Directed Evolution of \textit{de novo} Yeast Species for More Efficient Biomass Conversion

Frank Rosenzweig \textit{University of Montana, GCEP Co-Principal Investigator}
General nature of the problem

• Bio-ethanol production one of several strategies required to achieve energy independence & reduce greenhouse gas emissions

• To make biofuels production more economically attractive need to increase EtOH produced from non-comestible carbohydrates, decrease reliance on crops such as maize and sugarcane.

• Alternative feedstocks will include agricultural/silvicultural waste, as well as non-food crops grown for conversion to EtOH (e.g., switchgrass or softwoods).

Specific issues with yeast and biomass conversion

• $T_{opt}$ for growth and fermentation is 30-35C, at an industrial scale can only be maintained with considerable E input for cooling

• EtOH becomes toxic to yeast cells and inhibitory to fermentation at conc. >6-12%

• High %-age of CHO in alternative feedstocks pentose; not readily fermented by wildtype Saccharomyces species
**The quandary**
Yeast of the genus *Saccharomyces* grow on and aggressively ferment hexose, but do not grow or grow poorly on pentose.

Other yeast genera, such as *Pichia*, grow well on pentose, but relative to Bakers yeast, convert much more C to biomass, not ethanol.

**Solutions attempted**
2. Recombinant genetic engineering of Bakers yeast to ferment pentose (e.g., Ho et al. 1998. *AEM* 64:1852.)
Solutions attempted
2. Recombinant genetic engineering of Bakers yeast to ferment pentose (e.g., Ho *et al.* 1998. *AEM* 64:1852.)

Limited success with each
1. Hybrids quickly revert to parental genotypes
2. Fermentation yield < theoretical maximum; GMO
3. GMO; pleiotropy: *some* clones impaired ability to ferment hexose, *others* lose ability to ferment xylose anaerobically

GCEP projects 10, 17
Sherlock, G (PI), Rosenzweig, F (co-PI) An evolutionary genomic approach to industrial strain development
GCEP projects 10, 17
Sherlock & Rosenzweig, An Evolutionary Genomic Approach

1. Rather than start with single yeast clone (limited genetic variation), start with pool of MANY genetically diverse clones

2. Create pool of variants by *interspecific hybridization* within the *Saccharomyces* sensu stricto group → *de novo* species

3. Select over 100s generations for clones that have useful phenotypes
   A. Thermal tolerance
   B. EtOH tolerance
   C. Pentose utilization/fermentation

4. Recombine these traits by classical genetics (mating, sporulation).

5. Use genomics to discover mechanisms underlying complex traits.
Why this approach?
Wild-type *Saccharomyces* spp. do not ferment xylose, but they do have the genes required for this transformation!

Why don’t they use them?
Transcriptionally repressed?
Pseudogenes? (degenerate sequences that lack activity)
Metabolic constraints? (e.g. redox issues, *XDH* Sacch. vs Pichia)

“Genomic shock” of interspecific hybridization can
De-repress (or activate) genes
Restore activity in pseudogenes
Relieve metabolic constraints by altering regulation of related pathways
De novo yeast species for more efficient biomass conversion: An evolutionary genomic approach.

Implementing our approach
1. Identify & screen candidate parental strains

2. Create pools of de novo yeast species by interspecific hybridization*

3. Design culture conditions for experimental evolution

4. Develop & apply genomic tools to study adaptive evolution
   a. Assay large-scale chromosomal changes by PFGE
   b. Analyze genome content using 2-species microarrays

Research Plan and milestones for the coming year
Making interspecific hybrids

S. cerevisiae  ×  S. bayanus

Haploid Parent  Mate  Haploid Parent

Diploid F1 Hybrid

Sporulate

Haploid rare viable spores

Autofertilization

Diploid F2 Hybrids

Each a New Species!

10 Parental strains

S. bayanus

100 F1

F1

F2

1000s F2

+++

Gen var

Interspecific hybrids → de novo species → 100s Gens selection in chemostats

In independent, replicate exp select
A. ▲ thermal tolerance
B. ▲ EtOH tolerance,
C. Xylose metabolism

Analyze adaptive mutants
karyotype, genome content

Medium

Air

Culture

Temp Control

Effluent
1. Identify and screen candidate parental strains

Goals – Maximize standing genetic variation upon which natural selection act. Use “genome shock” of hybridization to uncover cryptic variation

