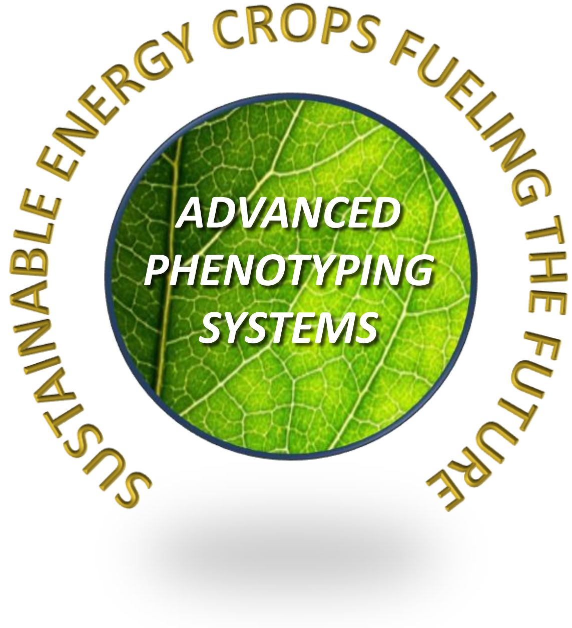
Accelerating Bioenergy Crop Development

Next Generation Phenotyping: Biology x Engineering x Computer Science



Workshop Conclusions

arpa·e
CHANGING WHAT'S POSSIBLE

Bioenergy Crop Breeding

- Poplar and sorghum are two promising bioenergy crops with good genomic resources available that could be used as test crops for a phenotyping program. It would be difficult to implement this program if existing breeding populations were not used.
- There is a clear correlation from the breeder's equation for genetic gain between the phenotypic variation (which can be represented by the number of varieties) assessed and yield improvement, so increasing the phenotyping throughput should increase genetic gain.
- It's not realistic to improve yield or other bioenergy traits in a three year program also aiming to develop field phenotyping technologies, but there may be more value in developing tools to predict yield gains. The data sets currently available are not comprehensive enough to do so.
- A goal of the program should be to correlate yield improvement and other phenotypes of interest with genetics (e.g. linking field phenotypes with SNPs). There are sufficient bioinformatic tools available for genetic analysis, but the current cost of $10 per plant for genotyping is too expensive for large breeding efforts.
- An automated sensory platform would significantly increase the throughput of screening lines in the field and the accuracy as compared to the existing approach (e.g. undergraduate students).

Bioenergy Crop Breeding

- Breeders should provide engineers with specific traits of interest in the field and the necessary specifications of the measurements. The engineering community felt confident that the necessary equipment exists and that sensor platforms could be readily built once the requirements were known.
- A number of traits that were of interest to measure in the field were identified along with some potential measurements. While a sensor platform capable of collecting data on a crop in the field on a daily basis would be ideal, there would not be much value in measuring some traits such as plant composition daily.

Trait	Potential phenotypic measurements
Carbon fixation	
Starch production	Delta during the day
Root architecture/ composition	Need radar or similar, surrogates such as free sugars in sorghum or cotton are correlated to total root length (IED backscatter)
Water use efficiency	Thermal imaging
Drought resilience	Measure of wiltiness, yellowing
Osmotic potential	
Crown architecture assessment	Sylleptic shoot production, correlated with productivity
Biomass	3-D imaging, volume reconstruction, LAI

Sensor platforms

- A wide variety of sensors have been demonstrated for non-agricultural applications (e.g. aerial remote sensing) that could probably be adapted to measure the desired crop phenotypes, though cost and data handling may be issues.
- Recording environmental data (temperature, humidity, light) will be key to assessing phenotypic responses, but these sensors can be external to the platform as long as the data can be integrated.
- One phenotype of interest that is currently not measurable in the field is root growth/ morphology, and may possibly require a dedicated effort to develop novel sensors.
- As techniques such as hyperspectral imaging and lidar will generate terabytes of raw data per data, some data reduction capability may be important, either in algorithm development to compress data or in identifying only specific spectral bands to acquire. However, in 3-5 years big data handling will not likely be an issue.
- Breeders want data on individual plants. Machine vision algorithms have been demonstrated to be capable of distinguishing individual plants in a population under certain circumstances but need to be improved. To get that resolution, aerial sensing will be limited to UAV platforms that do not fly too high above the field.
- Handling lighting conditions will be a challenge, both in terms of affecting passive sensors on the platforms and for correcting physiological responses in plants due to light intensity.

