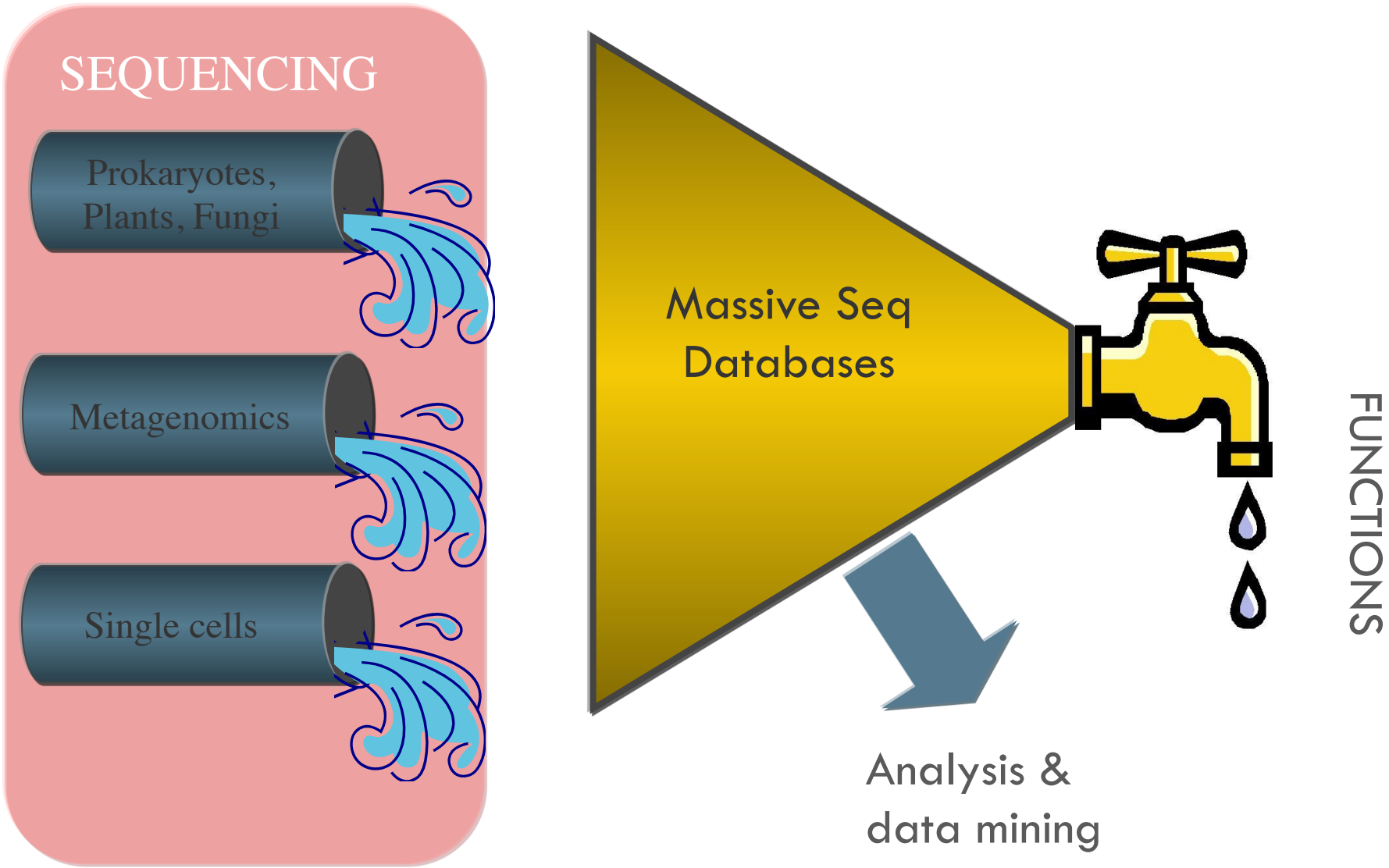
JGI JOINT GENOME INSTITUTE
UNITED STATES DEPARTMENT OF ENERGY

