments and microorganisms that have not spread everywhere.
    - Even if a population of algae is found everywhere, does having a large number of a species in one area change the risk? If there are hundreds of thousands of times more of those algae in ponds than in the surrounding ecosystem, does that affect its ability to survive and affect the ecosystem? One participant argued that it may affect dispersal events, but that if the organism is local, if it could be in that environment then it would be.
    - Dispersal events are important. Invasion rate matters regarding turnover of species populations. That problem is occurring now: As climate change occurs, it is opening new environments to different organisms. It is important to think about where the evolutionary constraints are.
- Florida has adopted an Invasiveness Index, first developed by Australia. Is this for consideration of synthetic organisms?
- An algae industry participant noted that they were using advanced technologies to sample and study organisms that cannot be cultured, and are culturing organisms better; this is where further development of tools and instrumentation would be important to understand existing biodiversity.
- Some participants compared synthetic biology to historical domestication of crops and argued that domestication is not the process of making organisms more fit for the environment. Others argued that there are many examples of organisms becoming more environmentally fit and invasive due to domestication, so an analogy with domestication does not suggest that one should not worry about domestic algae.

- These questions will look different to the environmental community.
- Need to consider fitness, genetic stability, and gene transfer.

Additional location considerations

- Participants mentioned dispersed production facilities, which people put where they wish.
    - The distributed approach includes only environment types already discussed for centralized facilities (cropland, urban, deserts, freshwater, saltwater).
    - Distributed systems still have to be big and near adequate infrastructure. They will have oil to be refined, which will need to be piped/trucked, so such facilities will still be located near refineries, with the ability to move oil/fuel to where it can be refined/distributed. They will need to be near existing infrastructure and labor pools to avoid creating a shantytown in the desert.
- The discussion has dealt with variables one at a time (siting, etc.), but the factors interact; for example, an organism's intended use may affect the ideal location.
    - For example, is making lipids dependent on availability of customers? Should they be grown in urban environments to use available CO_2? How should we separate the variables?
    - Techno economic approach to decisions. Social, labor impacts. It comes down to techno economic. It is all economics, particularly in energy, which is a low-margin commodity.
- A participant argued that the organism should be contained. If it is not contained, the environmental parties would oppose it because of concerns about its getting out, unless the organism was shown to be safe through better testing that was public and not deemed Confidential Business Information (CBI).
    - A participant argued that there is a need for a better tool than TSCA for regulating these things.
    - A participant stated that testing is not CBI. All testing must be publicly disclosed. EPA has a mechanism for looking at that and making it open under Freedom of Information Act (FOIA) requests.
    - Some participants stated that all participants could agree that they were looking for a set of data that shows release is safe.
- Facilities are now located in Hawaii, China, and other places. They are in places participants have noted as not being ideal, but some are in the "best" locations as well.
- What if more people were doing this? There was a problem with chemicals because of synergistic effects. No work was done on interaction effects. What if there were lots of GM or synthetic biology organisms out there? How would that affect things and how would it be regulated? Other than foods, assume they are making chemicals or other things that are in EPA's regulatory space.
    - Some argued that such a situation exists for bacteria now, and that an analogy between bacteria and algae is valid because algae from different facilities do not go and mate/cross with each other.

- Some argued that the existing bacteria are in reactors. The potential for interaction there are much less than when everyone is making their own "boutique bugs." What if there were many Bioprocess algae -type locations, in which one bug was making one product, and another was making another? Would it make a difference? It could be organized as an eco-industrial park, but it could just be co-located.
- Some argued that industry is not going to be making new molecules that have never been seen before, so that type of combinatorial interaction would not take place.

- A business plan could be to take CO_2 from a highly polluting ethanol facility to make fuels, not using GM plants. Cannot answer the question on what that would mean for possible synergistic effects. A participant stated that what industry is coming up with now is the most efficient way of making an existing industrial product (for now).
- One could make a species to take advantage of a niche, but more likely as the niche changes, new strains/characteristics develop, e.g. the tar sands produced a new environment, organisms have evolved to take advantage of it. There are many considerations in siting that have nothing to do with these considerations. The cases of dispersal and of evolution to take advantage of a niche are difficult because there is still a lack of understanding of what those endpoints are and how they can be addressed.
- It is important to be sensitive to variations in the release environment.
- Siting issues become connected with facility design issues. EPA does consider design of facility in approval.

Reductions in organism fitness

Assuming the "ideal organism" traits discussed in Session II, participants discussed claims of decreased fitness. What measures or tests would they look for to identify reliable evidence of reduced fitness?