**S. cerevisiae** parents

<table>
<thead>
<tr>
<th>Strain</th>
<th>Source</th>
<th>G418-resistant</th>
<th>flo8A::kanMX</th>
</tr>
</thead>
<tbody>
<tr>
<td>S288C</td>
<td>Lab strain, ATCC</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>CEN.PK113-1A</td>
<td>Lab strain, EuroSCARF</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>CEN.PK113-7D</td>
<td>Lab strain, EuroSCARF</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>FRY188</td>
<td>Rosenzweig collection</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>FRY198</td>
<td>Rosenzweig collection</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>FRY199</td>
<td>Rosenzweig collection</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>FRY200</td>
<td>Rosenzweig collection</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>FRY208</td>
<td>Rosenzweig collection</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>FRY212</td>
<td>Rosenzweig collection</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>FRY231</td>
<td>Rosenzweig collection</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>FRY232</td>
<td>Rosenzweig collection</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>FRY268</td>
<td>Rosenzweig collection</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>YJM270</td>
<td>vineyard</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>UC10</td>
<td>vineyard</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>M1</td>
<td>vineyard</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>YJM434</td>
<td>clinical</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>YJM280</td>
<td>clinical</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>YJM436</td>
<td>clinical</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>K9</td>
<td>sake</td>
<td>RES</td>
<td>RES</td>
</tr>
<tr>
<td>Y3</td>
<td>African palm wine</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Y5</td>
<td>African bili wine</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

+: G418-resistant isolate(s) recovered; -: resistant isolates not yet recovered; RES: strain displays high background resistance to G418

Strains screened for high T°C tolerance.

* FRY strains = Moroccan bread yeasts
**S. cerevisiae YJM270 – Y5**
Genetic, geographic, & ecological diversity

After Fey & Benavides 2005 PloS Genetics 1(1):66-71
### S. bayanus parents

<table>
<thead>
<tr>
<th>Strain (ATCC number)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSY31</td>
<td>Type strain</td>
</tr>
<tr>
<td>13055*</td>
<td>Umbrian wine must</td>
</tr>
<tr>
<td>204291*</td>
<td>Fermenting grape juice, South Australia</td>
</tr>
<tr>
<td>36023*</td>
<td>Champagne strain, Pasteur Institute</td>
</tr>
<tr>
<td>62014*</td>
<td>Wine, Spain</td>
</tr>
<tr>
<td>66283*</td>
<td>Champagne, France</td>
</tr>
<tr>
<td>76518*</td>
<td>Caddis fly, Spain</td>
</tr>
<tr>
<td>76673*</td>
<td>Beer, Rotterdam, the Netherlands</td>
</tr>
<tr>
<td>90738*</td>
<td>Wine, France; single-ascospore isolate</td>
</tr>
<tr>
<td>90919*</td>
<td>Wine, Montilla-Moriles, Spain</td>
</tr>
</tbody>
</table>

*ATCC number
1. **Identify and screen candidate parental strains**

High-throughput screening of candidate strains for growth on *xylose*

**Tecan Plate Reader**

- Reads optical density of 96 well plates at 15 minute intervals.
- Temperature-controlled
- Generate high-resolution growth curves for 96 cultures simultaneously.
• 672 Strains grown in YP xylose (2%).

• Sampled every 15 min for 68 h

• Data clustered to identify strains growing best
  - Best, ~3 doublings

• Selected strains will be re-tested YP xylose, YP no carbon, Minimal xylose, Minimal no carbon, to determine effect of xylose.
De novo yeast species for more efficient biomass conversion

Implementing our approach
1. Identify and screen candidate parental strains
2. Create pools of de novo yeast species

Mark prototrophic strains w/ traits that allow selection of interspecific F1 hybrids

Why prototrophs? Previous studies have used auxotrophic markers (e.g. requirement for an amino acid).
But . . .

- auxotrophic supplements (e.g. amino acid) provide C and N (~20%, and 50% of gluc and (NH4)2SO4, respectively); strong, but unwanted selective pressure
- auxotrophic mutants have altered metabolism; evolutionary changes they undergo may not be generalizable to wild-type industrial strains
- supplements required to “cover” auxotrophies too expensive to add to industrial-scale fermentations
2. Create pools of de novo yeast species

AbR to G418

**kanMX >**

pTEF tTEF

AbR to hygromycin

**hph >**

pTEF tTEF

Electroporation

Selection

Pools of antibiotic-resistant strains of *Saccharomyces species*
2. Create pools of de novo yeast species (cont.)

mass mating

selection for dual hygromycin & G418 resistant F1

sporulation and outgrowth

F2 interspecific hybrids between S. cerevisiae and S. bayanus (or S. paradoxus)
2. Create pools of de novo yeast species C

Introduce \( \text{Ab}^R \) cassettes in 2 forms

A. stably integrated at \( \text{FLO8} \) by homologous recombination

\( \text{FLO8} \) encodes transcriptional regulator involved in inducing flocculation; floc formation is undesirable trait in chemostat cultivation, so \( \text{FLO8} \) replacement will be advantageous to the process; presumed to be otherwise biologically neutral.