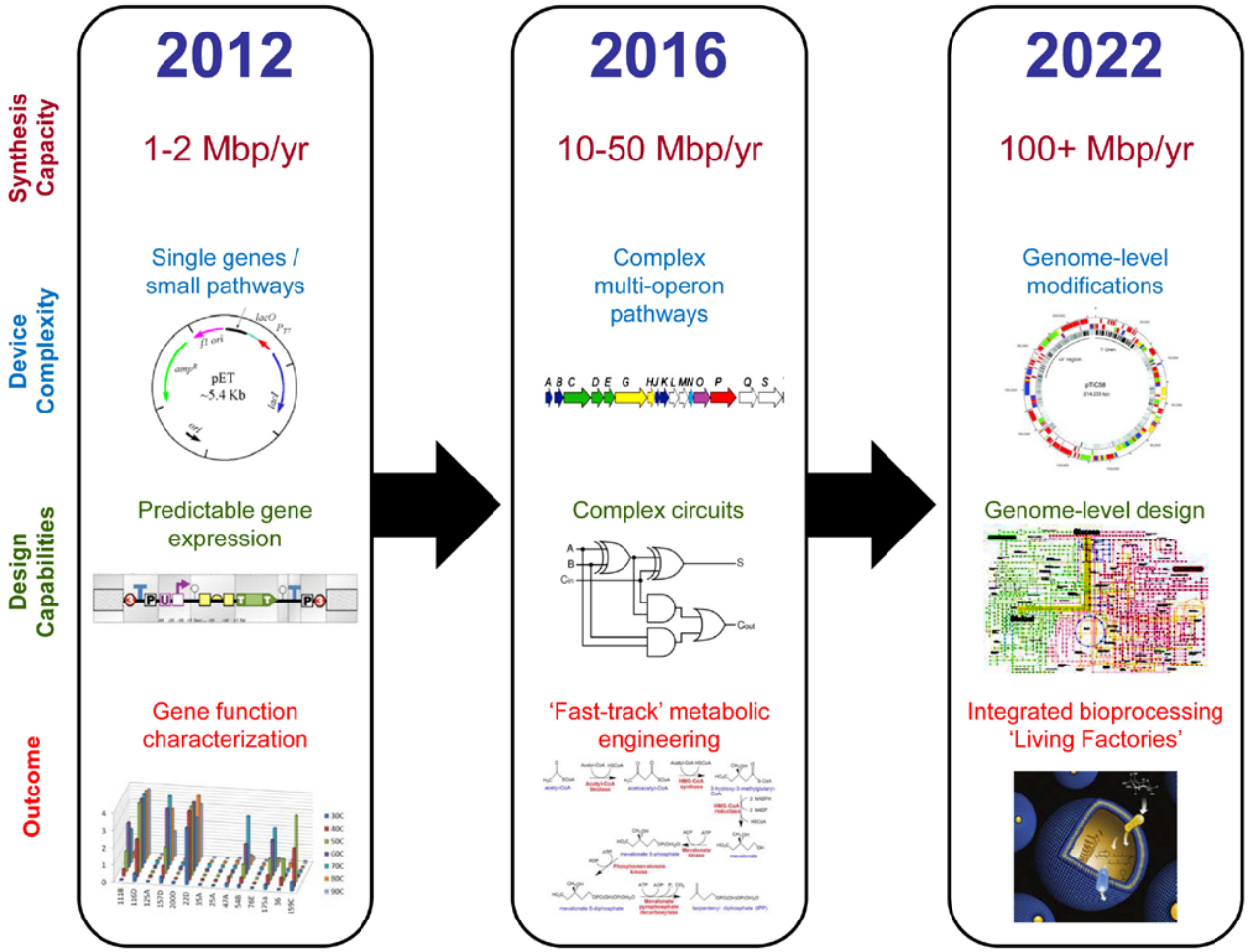
DNA synthesis Pipeline and applications

Sam Deutsch
DNA Synthesis Science

Why Synthesis ?