Sensor platforms

- Workshop participants discussed potential phenotyping systems that met the needs of breeders and the general approach was to utilize two sensor platforms. This would allow the system to rapidly assess certain phenotypes on a daily basis (e.g. plant height, water use efficiency) and more complex phenotypes on a weekly basis (e.g. plant composition).
- The high throughput platform would be aerial (UAV or gantry) based. The other would be ground based and travel through the field plots, allowing it to image the plants horizontally. The payload of a UAV will limit the suite of sensors that can be deployed, whereas a gantry system will be limited by the infrastructure necessary to cover large (> 5 acres) fields.
- HSI, lidar, thermal, and RGB cameras would be capable of obtaining all of the necessary imaging data for leaf area index, plant height and volume, and photosynthetic activity.
- Additional sensors could include ground penetrating radar for root structure, ultrasonic transducers, and micrometeorological sensors.
- The benefits of a fully automated platform are not necessarily outweighed by the additional cost and design complexity involved in removing a human operator.
- A comprehensive sensor platform could be built today with existing technology, and would be valuable to determine the specific sensors that would be essential for high throughput phenotyping of crops in the field. As technologies will improve rapidly over the next few years the program should not worry about the cost of the prototype, but should use a metric of cost/acre.

Data handling and analytics

- Today, producing multi-terabyte quantities of HSI and other raw image data each day from a field plot will require swapping hard drives from the sensor platform every day to process off-field. For the purposes of this program algorithms to reduce data size are important, but may not be when these systems expect to be commercially deployed due to technical advances in data processing.
- Performers will need to develop data standards for plant phenotyping to handle and make meaningful conclusions from the data generated. ARPA-E can provide a valuable service by helping coordinate the development of data standards for the phenotyping community.
- Open source tools are the norm in the bioinformatics field, and is desired for the phenotyping data analyses tools described in the workshop. Workshop participants envisioned a data and algorithmic platform whose backbone is open source, but modular enough to allow for performers to upload "black box" algorithms which are either open or closed.
- Kbase and iPlant were identified as existing data portals that could be utilized to handle field phenotyping data and build tools on, but currently not optimized for these purposes.
- There are currently good bioinformatic tools for linking genotypes with phenotypes, but would not be capable of handling the amount and novel types of data that would be generated by a phenotyping program.

Key take-away messages

- A field phenotyping program would be very timely because (1) genomic information for energy crops and individual lines is readily available due to advances in sequencing technologies; (2) there is no concerted effort in the public sector for developing high throughput phenotyping tools for energy crops, and tools being developed by the private sector are focused on traditional row crops and not available to public breeders; (3) advances in HSI sensors and UAVs can allow rapid assessments of a range of plant phenotypes; and (4) without a focused phenotyping program it seems unlikely that the necessary yield improvements for bioenergy crops will take place to meet the targets of the Department of Energy's Billion Ton Study.
- Predicting phenotypes from genotypes is an active area of research, but the field has been lacking in comprehensive enough data sets from field plants to build reliable models. A phenotyping program could generate the necessary data to enable this.
- Analytic tools developed under this program could be the biggest benefit. Private industry is already developing phenotyping hardware and may leapfrog anything developed under this program, but cannot establish standard data platforms and analytics that would advance the field as a whole.
- The near term commercial value of a phenotyping program would really be as a research tool to specify what sensors and capabilities are key to include on a products for the existing agriculture market (i.e. corn).