B. borne on an episomal \( 2\mu \) plasmid

plasmid transformation more efficient than integration; resulting \( \text{Ab}^R \) isolates can be cured of plasmid as needed by growth in non-selective conditions.
### S. cerevisiae strains *

<table>
<thead>
<tr>
<th>Strain</th>
<th>Source</th>
<th>G418-resistant Plasmid</th>
<th>flo8A::kanMX</th>
</tr>
</thead>
<tbody>
<tr>
<td>S288C</td>
<td>Lab strain, ATCC</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>CEN.PK113-1A</td>
<td>Lab strain, EuroSCARF</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>CEN.PK113-7D</td>
<td>Lab strain, EuroSCARF</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>FRY188</td>
<td>Rosenzweig collection</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>FRY198</td>
<td>Rosenzweig collection</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>FRY199</td>
<td>Rosenzweig collection</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>FRY200</td>
<td>Rosenzweig collection</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>FRY208</td>
<td>Rosenzweig collection</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>FRY212</td>
<td>Rosenzweig collection</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>FRY231</td>
<td>Rosenzweig collection</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>FRY232</td>
<td>Rosenzweig collection</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>FRY268</td>
<td>Rosenzweig collection</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>YJM270</td>
<td>vineyard</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>UC10</td>
<td>vineyard</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>M1</td>
<td>vineyard</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>YJM434</td>
<td>clinical</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>YJM280</td>
<td>clinical</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>YJM436</td>
<td>clinical</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>K9</td>
<td>sake</td>
<td>RES</td>
<td>RES</td>
</tr>
<tr>
<td>Y3</td>
<td>African palm wine</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Y5</td>
<td>African bili wine</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

+ = G418-resistant isolate(s) recovered; - = resistant isolates not yet recovered; RES = strain displays high background resistance to G418

* For further details see poster by Arle Kruckeberg
De novo yeast species for more efficient biomass conversion: An evolutionary genomic approach.

Implementing our approach
1. Identify & screen candidate parental strains
2. Create pools of de novo yeast species
3. Design culture conditions for experimental evolution

Continuous chemostat cultivation: ATR Sixfors fermenters; 6 vessels run simultaneously; indep. control of T°C, aeration, D h⁻¹, pH, medium comp. Medium fed from 10-L carboys; vessel V_{working} ≈ 0.3 L; cell density is ~3x10⁷ cell/mL Population size/vessel ~9x10⁹ cells.
Pilot evolutionary experiments

Single strains (S. cer, S. bay, F1, & 3 F2’s) used for proof-of-principle studies

**N-limitation**

- Excess hexose maintains catabolite repression of oxidative metabolism
- Approximates industrial fermentation: excess carbohydrate, low assimilable N

**Operating conditions**

- Dilution rate 0.10 – 0.17 h\(^{-1}\)
- Delft min. media 0.15 g/L \((\text{NH}_4)\_2\text{SO}_4\) and 9 g/L glucose
- Estimate: cell density \((\text{OD}_{600}, \text{cfu/mL})\), extracellular [Gluc], [EtOH], [NH\(_4\)]
- Sample every 2d \((\approx 10\text{ Gens})\)
- Archive samples as -80°C 15% glycerol stocks

At steady-state, residual glucose \(\approx 3\text{ gL}^{-1}\) extracellular EtOH \(\approx 2\text{ gL}^{-1}\)
De novo yeast species for more efficient biomass conversion: An evolutionary genomic approach.