10-year Vision for Evolving Capabilities



10 year vision for DNA Synthesis at the DOE JGI.

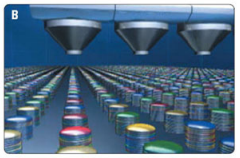
JGI 'Build' Pipeline

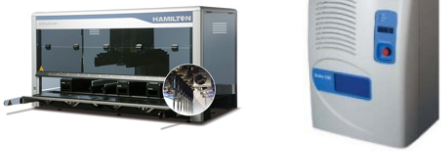
Project Management
Calls
Review
ELSI
Initiation

phytozome
UniProt
img
Data mining

GENE DESIGN
j5
BioStudio
Design Software

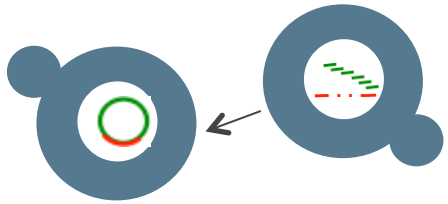
Gen9
IDT INTEGRATED DNA TECHNOLOGIES
1-Kb blocks



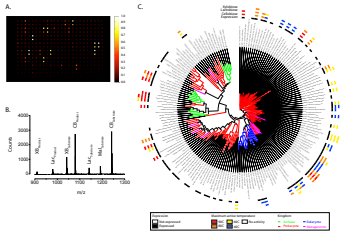


Liquid handling systems

Cloning and picking



Multi-fragment assembly



Functional Characterization

Design

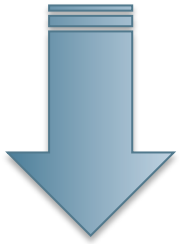
Goal: To enable scalable pathway refactoring from coding sequence information (DBs) to order sheets and protocols

Multiple tools:

- GeneDesign/DNAssembler (Codon optimization, partition)
- DIVA/J5 (Combinatorial library design)
- ICE (Parts repository)
- SPL (sequence polishing library… New)
- IMG2DE (from databases to Synbio design… New)

Interact with other external synbio tools: RBScalculator, Pigeon

- SeqScreen (Biosecurity screening tool.. not really design)



Automatically Design 1000s of pathways

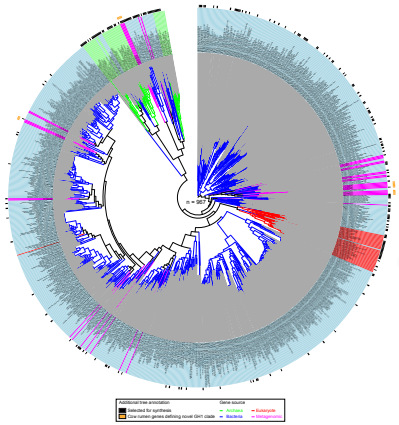
Industry relevant GH1s

JGI-JBEI collaboration

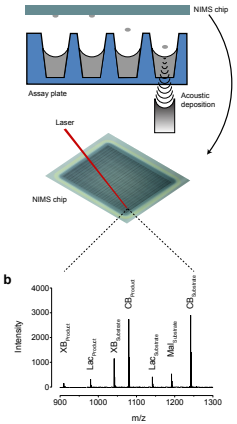
- Interested in GH1 enzymes for biomass deconstruction
- Looking for enzymes that can operate under industrial conditions (70 ºC, 20% ILs)

