

Generating a Synthetic Biology Toolbox for Nitroorganics

WP-2339

Jon Magnuson, PI

Shuang Deng (presenting)

Pacific Northwest National Laboratory

Clay Wang, co-PI

University of Southern California

Project Outbrief

May 5, 2023



REPORT DOCUMENTATION PAGE

*Form Approved
OMB No. 0704-0188*

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

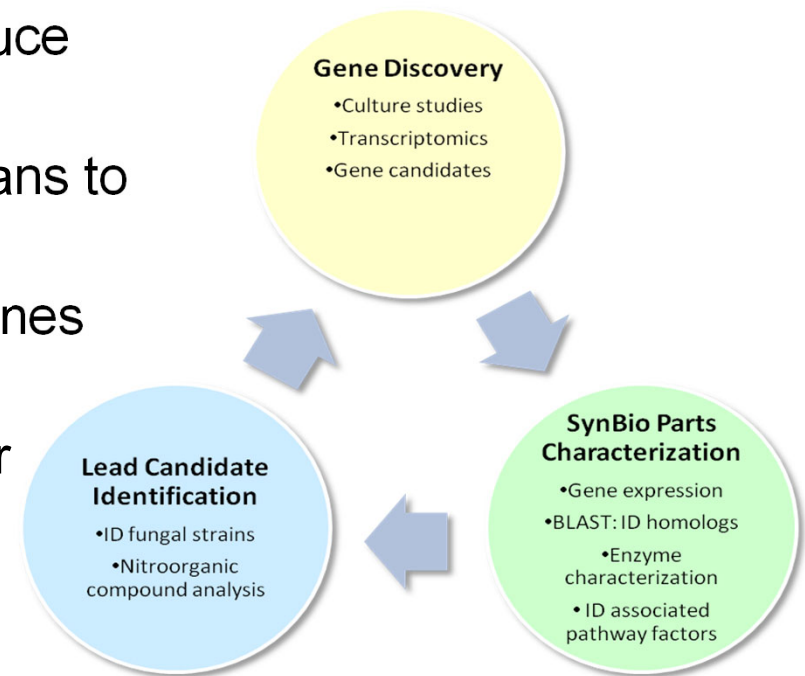
1. REPORT DATE (DD-MM-YYYY) 05/05/2023			2. REPORT TYPE SERDP Project Outbrief		3. DATES COVERED (From - To)	
4. TITLE AND SUBTITLE Generating a Synthetic Biology Toolbox for Nitroorganics					5a. CONTRACT NUMBER	
					5b. GRANT NUMBER	
					5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Jon Magnuson Shuang Deng Pacific Northwest National Laboratory Clay Wang University of Southern California					5d. PROJECT NUMBER WP-2339	
					5e. TASK NUMBER	
					5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Pacific Northwest National Laboratory 902 Battelle Blvd P.O. Box 999, MS P8-60 Richland, WA 99352					8. PERFORMING ORGANIZATION REPORT NUMBER WP-2339	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Strategic Environmental Research and Development Program 4800 Mark Center Drive, Suite 16F16 Alexandria, VA 22350-3605					10. SPONSOR/MONITOR'S ACRONYM(S) SERDP	
					11. SPONSOR/MONITOR'S REPORT NUMBER(S) WP-2339	
12. DISTRIBUTION/AVAILABILITY STATEMENT DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.						
13. SUPPLEMENTARY NOTES						
14. ABSTRACT Enabling synthetic biology routes to energetic materials through discovery of the pathways (genes/enzymes) needed for production of the necessary functional groups.						
15. SUBJECT TERMS Synthetic Biology Toolbox, Nitroorganics						
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UNCLASS	18. NUMBER OF PAGES 38	19a. NAME OF RESPONSIBLE PERSON Jon Magnuson	
a. REPORT UNCLASS	b. ABSTRACT UNCLASS	c. THIS PAGE UNCLASS			19b. TELEPHONE NUMBER (Include area code) 509-372-4119	

Goal and General Approach

Goal: Enabling synthetic biology routes to energetic materials through **discovery of the pathways (genes/enzymes)** needed for production of the necessary functional groups.

Approach:

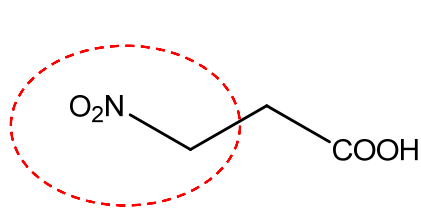
- Identify fungi and culture conditions that produce nitroorganic compounds
- Use comparative transcriptomics or other means to identify candidate genes
- Deletion analysis to determine if candidate genes are necessary for production of the target
- Express candidate genes to demonstrate their biosynthetic utility and confirm their products



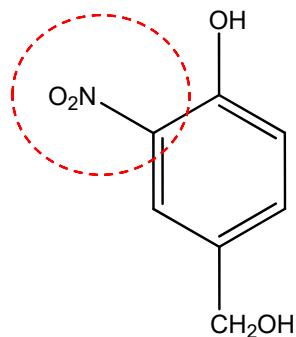
Technical Background

Biology of Nitroorganics

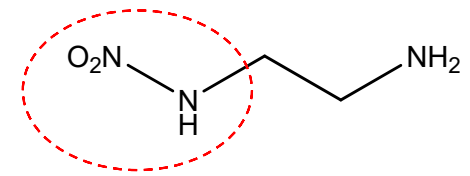
- Many microorganisms (fungi, bacteria, etc.) produce them
- 100's of natural **nitroorganic** compounds known
(Parry, Nishino, & Spain. 2011. Naturally-occurring nitro compounds. Nat Prod Reports 28:152-167)
- Contain a variety of functional groups that are found in **energetic materials**: $R-NO_2$, $Ar-NO_2$, $R-NH-NO_2$, $R-O-NO_2$



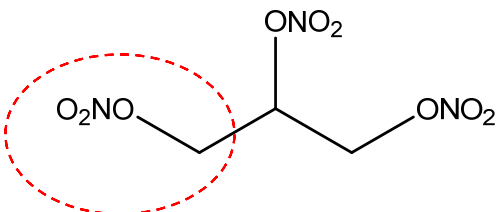
3-nitropropanoic acid



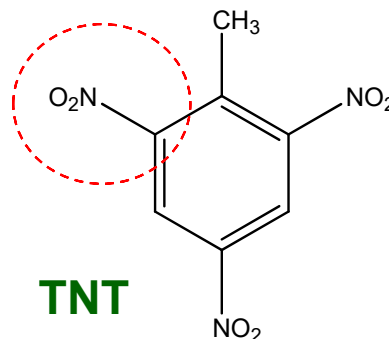
4-hydroxymethyl-2-nitrophenol



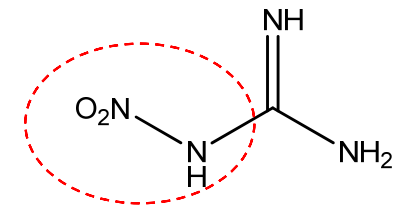
N-nitro-1,2-ethanediamine



nitroglycerin



TNT



1-nitroguanidine

Candidate Fungi and Nitroorganics

- Identify fungal strains making nitroorganics from the literature and obtain fungal strain candidates (targets prioritized)
- Our **top three candidates** are below