Implementing our approach
1. Identify and screen candidate parental strains
2. Create pools of *de novo* yeast species
3. Design culture conditions for experimental evolution

4. Develop and apply genomic tools to analyze adaptive evolution
   a. Screen for chromosomal changes by PFGE
   b. Analyze genome content using 2-species microarrays
Ancestral strain …..[100s generations of selection]…… Evolved strain

embed intact chromosomes in agarose plugs

run plugs in a PFGE apparatus

compare ancestral chromosome bands to evolved chromosome bands
PFGE comparing founders with 3 independent isolates after >100 Gens

1-18: *S. cerevisiae* YNN295 (BioRad)
2-5: GSY86 (F1; founder & 3 evolved clones)
6-9: GSY121 (F2; founder & 3 evolved clones)
10-13: GSY144 (*S. cerevisiae*; founder & 3 evolved clones)
14-17: GSY31 (*S. bayanus*; founder & 3 evolved clones)
Evolution observed in F2 Exp populations by Gen 50
GSY86 (F1), GSY122 (F2), GSY123 (F2')

Sample from same exp vessel yields:
L = Lg colonies
S = Sm colonies

M and M’ S. cerevisiae marker
Lane 2: GSY86 Founder; 3: GSY86 A5; 4: GSY 86 A5; 5: GSY86 B5; 6: GSY86 B5;
15: GSY122 D6, “S”;
CONCLUSIONS from PFGE

Multiple examples of phenotypic & karyotypic evolution

In one replicate exp founded by S. cerevisiae, chromosome rearrangements appear by Gen 100

By Gen 50, cells in replicate F2 hybrid populations differentiated with respect to colony size on agar

“L” and “S” morphs have different karyotypes
De novo yeast species for more efficient biomass conversion: An evolutionary genomic approach.

Implementing our approach
1. Identify and screen candidate parental strains
2. Create pools of de novo yeast species
3. Design culture conditions for experimental evolution

4. Develop and apply genomic tools to study adaptive evolution
   a. Screen chromosomal changes by PFGE
   b. Analyze genome content by 2-species microarrays
**DNA microarrays** can be used to investigate genome content & gene expression globally.

Because genome sequence known for *Sacch* sensu stricto, possible to design microarrays that interrogate **all** genes from **both** parents used to make new species.

**Design of 2-species arrays to assay genome content (G. Sherlock)**

- Find 60-mer oligos within tiled 2-kb regions of *S. cerevisiae* and *S. bayanus* genomes
- Blast each oligo against combined *S. cer* and *S. bay* genomes to obtain probes unique to each genome and to rank
- Pick ~5000 probes per genome
- Include oligos that can be used as controls for normalization
- Use custom Agilent arrays—flexibility nice
Testing the 2-species array

- Label *S. cerevisiae* gDNA with green fluor (Cy3),
- Label *S. bayanus* gDNA with red fluor (Cy5)

(A) 2-species Agilent array of *S. cerevisiae* (red) vs. *S. bayanus* (green); (B) Histogram of log10 ratios of *S. cerevisiae* (red) vs. *S. bayanus* (green) spots. For most probes there is 10 to 100-fold discrimination between species and therefore, little cross-hybridization.
2-spp. array-Comparative Genomic Hybridization: a-CGH

Reference Sample → ~100 Gen → Experimental Sample

Ancestral Hybrid

Genomic DNA

Labeled DNA

DNA microarray

Evolved Hybrid
a-CGH provides further evidence for genome evolution in F2 hybrids

Ancestral F1

Ancestral F2

S. cer

S. bay

Evolved F2 “L” 100 Gens

Evolved F2 “S” 100 Gens
Outlook and Timeline

**Sept 06**
- Construction of hybrid pools
  - Selection EXP-A
    - S. cer X S. bay
    - T°C ramp-upX6
  - PFGE ExpA
  - a-CGH ExpA
  - Expression Profiling ExpA
  - RECOMBINATION ExpA X Exp B X ExpC

**March 07**
- Selection EXP-B
  - S. cer X S. bay
  - Xylose ramp-upX6
  - PFGE ExpB
  - a-CGH ExpB
  - Expression Profiling ExpB

**Aug 07**
- Selection EXP-C
  - S. cer X S. bay
  - EtOH ramp-upX6
  - PFGE ExpC
  - a-CGH ExpC
  - Expression Profiling Exp C
- Selection EXP-D
  - S. cer X S. par
  - Xylose ramp-upX6
  - PFGE ExpD
  - a-CGH ExpD
  - Expression Profiling Exp D
Acknowledgements

Gavin Sherlock
Assistant Professor, Stanford
Principal Investigator

Barbara Dunn
Research Associate, Stanford

Katja Schwartz
Research Scientist, Stanford

Arle Kruckeberg
Research Associate, Montana

Boris Stambuk
Visiting Professor

Heidi Kuehne
Postdoctoral Fellow, Montana

Paul Levine
Emeritus Professor, Stanford

Terry Paulish
Graduate Student, Montana