Mine Databases

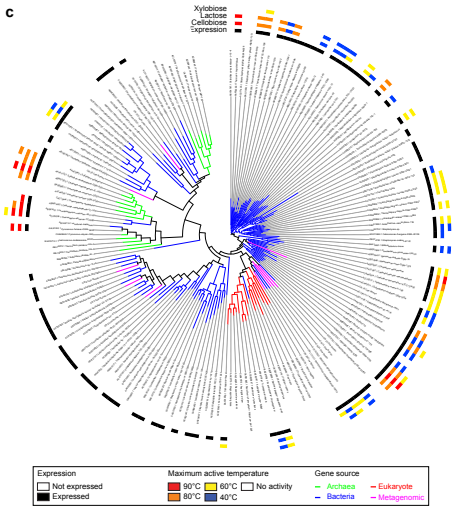
Phylogenomic selection



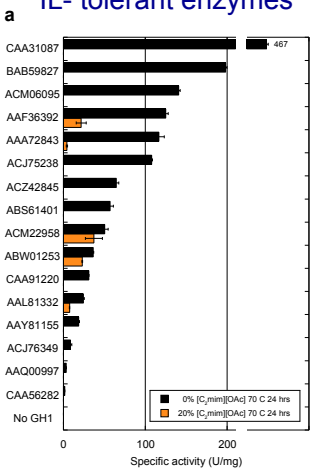
HTP characterization



Annotated tree



IL- tolerant enzymes



ACS Chem Biol,2014

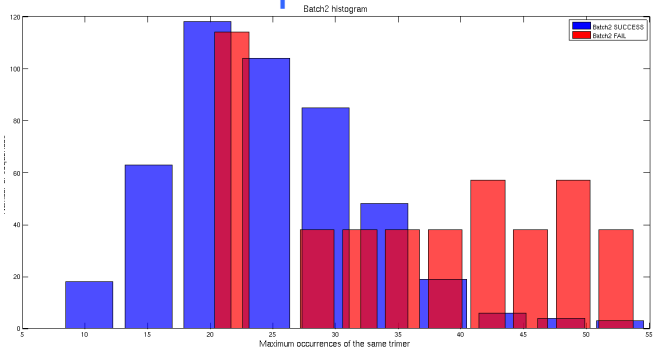
Characterize transcription factors in model plants

CSP: Sam Hazen (UM Amherst)

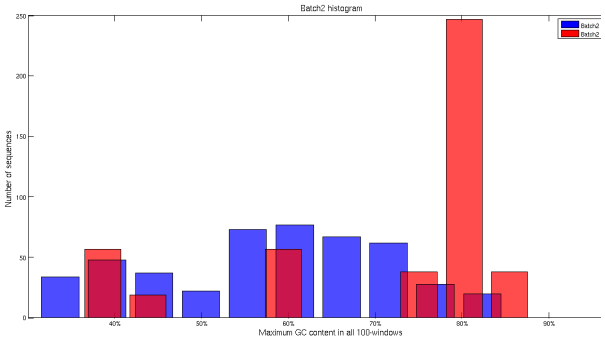
Synthesis of about 600 TFs governing cell wall metabolism (Y1H, overexpression, KOs).

Round 1= 300 TFs: Success rate: 55%, Cycle time ~ 1 Year

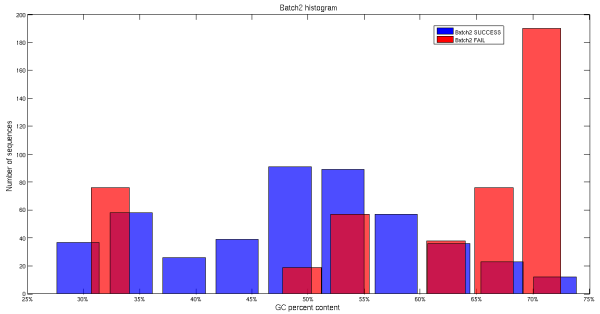
Trimer repeats



Global GC



Local GC



Successful synthesis Failed synthesis

Other significant features:

- Direct repeats
- Homopolymers
- Hairpins
- Composition of sequence ends

Implement rules during design stage
Higher success rate ??