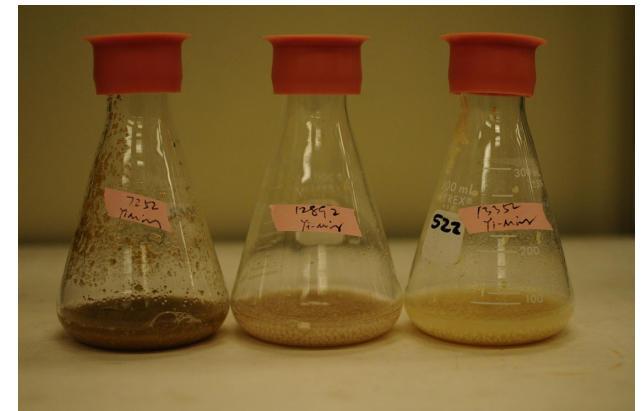
Fungus	Nitroorganic compound
<i>Aspergillus oryzae</i>	3-nitropropionic acid
<i>Penicillium citrinum</i>	Citrinalins A & B
<i>Aspergillus wentii</i>	1-amino-2-nitrocyclopentanecarboxylic acid
<i>Coniothyrium sp.</i>	1-hydroxy-5-methoxy-2,4-dinitronaphthalene
<i>Agaricus subrutilescens</i> & <i>A. sylvaticus</i>	N-(2-aminoethyl)nitramide
<i>Lepista diemii</i>	4-nitrobenzaldehyde
<i>Phellinus robiniae</i>	1,2,4-trichloro-3,6-dimethoxy-5-nitrobenzene
<i>Pyricularia oryzae</i>	4-hydroxy-3-nitrobenzyl alcohol
<i>Penicillium cyclopium</i>	Cyclopiamine A & B

Target 1: 3-nitropropionic acid (3-NPA)

- Obtained three different fungal strains reported to produce 3-NPA and screened them for 3-NPA production
- Selected the **best 3-NPA production strain, *Aspergillus oryzae* 12892**
- Determined culture conditions and sampling times to enable **transcriptomics** experiment

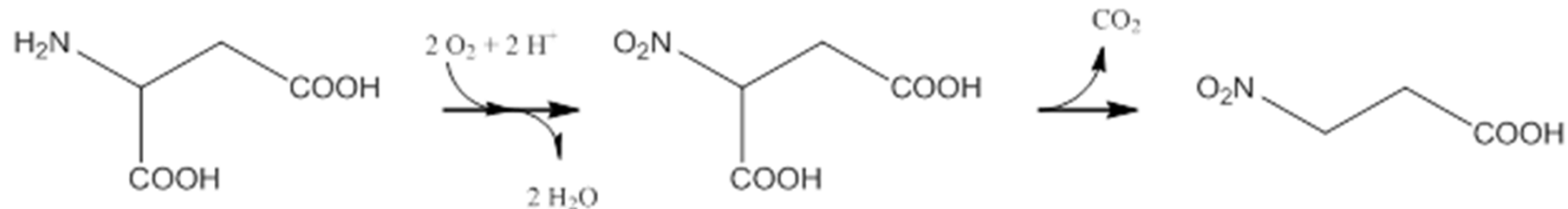
Strain	Diammonium tartrate (g/L)	Culture Conditions	3-NPA (ppm) at Day:			
			1	3	5	7
A. oryzae 7252	8.0	still, dark	0.2	nd	856.9	0.7
A. oryzae 12892	8.0	still, dark	0.2	0.1	308.5	793.7
P. melinii 13352	8.0	still, dark	nd	0.2	2.1	248.6
A. oryzae 7252	2.7	still, dark	nd	0.1	nd	nd
A. oryzae 12892	2.7	still, dark	0.2	0.2	nd	nd
P. melinii 13352	2.7	still, dark	nd	21.9	nd	1.4
A. oryzae 7252	8.0	shaken	nd	0.2	nd	709.5
A. oryzae 12892	8.0	shaken	0.1	6.6	308.4	339.8
P. melinii 13352	8.0	shaken	nd	nd	212.6	3.8
A. oryzae 7252	2.7	shaken	nd	85.7	nd	nd
A. oryzae 12892	2.7	shaken	0.2	121.7	nd	nd
P. melinii 13352	2.7	shaken	nd	1.0	1.6	nd

nd = none detected



Target 1: 3-NPA

3-Nitropropionic Acid (3-NPA) Biosynthesis, Proposed

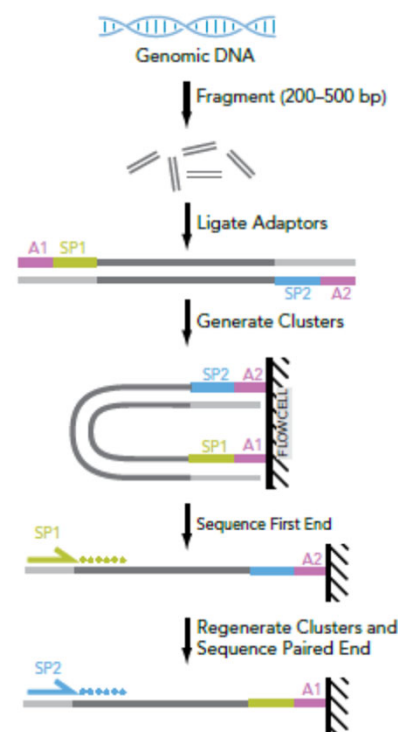
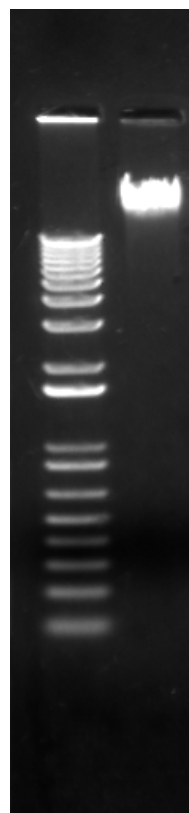


3-NPA

- The 1st step is **enzymatic**; aspartate-N-oxygenase—gene(s) unknown
- The 2nd step may be spontaneous
- Alternative more complicated pathways have been proposed
- Utilize **comparative transcriptomics** to identify gene candidates

Target 1: 3-NPA

De novo genome assembly



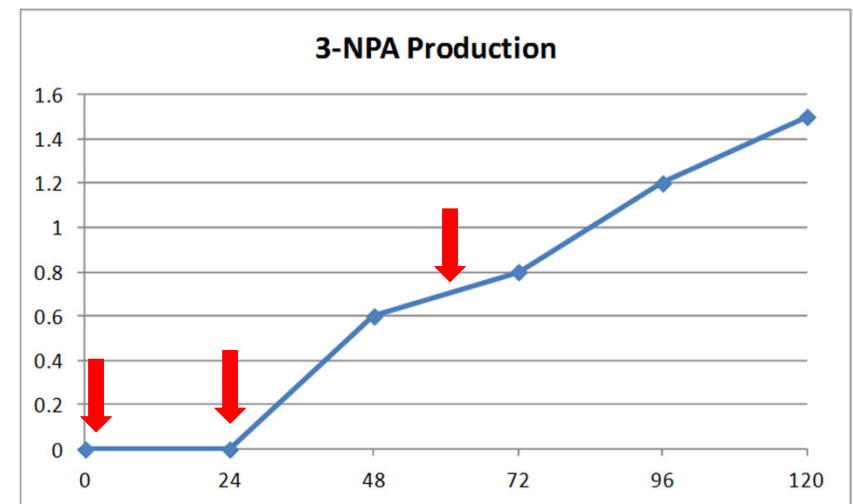
	A. oryzae genome DNA
Name	
# of reads	28,046,040
Avg. length	149.9
# of reads after trim	28,035,212
% of remaining after trim	99.96%
Avg. length after trim	147.8

Summary of sequencing Data

Illumina TruSeq
Paired-End Sequencing workflow

3NPA: Gene Discovery

1. Grow fungi under different conditions to obtain product biosynthesis profile
 2. Select time points for RNA sampling
 3. Obtain transcriptomics data
 4. Perform comparative analysis of transcriptomics data to identify likely candidate genes for nitroorganic biosynthetic enzymes, export, “cluster” regulation
 5. Parts Characterization
- **Red arrows** indicate **RNA sampling** times to catch onset of expression of biosynthesis genes