- Experimental data should be obtained to determine how an organism survives in the environment and how it competes against local organisms. For example, one might put the organism into ponds and see how it survives, or put algae into samples from nearby water and see how they do. It is important to include markers to be able to track the organism.
- One may also create microcosms and test fitness. These are not expensive, and are doable.
- Genetic stability: What if the organism loses a trait that had lessened its fitness?
  - Dependent on method. In vitro is point mutation, so it will drift immediately back when the selection pressure is relieved. If one does a stable transfer, then the genes are more stable. So methods affect stability. Process becomes important.
  - Reversion data exist, but are just never seen as worth publishing?

  - With drug resistance experiments, as long as the selective pressure remains, only the ones that survive continue. GFP (green fluorescent protein) has existed forever.
  - It would be valuable to test the degree to which traits shed. For example, tests conducted with plants and stack traits, but used to test one-by-one. Tests are only as valuable as the settings acknowledged. How to design tests for this? To check stability of attributes? How to test for hazard and likelihood?
- Bloom algae are rapid reproducers, strong against grazers; they accumulate, and then form harmful algal blooms.
- A participant suggested that lessons may be learned from earlier soil microbiology: When rhizobia were first being introduced, everyone said that they would die out rapidly in the soil. Later, someone planted sensitive legumes in the area where the rhizobia could no longer be found, and found them a decade after they had disappeared. Similar studies were conducted in Oregon and there too, it was found that organisms can be undetectable until the perfect conditions arise.
  - For algae, do the organisms die or do they just become undetectable?
  - A participant stated that algae do form spores. Large proportions of algae are rare and become abundant only once in a while.
  - How would the knowledge gained about rhizobia have changed the risk assessment decision in that case? A participant argued that the situation with rhizobia had been anticipated, but that the question at the time was what the hazard consequences are, and stated that risk assessments for algae would be approached in the same way.

Considerations of horizontal gene transfer (HGT)

- To put this in context, the question has been do they survive or not, but the real question needs to include a time variable. How many generations? 50? 100? Some participants argued that it should be assumed that the organism will survive long enough to transfer. Much more is known for cyanobacteria, but data are still needed for eukaryotic gene transfer.
- There has been some indication of rates and environment, but it would be good to know information specific to eukaryotic algae. How do they work, and are there appropriate recipients out there? This is an information gap that requires more data.
  - Dick Sayre’s paper suggests there may be more transfer than expected, but there has been much less sequencing. Data are expected to come out within the next year or so; the extent of transfer has not been seen yet, just because not enough sequencing has been done yet.
  - One company showed no HGT against local organisms. Is that useful? Should companies be conducting such studies?
  - One reference is looking at gene transfer between alga and viruses. Is this really rare, or does it require further study?
- What is the significance of a high probability of naturally occurring transfer? How much has changed, how unique is the gene? Not all genes require further study;

how should it be determined which do? Transfer is not always hazardous, so the consequences of transfer need to be known. How much transfer is too much? Some participants argued that the limit cannot be kept to zero. Some organisms will take anything (e.g., rotifers take on everything they eat), some will not. There is a lack of sufficient data.

Who should be performing the needed studies and how should data be made available?

- Models for who funds and who performs studies:
    - Some participants argued that it is the role of government to perform these basic science studies. There are many experts in these areas, so the government should spend some of its basic research dollars on this. Private companies doing this research are always seen in a tainted light. The companies want to know what tests to do and they do not want to wait.
    - Petroleum Environmental Research Forum (PERF) model: a number of petroleum companies have the same questions, so they pool money, DOE matches (-ish), then national labs do the research and publish the answers. Some proprietary data are kept confidential, but at least the answers become available, and everyone benefits.
    - National Institutes of Health (NIH) model: “we fund, you publish, we post findings and data.” Some participants stated that this is the type of model they would like to see.
- Many questions are being explored, but by proprietary entities. Who owns the data and how public are they? If they are not presently available, how can they be made public?
    - A participant FOIA’d DOE, and some labs gave everything while others kept nearly everything confidential. The participant argued that the agency needs a clear policy. The Department of Defense, on the other hand, gave lots of information (the Defense Advanced Research Projects Agency (DARPA), however, is entirely secret).
    - CBI and protecting firms, vs. protecting the public interest; is this a “collective interest free-rider problem”? Without scrutiny, is trust lost?
    - A biofuels industry member asked other participants what information they would like to see. Participants responded that they would like to see data from studies, and to know where variables were modified and how, etc.
    - In pharmaceuticals, the European Union has shifted to public dossiers.
    - Floating around Congress are revisions to TSCA that put more of a burden on the submitter for why their data should be confidential. How that will play out is unknown.

Summary of Session VI: Methodology and Protocols

Participants were asked to discuss what kind of advances in instrumentation and measurement methods might be required to evaluate the potential or actual environmental impact of GM algae. The discussion focused on four areas: mesocosm experiments, ways of measuring genetic stability and gene transfer, the need for data on a "base set" of organisms, and the need for modeling, validation, and reproduction of all experiments. Many participants repeatedly emphasized the need for cooperation, data sharing, and replication/parallelization of experiments between industrial, government, and academic labs.