Round 2= 300 TFs: Success rate: 95%, Cycle time: 60 days

Apply learning strategy to all Synbio steps: Synthesis, Assembly, Expression, Phenotype (improve Design)

Characterize Orphan Fungal PKS Clusters

CSP: Istvan Molnar (Arizona State)

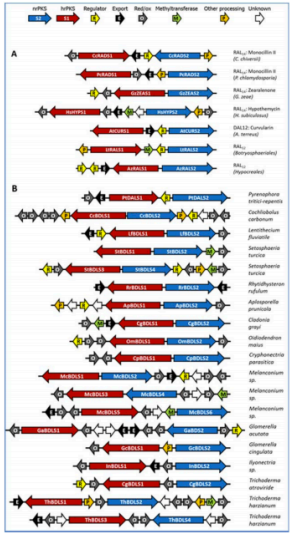
- Looking for Biosynthetic clusters encoding Reduced PKS
- Biofuel replacement molecules
- Monomers for new materials
- Molecules mediate ecological interactions

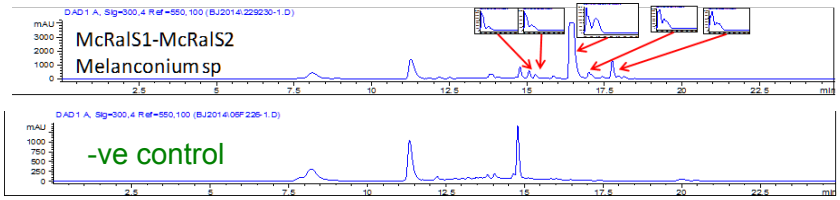
Mine JGI Databases

MycoCosm

JGI:
RNA seq
Manual curation

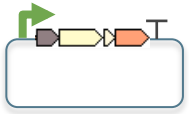
Curated Biosynthesis clusters





8/10 strains generate novel compounds

Strain generation



JGI:
Synthesis & Assembly
QC

GENE DESIGN

j5

JGI:
Construct Design
Expression :Promoters/Terminators, Topology
Optimization + Assembly strategy

37/40 delivered to user in ~100 days

Focus Area: Secondary Metabolism

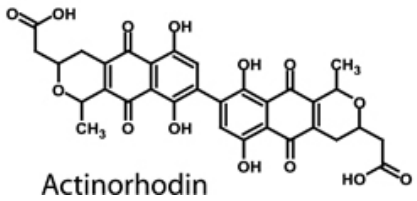
Characterize the role of secondary metabolites in mediating microbe-microbe and plant-microbe interactions and their effect on biological systems

Why study secondary metabolites ?

- **Constitute 15-20 % of gene content in certain environments. Poorly characterized, and ecological role is largely unknown**
- **Thought to be important for signaling (e.g., Quorum sensing, Flavonoids-RNB) and impact physiology (competence)**
- **Thought to be key determinants of community composition, and thus has large impact on biogeochemstry.**

Actinorhodin

Actinorhodin: polyketide antibiotic produced by *Streptomyces coelicolor*, requiring 22 genes for biosynthesis. Cluster ~ 25 kb, %GC >70



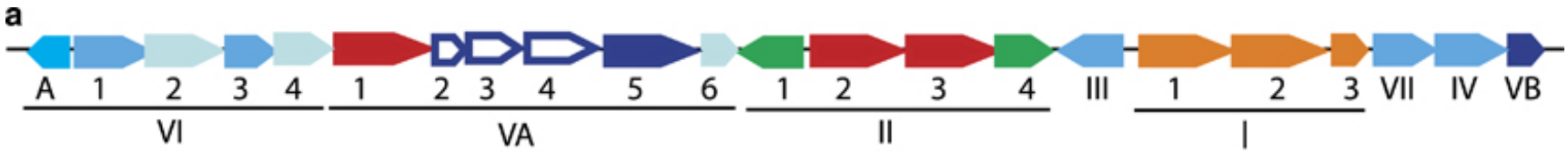
Why Actinorhodin?