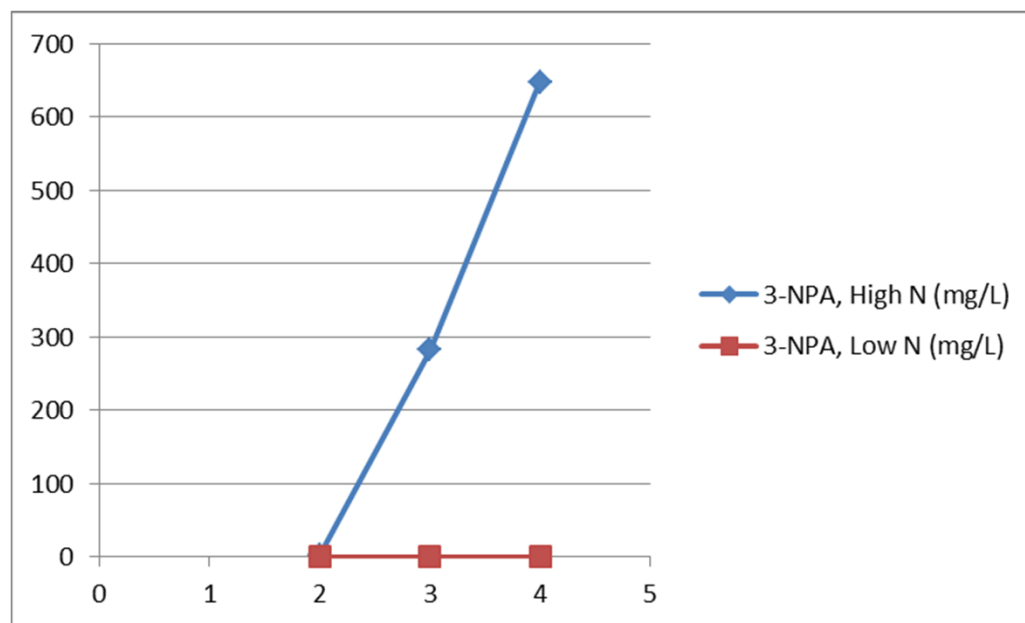


Time (hr)

3NPA Production and Suppression Conditions

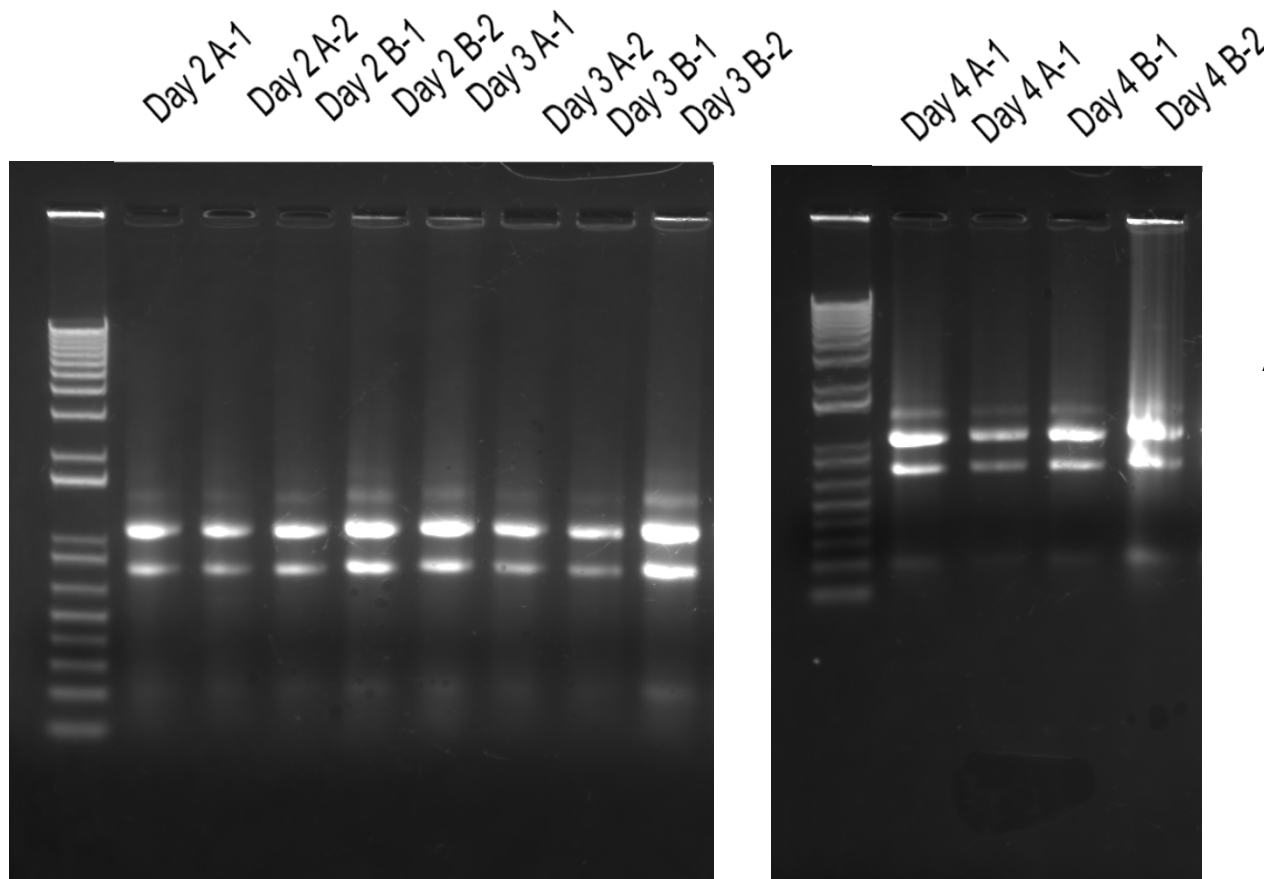
High ammonium tartrate medium gives high production
 1/3 of that concentration eliminates detectable 3NPA

Raulin-Thom Solution	Production g/L	No Production g/L
Glucose	45.5	45.5
K ₂ CO ₃	0.4	0.4
MgCO ₃	0.27	0.27
FeSO ₄ ·7H ₂ O	0.047	0.047
ZnSO ₄ ·7H ₂ O	0.047	0.047
ammonium tartrate dibasic	8	2.7
tartaric acid	2.7	2.7
(NH ₄) ₂ SO ₄	0.17	0.17
(NH ₄) ₂ HPO ₄	0.4	0.4