1. **Technical Aspects of Mesocosm Experiments**
   The first question is the choice of which environments to simulate. Sterile water is not a representative environment, although it could be used in a control experiment. Participants emphasized the need to choose a representative sample of the environments that an escaping organism is most likely to encounter near the site of cultivation. Although earlier discussions of land-use and siting focused on polluted and degraded land, one participant described a practice of choosing to sample healthy local environments, rather than those already perturbed by human activity.

   Self-contained mesocosms must be technically sophisticated enough to accurately simulate the natural environment. They must include diurnal variation of conditions, water flow, and replenishment of resources, and should not allow unnatural chemical buildup. Any mesocosm experiment should be run long enough to be meaningful (possibly as long as multiple years). The appropriate duration of an experiment will depend on the effect being studied.

   There is great desire for validated, standardized mesocosms; one participant said, "I'll buy a dozen of those reactors if it'll help me standardize." There is an existing field of mesocosm studies, but its methods and apparatuses may need to be modified to fit the needs of algae researchers. ORNL has developed flow-through systems that are germane to mesocosm development. Existing and proposed test bed facilities should be built with an eye to scaling up from test-tube to microcosm to mesocosm experiments. JBEI reportedly has a full range of production models (labs, greenhouses, open ponds).

2. **Measuring Genetic Stability and Gene Transfer**
   Several participants stated that as the study of horizontal gene transfer (HGT) has advanced, researchers have found that far more HGT is taking place than was previously thought. This is due partly, but not wholly, to improvements in sequencing and metagenomics techniques. There are good data on transfer between bacteria, but not for eukaryotes. It is now known that gene transfer can

cross kingdoms. Thus, there is a great need to narrow the search space. It is infeasible to do pairwise HGT tests of all organisms in the environment. With current methods, researchers can establish whether or not HGT took place, but there is currently no way to predict whether one organism will transfer genes to another. Given the large amount of HGT already occurring in nature, experimenters must also consider exactly which genes are being transferred. If a particular HGT event would have occurred in nature even without human intervention, then that event may not be of concern if it occurs due to a GMO. Presumably, novel/engineered genes are of greater concern than naturally occurring ones, but not all novel genes are equally problematic.

The stability of genes introduced into an organism is often tested by growing the organism for several generations without selection pressure for those genes, and seeing if the genes are retained. The stability of an introduced gene depends strongly on the method by which it was introduced. Plasmids are lost relatively rapidly, while genes inserted by chromosomal integration are far more stable; one participant cited stability of up to "decades in the lab." Chromosomally integrated genes are not perfectly stable forever, but it would be difficult to detect reversion events due to their rarity. The current EPA regulatory approach does take into account the method of insertion of genes, and treats plasmids differently than genes integrated into the chromosome.

Several participants stated that the group attending the current workshop did not have all the necessary expertise to have a comprehensive discussion on the topic of genetic stability. They suggested holding another workshop with evolutionary biologists, specifically to address this question, and named some potential attendees.

3. A "Base Set" of Organisms

The published literature on algae is very thin. Most potential commercial biofuel organisms have not been thoroughly studied in a way that addresses the concerns brought up at this workshop. A few food-relevant algae, such as Spirulina, have been so studied, but they may not be suitable analogues for biofuel-producing algae. One participant called attention to ORNL's extensive study of vascular plants for bioenergy applications in which the group examined hundreds of species in cooperation with USDA and others. The participant suggested that the algae community could do similarly wide-ranging studies.

Many participants agreed that it would be useful to conduct comprehensive studies on a "base set" of algae species, evaluating their safety in specific environments. The establishment of this "base set" could go hand-in-hand with the earlier suggestion of establishing a "bad bug list" of organisms to avoid using. Members of the base set could be considered as analogues when evaluating a novel GMO similar to one in the base set. The eventual goal would be to make the leap from specific conclusions like "organism X causes harmful outcome Y in environment Z," to general statements like "engineered feature A will likely present problems, and feature B will not."

One participant proposed a basic experiment: take a wild-type strain, modify it to include green fluorescent protein (GFP) or to be traceable in some other way, and release it into the environment. (This strain would be relatively safe to release, but would still require a TERA.) Experiments like this one would be part of the "base set" species evaluations, but this particular experiment would only be a starting point.

4. **Experimental Design / Need for Modeling, Validation, and Reproducibility**
Measurement criteria should be defined ahead of time, and include consideration of factors like how long the experiment must be run. In every case, the experimenters must decide: what precisely is the effect of interest, and how long must they wait for it to occur? For example, the concept of evolutionary fitness includes much more than "yes, the organism survived" or "no, it did not survive." It also includes the relative growth/success of different species, and their prowess at nutrient utilization.

