Science & Technology Opportunities at EMSL

David W. Koppenaal
Chief Technology Officer

JGI Genomics Technologies Workshop - March 24, 2015

Scientific Innovation Through Integration  ■  www.emsl.pnl.gov
Our vision is to pioneer discoveries and mobilize the scientific community to provide the molecular science foundations for BER research priorities and our nation’s critical biological, environmental and energy challenges.

**SCIENCE THEMES**

- **Biosystem Dynamics & Design**
- **Atmospheric Aerosol Systems**
- **Terrestrial & Subsurface Ecosystems**
- **Energy Materials & Processes**

**CAPABILITY AREAS**

- **Mass Spectrometry**
- **NMR & EPR**
- **Molecular Science Computing**
- **Microscopy**
- **Spectroscopy & Diffraction**
- **Deposition & Microfabrication**
- **Subsurface Flow & Transport**
- **Cell Isolation & Systems Analysis**
Today’s topics

- EMSL-JGI Joint User Call – status and overview
- Science highlights
- New capabilities
EMSL-JGI Joint Science Calls

- EMSL-JGI Joint Science Call established in 2014 to promote collaborative use
- Of 58 proposals submitted, 20 accepted in 2014-2015
- Focus on bioenergy, terrestrial carbon, microbial dynamics

LIVE Google+ Hangout on March 4

FY2016 Proposals
JGI-EMSL Project: Isolate cellulose-assimilating cells from soil microbial communities

**Goal:** Identify key organisms involved in soil carbon cycling (e.g. cellulose) and determine their metabolic functions by single cell- and meta-genomic/transcriptomic analyses.

Blue – fluorescent DNA probe
Red – fluorescent cellulose
Magenta – co-localized blue and red

Influx Flow Cytometer

Kirsten Hofmockel, Iowa State University
Deciphering the composition of fungal cellulosomes

- Fungal cellulosomes from *Piromyces sp finn* are readily precipitated from anaerobic cultures
- Cellulosome complexes are large, and have 10+ putative subunits

What is the structure/composition of these fungal cellulosomes?

Fungal Cellulosome Architecture?

Fungal Cellulosome Components

All proteins were predicted by the transcriptome and contain dockerin domains

Michelle O’Malley, Charles Haitjema, UCSB
Proteomic studies reveal a putative scaffold in gut fungi

- ~600 kDa protein with interspersed “repeat” motifs
- Motif is also found in GH48 and GH6
- Downregulated on glucose
- Conserved across all gut fungal genera we have isolated and characterized
- No homology to other sequences in the NCBI database

Michelle O’Malley, Charles Haitjema, UCSB
Aaron Wright, Sam Purvine, PNNL
EMSL Research Campaign: Cyanobacterial Synthetic Biology

- **Synechococcus elongatus** - UTEX 2973
- Unicellular, oxygenic photosynthetic microbe
- Most rapid growth recorded for a cyanobacterium to date
- Ideal characteristics for synthetic biology chassis

Himadri Pakrasi, WUSTL
Limited number of genomic changes confers high growth rate for UTEX 2973 compared to PCC 7942

- Closest genomic relative, *Synechococcus elongatus* sp. PCC 7942 was used as comparator strain
- 99.8% sequence identity between UTEX 2973 and PCC 7942
- Yet markedly faster growth...how/why?

99.8% sequence identity between UTEX 2973 and PCC 7942

Visual culture densities across 16 hrs; Growth curves of strains at optimum conditions.

EMSL proteomic capabilities helping to explain

- Proteomics validated specific SNPs, and confirmed key protein abundances
  - 1754 proteins; 66% coverage
  - Peptide sequence validation of 5 detected UTEX 2973 SNPs
  - Quantitative validation of 5 of 6 unannotated genes expressed in PCC 7942 but deleted from UTEX 2973

- Proteomics a vital tool in linking protein expression with physiological observations

- Top-down proteomics can help to decipher regulatory mechanisms behind fast growth

Yu, J. et al. 2015
Probing single-cell fungi growth with microfluidic cell culture arrays

Hyphal growth channel: 2.5 mm X 10 µm X 10 µm

Cell trapping channel: 10 µm X 7 µm X 2 µm
Single-conidium trapping and compartmentalized hyphal growth
Hyphal growth observation - *Neurospora crassa*

2.5 mm long, 10 µm wide channels

In both strains, single hyphae exhibit distinct germination time and growth rate behavior (cellular heterogeneity).
Long-term observation of xylose transporter-GFP expression in response to carbon source change
HRMAC (21T FTICR MS) – A new EMSL capability
- Ultra-high resolving power and mass accuracy will provide near unequivocal biomolecular species identification
- Improved ability detect & monitor intact level protein transformations
- Increased sensitivity will provide higher imaging resolution (< 1 µ)

First Spectra!

Peptide (MRFA)
HRMAC will address environmental and biological complexity at the molecular level.
Computation modeling is an integral part of many research projects.

Molecular Science Computing Facility - integrated production computing environment

- Chinook supercomputer - 18,480 processor cores (163 teraflops peak performance)
- Data archive - 7.5 petabytes of storage capacity
- High performance software development – NWChem

Unique environment for integration of experiment and simulation to address complex molecular science problems.
Starting guess came from 1RUV (RNaseA - uridinium vanadate complex)
While protein backbone changed little, there are significant changes in the active site
Calculation of Redox Potentials

- 1QJD in rectangular box (81x97x103 Å) of SPC/E water ~ 77,000 particles
- 109 atoms QM region
- DFT/B3LYP Ahlrichs pVDZ

Redox Potential of heme groups (kcal/mol)

- Important for bioremediation efforts (metal-ion reduction)
- Four iron heme groups facilitate electron transfer
EMSL combines multiple approaches and high-performance computing for complex molecular science studies

What we offer:
- Expert staff
- Specialized facilities
- Unique instrumentation
- Science opportunity

Call for Proposals Open:
- Science Theme research
  - Opens in December/January
- JGI-EMSL Collaborative Science research
  - Letters of intent due April 6

http://www.emsl.pnl.gov
Questions?