3NPA: Gene Discovery

Transcriptomics analysis



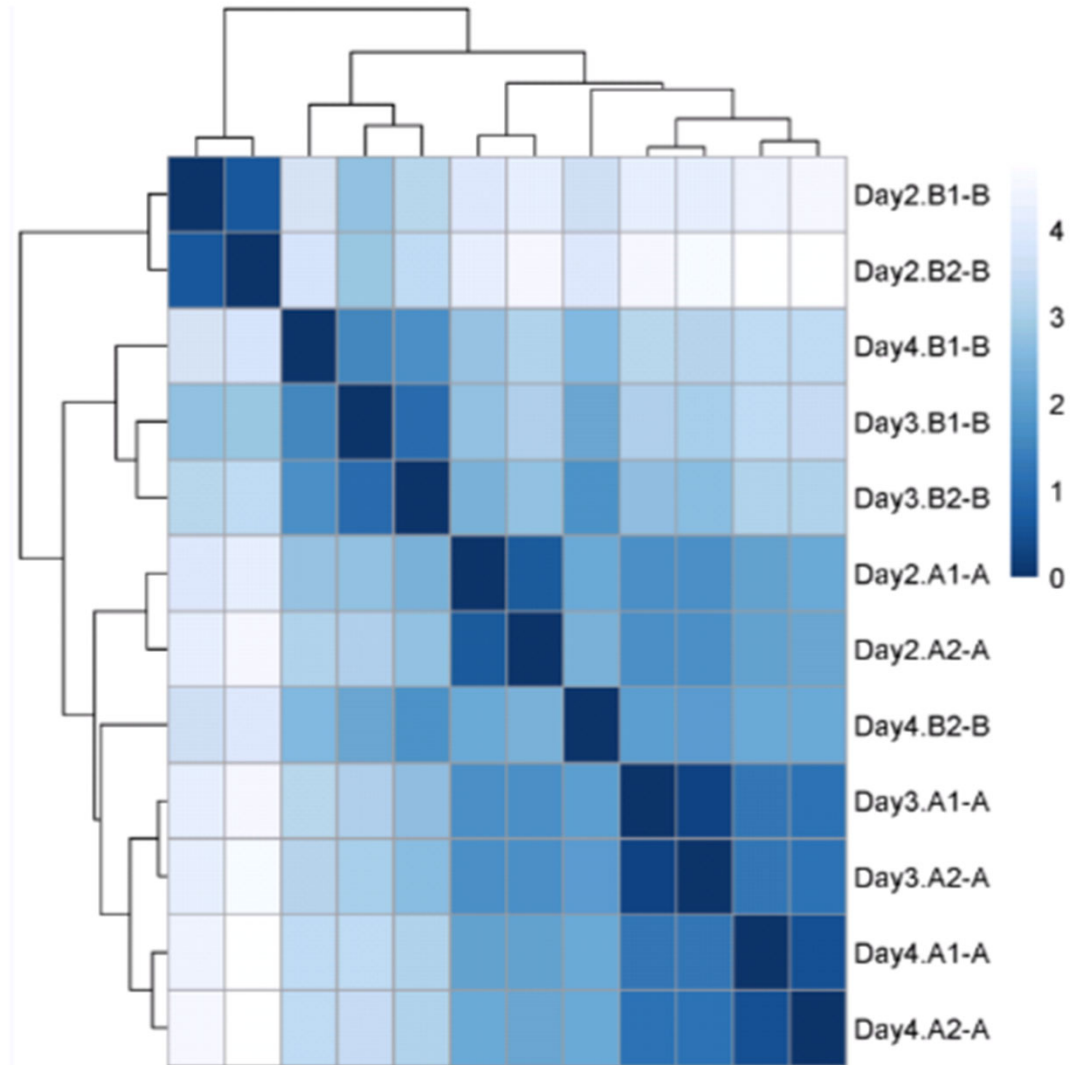
A: No production medium
B: Production medium
1,2: Duplicate samples

RNA Extraction

3NPA: Comparative transcriptomics

Samples sequenced by Illumina

- Duplicate samples for days 2, 3 & 4
- Production vs. non-production media



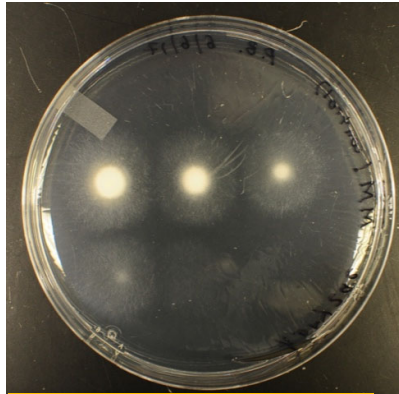
Difference in gene expression between groups

3NPA: some Potential Targets

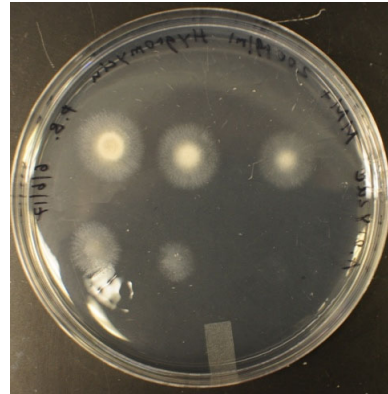
log2 Fold Change	Gene Description
9.39	oxidoreductase, zinc-binding dehydrogenase family, putative
9.24	alkanesulfonate monooxygenase
7.00	Nitrilotriacetate monooxygenase component A / pristinamycin IIA synthase subunit A
5.91	Naringenin, 2-oxoglutarate 3-dioxygenase
4.92	alpha-ketoglutarate-dependent taurine dioxygenase, partial
4.43	Nitrilotriacetate monooxygenase component A
3.29	monooxygenase FAD-binding protein; 2-polyprenyl-6-methoxyphenol hydroxylase and related FAD-dependent oxidoreductases
2.14	FAD-dependent oxygenase, putative
1.86	putative flavoprotein involved in K ⁺ transport

3NPA : Parts Characterization

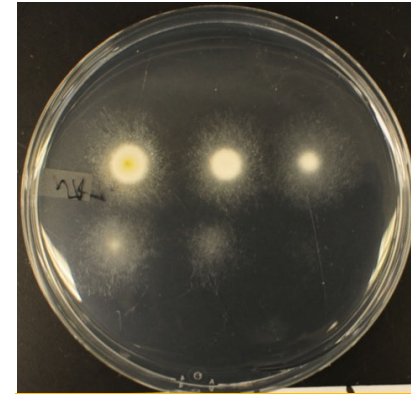
Screening for antibiotic sensitivity in *A. oryzae*



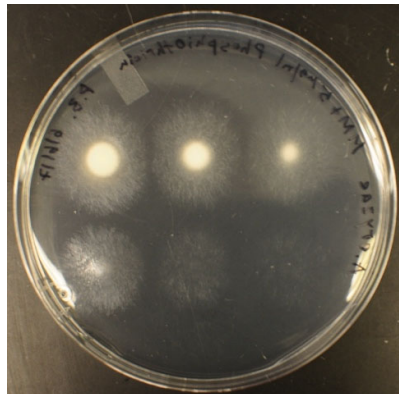
No antibiotic



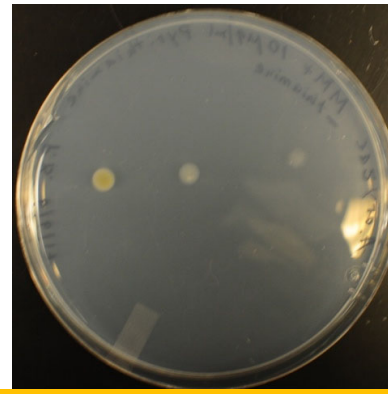
Hygromycin 200 ug/L



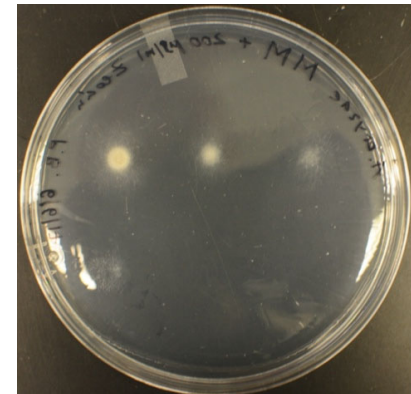
Nourseothricin 30 ug/L



Phosphinothricin 5 ug/L



Pyrothiamine 10 ug/L

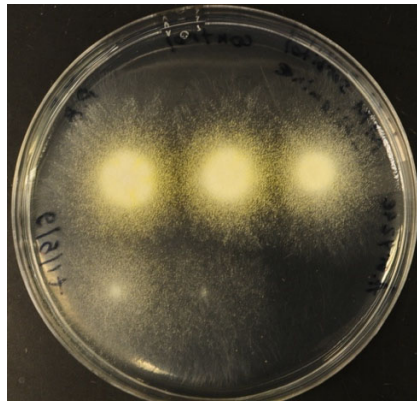


Zeocin 200 ug/L

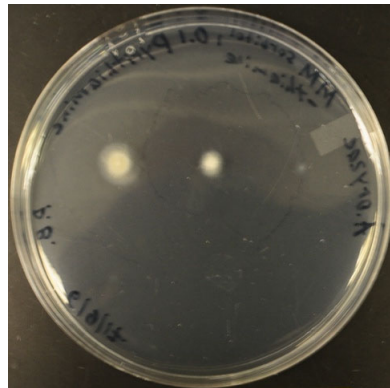
***A. oryzae* is sensitive to Pyrothiamine and Zeocin**