- Representative of many biosynthetic clusters (size, complexity, GC content)
- Pathway is well characterized (but never before refactored)
- Actinorhodin is readily detectable (blue/red pigment depending on pH)

Actinorhodin pathway refactoring

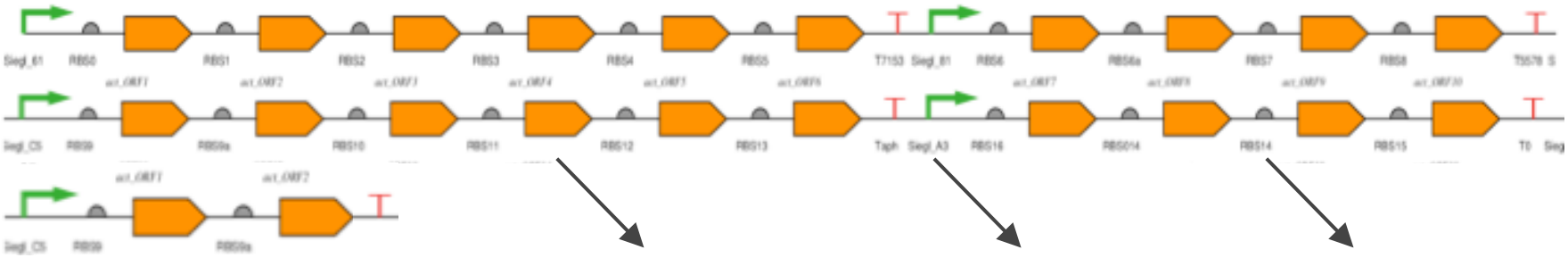
"Refactoring": Redesigning a complex set of operons that are highly regulated under native conditions to predictably achieve a phenotypic outcome under laboratory conditions

Native pathway:



Highly regulated expression; unknown control system

Refactored Design:



Codon optimized genes: lower GC, remove repeats & secondary structure, preserve high translational potential

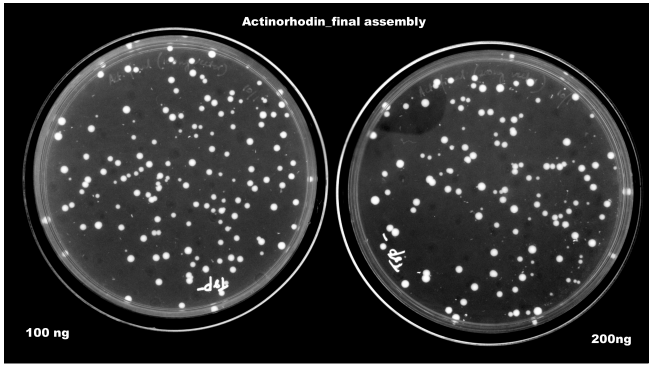
Previously validated constitutive promoters and terminators

