Resolving Gene Regulatory Networks for Next Generation Strain Engineering

Sarah R. Smith

Andrew E. Allen
Scripps Institution of Oceanography
UCSD
J. Craig Venter Institute
Strain Improvement: Major Challenges

- Molecular Toolbox
  - Model organisms vs. new isolates
  - Reverse genetics
  - Transformation techniques
  - Karas et al. 2015 Designer diatom episomes delivered by bacterial conjugation. Nat. Commun. 6: 6925

- Gene Target Selection
  - A priori knowledge
  - Hypotheses (high-throughput)
  - Targeted approach

Multiple successes to metabolically engineer algae for enhanced production of desired compounds
Synthetic Biology as Next Generation Algal Strain Engineering

• What is Synthetic Biology?
The design and construction of new biological parts, devices, and systems, and the re-design of existing, natural biological systems for useful purposes. (OpenWetWare)

• Capabilities are expanding - brings new challenges
  – Introduction of entire episomes (~10kb), how would one rationally design a synthetic chromosome?

• Lot of unknown principles in genome regulation and even in well-studied organisms, gene regulatory networks remain unknown
Gene regulatory networks control cellular structure and function

INPUT (signal A):
Receptor proteins → Signal cascades

INPUT (signal B):
Receptor proteins → Signal cascades

Cell Function

Output

Cis-regulatory DNA sequence elements
Regulatory significance of gene order

Primary (~10 kb) cis-acting elements

Secondary (~100 kb) histone modifications

Highly-resolved time course transcriptomes

Phaeodactylum tricornutum

Thalassiosira pseudonana

Hurst et al. 2004
Regulatory significance of gene order

Shrestha et al. 2012, Smith et al. submitted, Allen et al. unpublished data
Regulatory significance of gene order

Shrestha et al. 2012, Smith et al. submitted, Allen et al. unpublished data
Regulatory significance of gene order

Shrestha et al. 2012, Smith et al. submitted, Allen et al. unpublished data
Regulatory significance of gene order

Regional co-expression: Chromatin dynamics

Shrestha et al. 2012, Smith et al. submitted, Allen et al. unpublished data
Regulatory significance of gene order

Shrestha et al. 2012, Smith et al. submitted, Allen et al. unpublished data
Microsynteny of Key Diatom Carbon Partitioning Genes

TPI-GAPDH fusion

GAPDH

T. pseudonana
P. tricornutum
F. cylindrus
P. multiseries

Smith et al. 2012 Algal Research
Inverted gene pairs are co-expressed *and* more common in chromalveolate genomes.

<table>
<thead>
<tr>
<th>Organism</th>
<th>Size (Mbp)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Guillardia theta</em></td>
<td>87.16</td>
</tr>
<tr>
<td><em>Emiliania huxleyi</em></td>
<td>167.7</td>
</tr>
<tr>
<td><em>Fragilaripsis cylindrus</em></td>
<td>80.5</td>
</tr>
<tr>
<td><em>Thalassiosira pseudonana</em></td>
<td>31</td>
</tr>
<tr>
<td><em>Phaeodactylum tricornutum</em></td>
<td>26.1</td>
</tr>
<tr>
<td><em>Pseudo-nitschia multiseries</em></td>
<td>218.73</td>
</tr>
<tr>
<td><em>Aureococcus anophagefferens</em></td>
<td>56.7</td>
</tr>
<tr>
<td><em>Cyclotella cryptica</em></td>
<td>180</td>
</tr>
<tr>
<td><em>Schizochytrium aggregatum</em></td>
<td>40.85</td>
</tr>
<tr>
<td><em>Aplanochytrium kerguelense</em></td>
<td>35.77</td>
</tr>
<tr>
<td><em>Aurantiopycnos limacinum</em></td>
<td>60.93</td>
</tr>
<tr>
<td><em>Phytophthora sojae</em></td>
<td>82.6</td>
</tr>
<tr>
<td><em>Naegleria gruberi</em></td>
<td>41</td>
</tr>
</tbody>
</table>

Conclusions:

- Strong functional significance of gene order in transcript-level gene regulatory networks (bidirectional pairs)
- Bidirectional genes common in central carbon and energy metabolism
- Bidirectional gene pairs enriched in photosynthetic organisms with compact genomes

Implications for the bioengineering of carbon metabolism and the efficient design of synthetic chromosomes
Acknowledgements

Andrew E. Allen

JCVI
Jonathan Badger
John McCrow
Flip McCarthy
Jeroen Gillard
Hong Zheng

SIO
Mark Hildebrand
Jesse Traller
Raffaela Abbriano
Roshan Shrestha

UCLA
Sabeeha Merchant & Lab
Matteo Pellegrini & Lab