3NPA : Parts Characterization

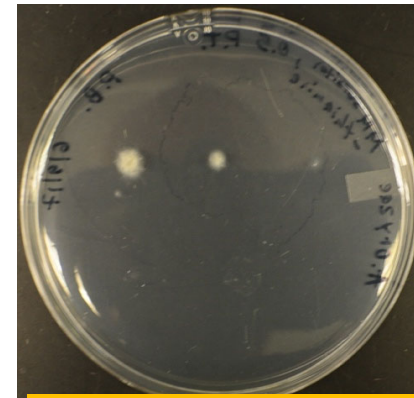
Growth on MM-thiamine with different Pyrothiamine concentrations



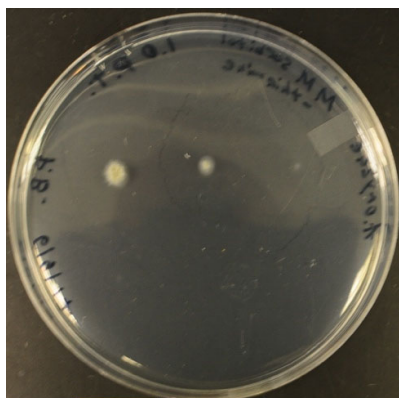
No antibiotic



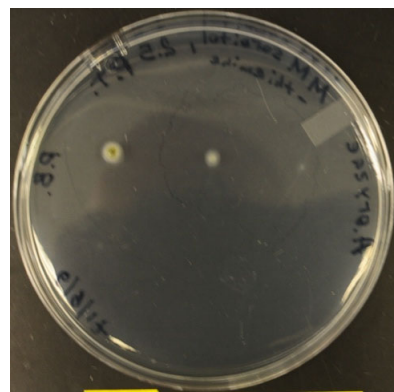
Pyrothiamine 0.1 ug/L



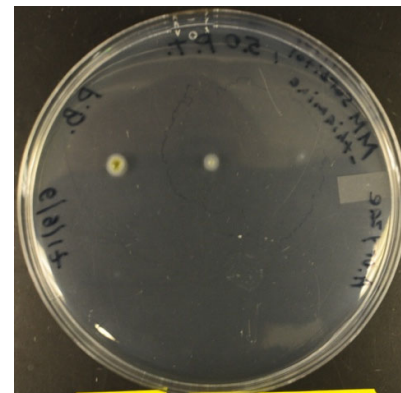
Pyrothiamine 0.5 ug/L



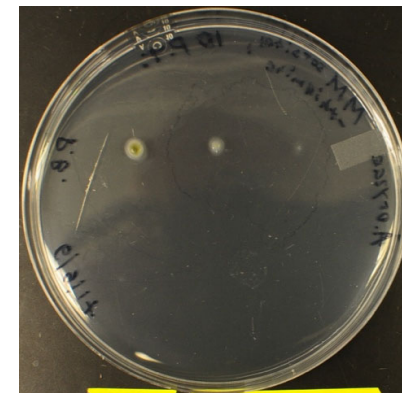
Pyrothiamine 1.0 ug/L



Pyrothiamine 2.5 ug/L



Pyrothiamine 5.0 ug/L

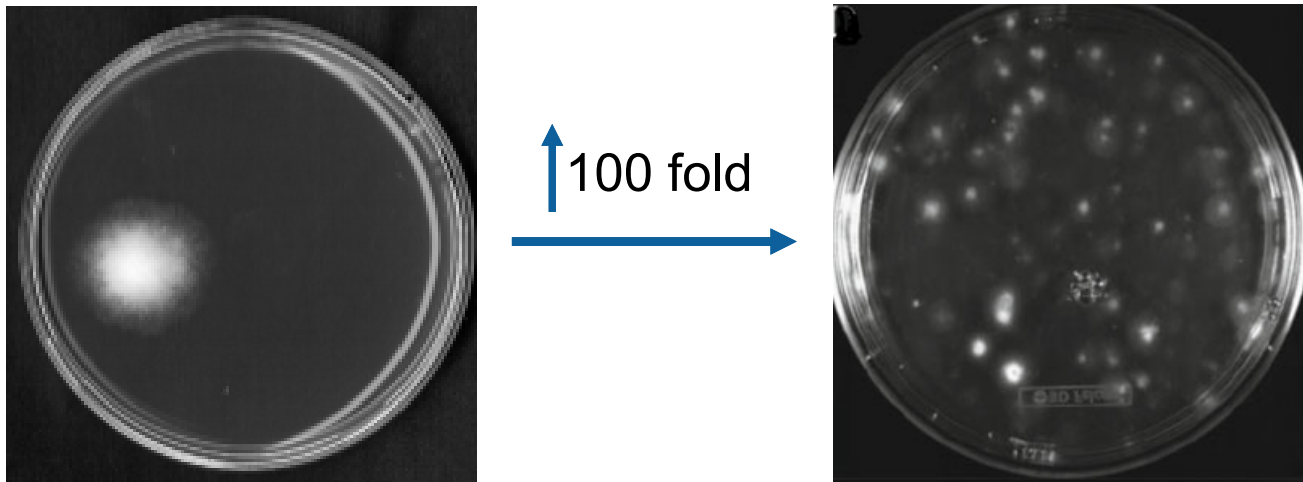


Pyrothiamine 10 ug/L

Pyrothiamine 1ug/ul will be used as selection

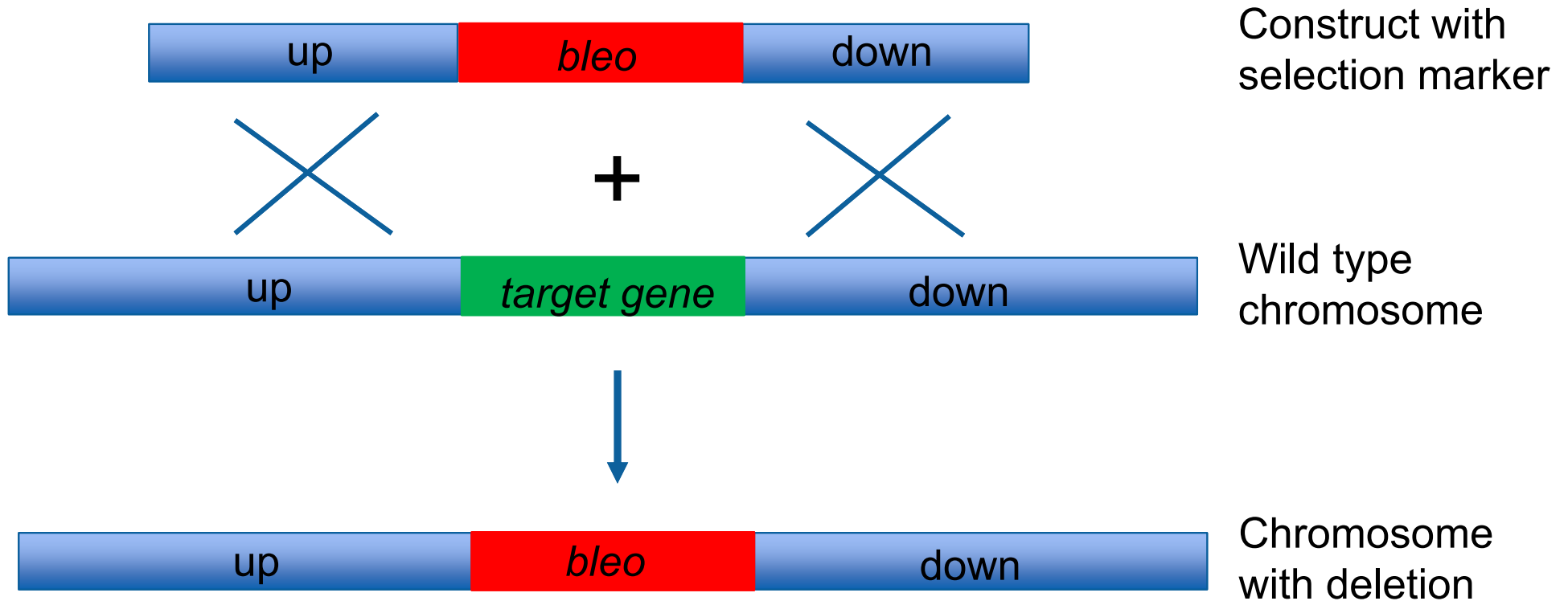
3NPA

Optimizing protoplast transformation protocol



- Protoplasting transformation efficiency was improved dramatically, from one or two per plate to around 100 per plate
- This enables screening for deletion mutations