Mesocosm experiments need to be extensively replicated, and parallelized between different laboratories – academic, industrial, and governmental. This point was made repeatedly by multiple participants, one of whom told an anecdote about losing an entire summer's worth of data from open mesocosms in Lake Erie because a great blue heron defecated in one replicate and not in another.

Once mathematical models have reached an adequate level of sophistication, they could help ease the burden of replication. Models and replicable analyses would also help experimenters plan their measurements, controls, and duration of testing. Besides the inherent usefulness of models, as one participant noted, the algae community needs to reach the level of sophistication needed to produce relevant models in order for anyone to have confidence in the results of their experiments, whether virtual or physical.

Both models and physical experiments must always be validated against field observations. Although comparison between a mesocosm and a natural stream is the first step of validation, the community must decide on specific validation criteria beyond "it seems similar to nature." One proposed experiment was to see if natural environments and self-contained mesocosms react similarly to the introduction of a "somewhat exotic," but non-GM, organism. Following this experiment, a GMO could be introduced to the mesocosm, and its effects extrapolated to nature.

Participants also mentioned the possibility of genetic manipulation having unintended side effects. Metabolic networks are highly complex and redundant, and a modification to one gene may change the behavior of many others, and/or have biotrophic effects. In general, EPA expects applicants to know what changes they have made in a GMO relative to the parent organism, but these side effects present a difficulty. Applicants might address this difficulty with a comprehensive set of gene microarrays, protein, RNA, and metabolite measurements. TSCA requires applicants to submit all data relevant to health or environmental consequences, and does not set a standard list of testing procedures. This allows EPA to be flexible, in case an applicant presents a highly unusual organism.

Finally, one participant asked whether the current discussion was taking place within an appropriate conceptual framework. There was some agreement that a good framework is needed for any such discussion in order distinguish which questions are most important and why. Several potentially applicable frameworks already exist:

- The TSCA statute and regulations themselves constitute a framework, but it may not be the most useful one and should certainly not be the only one used.
- The EPA is using a comprehensive environmental assessment framework for the evaluation of nanomaterials, and a prior Wilson Center workshop on synthetic biology used this same framework to guide the conversation on possible hazards from cyanobacteria.
- There also exist tools specific to the field of risk assessment, such as fault trees, which could provide guidance.

Summary of Session VII: Wrap-Up

To close the workshop, participants were asked: "If you had $100 million to spend on research to address the questions raised today, how would you allocate it?" Answers focused on three areas: meta and organizational efforts; basic algal studies; and specific experiments to do or technologies to develop.

Meta / Organizational Goals

- Carefully planning the studies to be done, in accordance with an agreed-upon set of priorities.
- Rigorously defining concepts like "fitness"—which has so far been used in a vague way—in terms of measurements, baseline conditions, environmental contexts, safety targets, and best management practices to reach those targets.
- Performing literature reviews, and examining the work of other countries on other types of GMOs such as fish.
- Examining algae-farming endeavors from the perspective of social/economic sustainability, as well as environmental sustainability.
- Supporting education and outreach to ensure that the algal research community serves the public as well as the expert audience.
- Wrapping up results from across the range of future studies into a coherent body of knowledge.

Basic Algal Studies

- Studying general algal biology to increase the community's knowledge and tool-set, ultimately endeavoring to approach the level of characterization currently available for other industrially useful organisms.
- Establishing environmental reference data on natural algal communities so that in the future it will be apparent if a change has taken place. This could include taking baseline data on natural algae, their effect on the environment, and their natural lipid production. One participant considered whether baseline environmental monitoring should be a condition of DOE awards or other grants for biofuels development.
- Establishing a "base set" of useful organisms—and accompanying comprehensive characterizations—as well as a list of organisms to avoid using.
- Studying what role the "base set" organisms play in natural microbial communities in order to know what one might expect to see, or what one should plan to measure, in a mesocosm experiment.
- Studying the ability of harmful algal blooms (HABs) to produce neurotoxins, and the ability of engineered organisms to do the same; sequencing genomes to look for toxin-producing or allergenic gene products.

Specific Experiments or Developments

- Developing replicable and realistic mesocosm apparatuses and protocols for use.
- Modeling relevant phenomena that are already well understood, such as the airflow over a typical open-pond facility.
- Taking an engineered organism, knocking out the inserted genes one at a time,

and measuring how well it competes in a natural population to simulate the effect of genetic instability.

- Studying GMOs in the context of “non-natural” environments, such as depleted farmland, and the combined effect of farming/depletion and algae culture over time.
- Expanding previous work on probabilities of various adverse outcomes to incorporate the full range of spatial and temporal scales. (See: Martin Alexander, “Ecological consequences: reducing the uncertainties”. Issues in Science and Technology 1:57-67 (1985).)