Calculated optimal RBSs specific for each rCDS

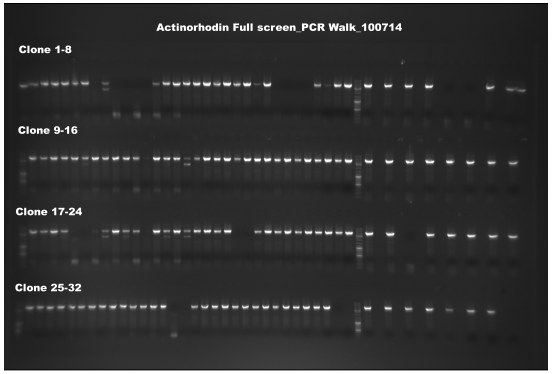
Goal: predictable expression

Actinorhodin pathway construction

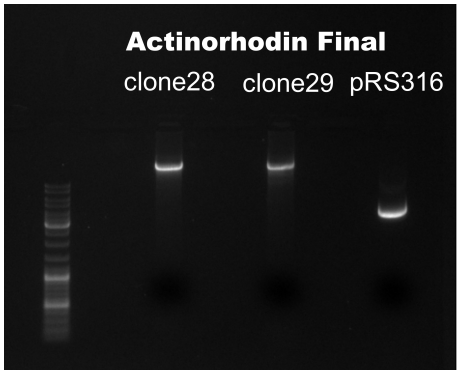
- Pathway contains 22 genes for a total of 25 kb
- Synthesized as randomly partitioned 1 kb blocks
- Pre-assembled into cloned and sequenced 5 kb blocks via Gibson assembly
- Final assembly of 5 kb blocks via yeast TAR cloning into *Streptomyces* integration vector



High efficiency yeast TAR cloning

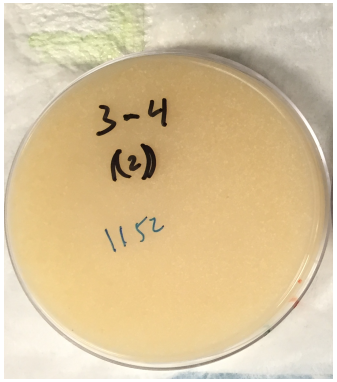


Screened clones via PCR walking;
Sequenced 8 positive clones
using PacBio

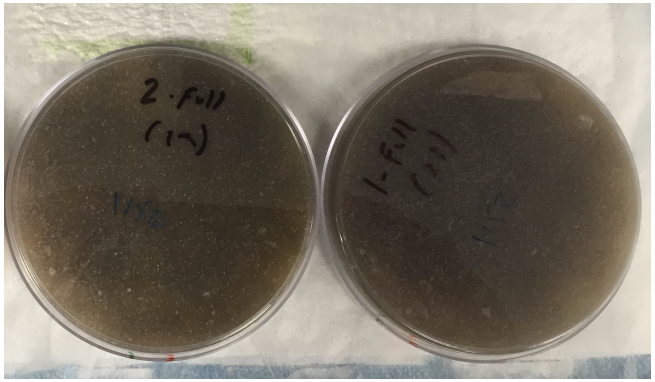


Two preps used for
Streptomyces coelicolor
strain generation

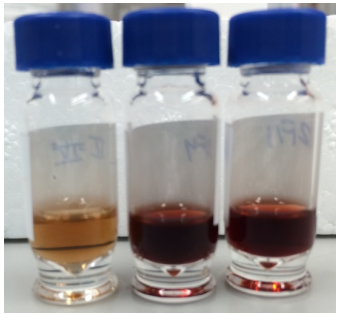
Actinorhodin strain characterization



S. coelicolor Δ*act strain* (control)

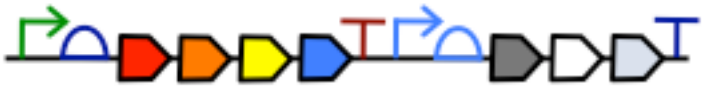


Refactored actinorhodin strains



Extracts for LC-MS analysis

JGI Synthesis Science: Not Just Synthesis!



Gene or Pathway Design

phytozome

img/m

MycoCosm
the fungal genomics resource

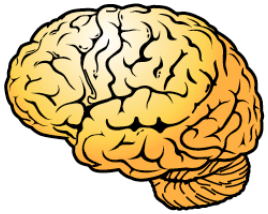
Automated DNA Construction

NGS sequencing

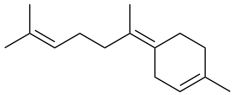
Commercial space

Hosts/Strain

High-throughput functional assessment



Success/failure learning



USER

Summary

- Described rationale for DNA synthesis Science at JGI
- Described pipeline
- Emphasis on integrated approach, where synthesis is only a component of the program. Capability suite includes data mining, design, construction and characterization
- Examples of science applications and focus area