3NPA: Delete target genes

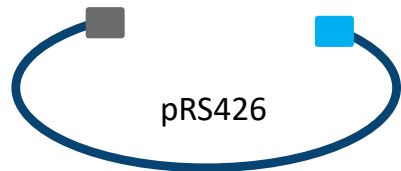


Three requirements:

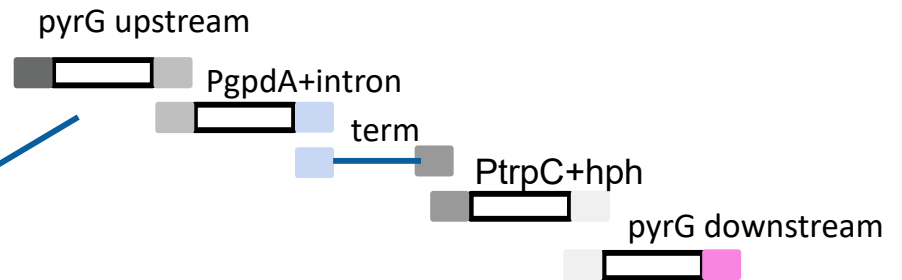
1. Genome sequences
2. Efficient Transformation
3. Homologous recombination, ku70 deletion

3NPA: Gene Expression Vector

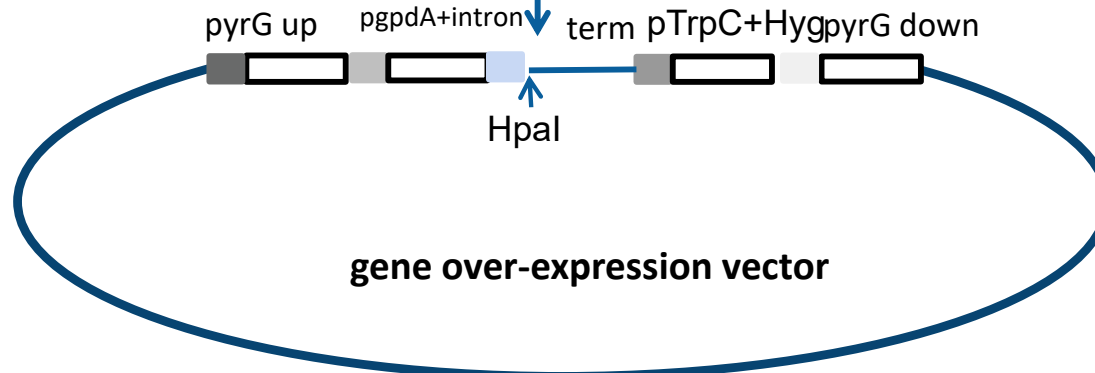
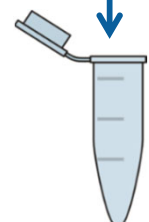
1. Generate a linear vector by NcoI digestion



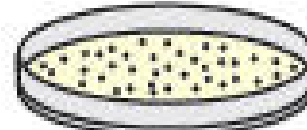
2. PCR amplify DNA inserts with overlapping ends



3. Add fragments and vector to **Gibson assembly** Master Mix.



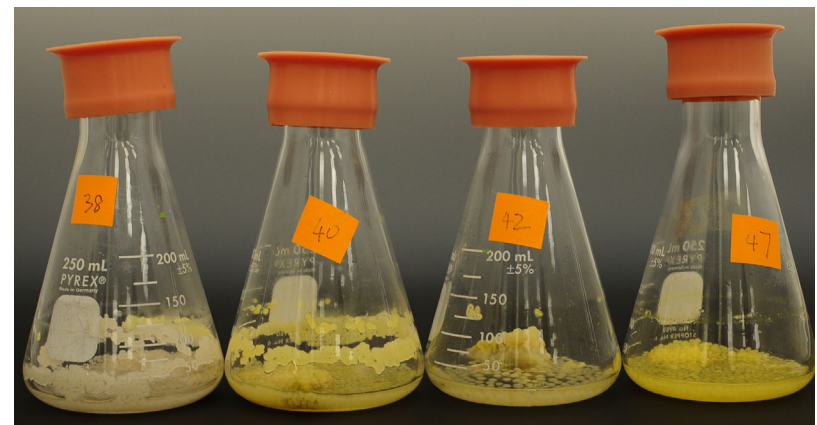
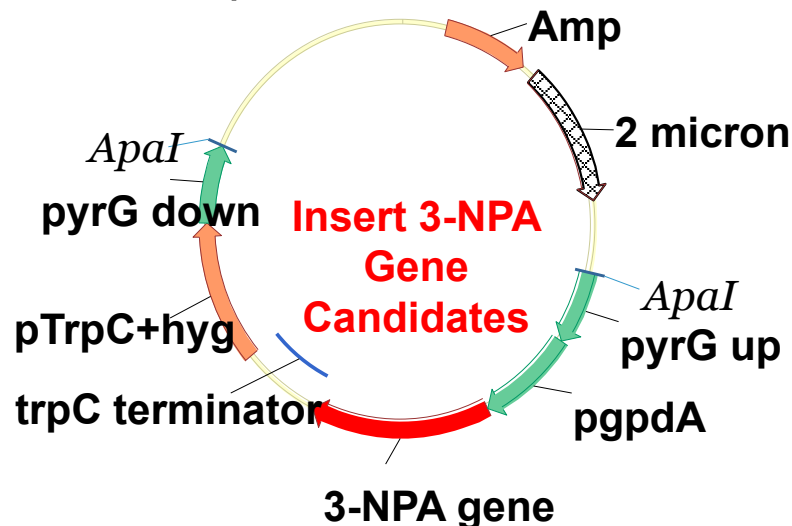
4. Transform competent *E. coli* with Gibson Assembly Reaction



5. Screen correct transformants by PCR

3NPA: Expression analysis

- Selected 9 first tier gene candidates from the list
- Built an overexpression vector for each of the 9 genes, transformed *A. niger*, picked clones, cultured, analyzed for 3-NPA
- None of them produced detectable 3-NPA
- Possible Issues:
 - ◆ May be a multi-subunit enzyme
 - ◆ May require more than one enzyme in the pathway
 - ◆ Poor expression or inactive enzyme produced in the different host
 - ◆ Metalloproteins are often difficult to express in an active state



3NPA

Streptomyces FzmM homologs in *A. oryzae*

- Blast search of *Streptomyces* FzmM against *A. oryzae* transcriptome

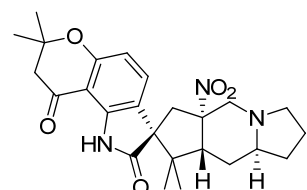
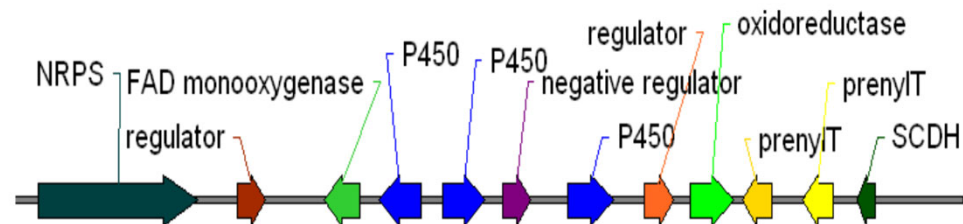
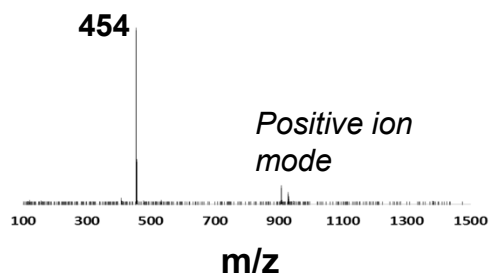
#query_acc	target_acc	avg_per_id	min_Evalue	query_match_range	target_match_range	pct_query_len	pct_target_len	max_pct_len
WP_053787792.1	TRINITY_GG_373_c290_g1_i1	52.56	2.110000e-171	19-578	5-1663	86.96	93.78	93.78
→ WP_053787792.1	TRINITY_GG_373_c533_g1_i1	85.71	7.950000e-09	590-617	135-218	4.35	37.84	37.84
WP_053787792.1	TRINITY_GG_79_c221_g1_i1	28.87	1.000000e+00	425-520	100-312	14.91	67.62	67.62

- TRINITY_GG_373_c290_g1 was the second most upregulated transcript, but was not selected due to poor annotation

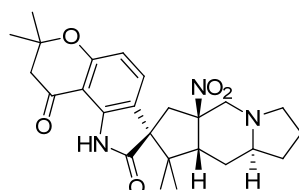
Gene	Log2 Fold Change (B vs A)			Sum	Annotation		uniprot_annotation
	Day2	Day3	Day4		contig_#	contig_annotation	
TRINITY_GG_217_c13_g1	5.8	10.4	11.9	28.1	13476	XP_002375347.1 conserved hypothetical protein	Probable tautomerase YrdN OS=Bacillus subtilis
→ TRINITY_GG_373_c290_g1	5.7	11.1	11.0	27.8	12120	XP_001825020.1 hypothetical protein AOR_1_48	
TRINITY_GG_373_c125_g1	6.1	9.7	11.7	27.5	4038	KOC15298.1 hypothetical protein AFLA70_338g0	Uncharacterized transporter C36.02c OS=Schizos
TRINITY_GG_373_c515_g1	8.7	8.7	9.7	27.1	11969	KOC15296.1 ankyrin repeat containing protein [A	
TRINITY_GG_373_c385_g1	7.7	9.1	10.1	27.0	11969	KOC15296.1 ankyrin repeat containing protein [A	Ankyrin-repeat domain containg transcription co
TRINITY_GG_373_c96_g1	5.7	9.9	11.2	26.8	4038	KOC15298.1 hypothetical protein AFLA70_338g0	Multidrug transporter TPO1_1 OS=Candida glabr
TRINITY_GG_373_c533_g1	5.7	11.0	9.9	26.5	12120	XP_001825020.1 hypothetical protein AOR_1_48	
TRINITY_GG_43_c87_g1	2.8	9.7	11.0	23.6	16905	XP_001823268.1 hypothetical protein AOR_1_77	Hydroxylase/desaturase asaB OS=Aspergillus fla
TRINITY_GG_79_c97_g1	5.7	8.1	9.4	23.1	7038	XP_001827032.1 oligopeptide transporter [Asper	
TRINITY_GG_373_c488_g1	4.9	8.9	9.2	23.0	11969	KOC15296.1 ankyrin repeat containing protein [A	

Target 2: Citranalin from *Penicillium citrinum*

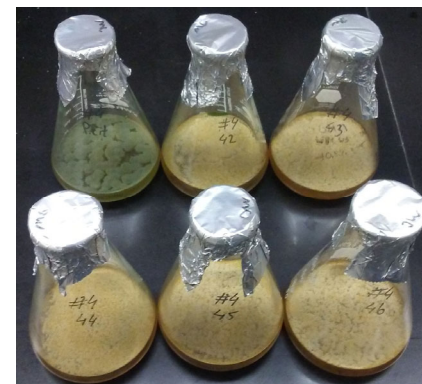
- *Penicillium citrinum* grew in culture at USC
- Identified the compounds by HPLC/UV/MS
- Citranalin is produced by a Non-Ribosomal Peptide Synthase, **NRPS, gene cluster**
- There are 17 genes present in the 53,287bps cluster, including three encoding Cytochrome P450 (oxidative) enzymes and another oxygenase
- One of the P450s is likely involved in the synthesis of the nitro group!
- Culture studies of *P. citrinum* resulted in the identification of the native citranalin isomers



Citranalin A (1)
Chemical Formula: $C_{25}H_{31}N_3O_5$
Exact Mass: 453.2264

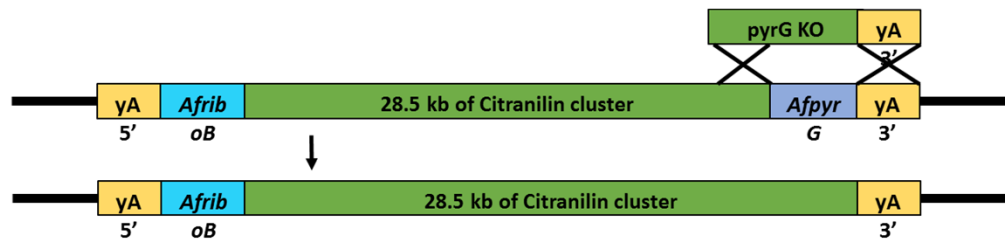


Citranalin B (2)
Chemical Formula: $C_{25}H_{31}N_3O_5$
Exact Mass: 453.2264



Target 2: Citranalin

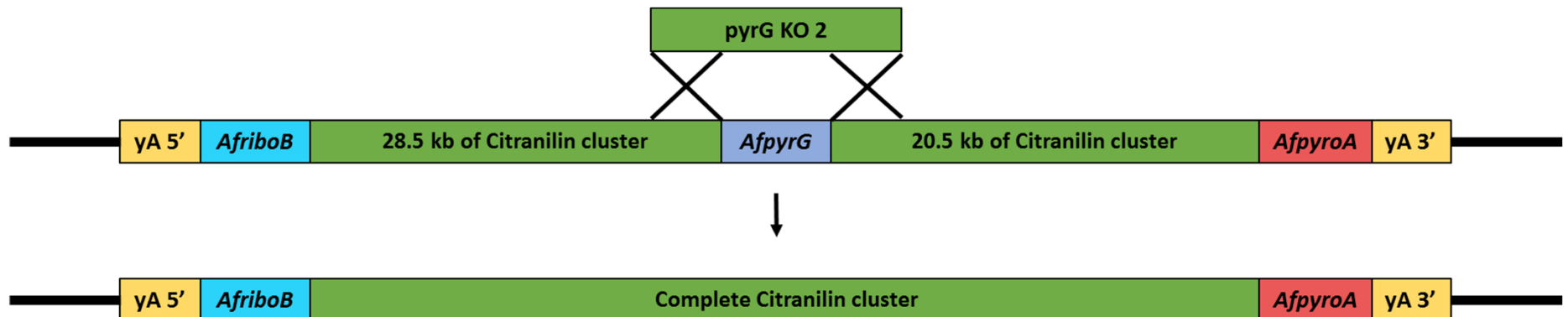
Heterologous expression of Citranalin biosynthetic cluster in *A. nidulans*



Assembly of second half



Final Assembly of 53kbp Cluster

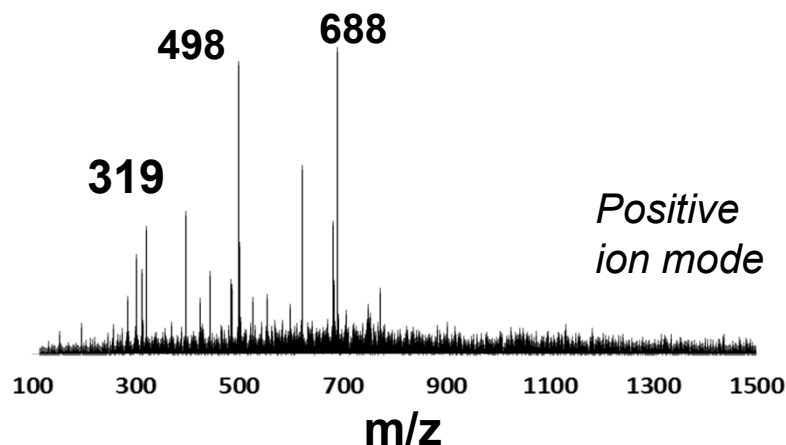


Target 2: Citrinalin

Heterologous expression of Citrinalin biosynthetic cluster in *A. nidulans*

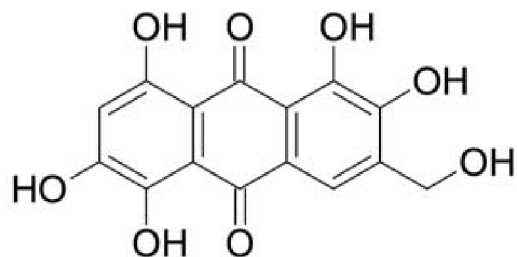
Culture Studies

- A. nidulans* + 53kbp citrinalin cluster: no citrinalin observed



Asperthecin?

MW: 318.24 Da

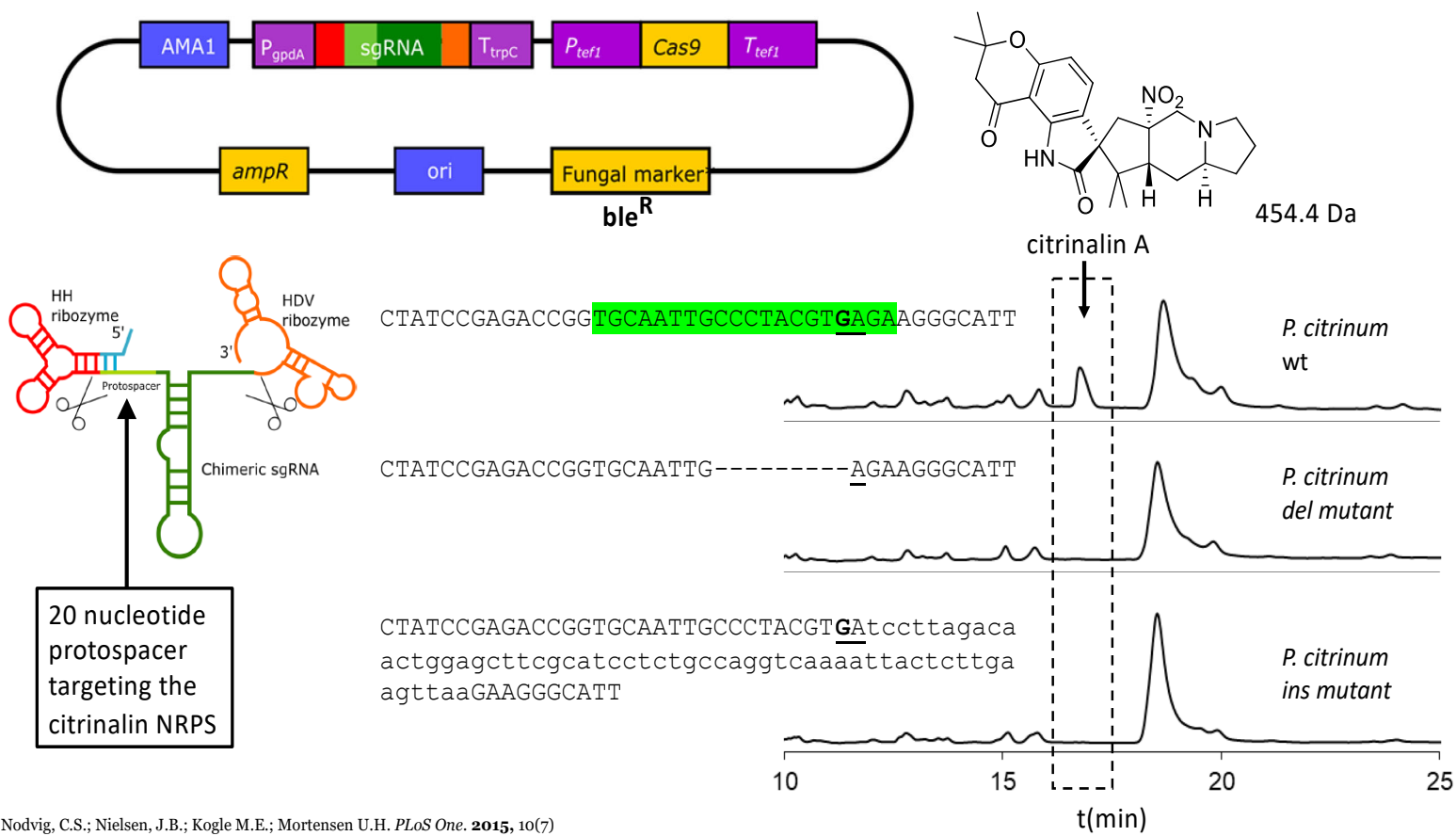


Possible issue

- Promoters may not have been activated in the *A. nidulans* background
- Drive the native promoter with *laeA* over-expression or insertion of a constitutive promoter upstream of the cluster transcription factor

Target 2: Citranalin

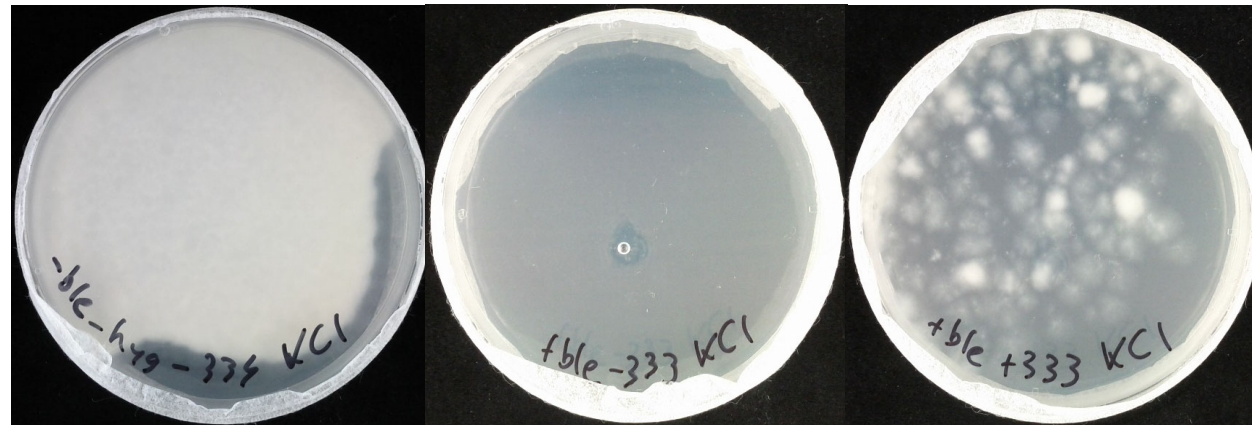
Enabling Gene Deletion Analysis with the CRISPR-Cas9 system in *P. citrinum*



Nodvig, C.S.; Nielsen, J.B.; Kogle M.E.; Mortensen U.H. *PLoS One.* **2015**, 10(7)

Target 2: Citranalin

Enabling Gene Deletion Analysis with the CRISPR-Cas9 system in *P. citrinum*

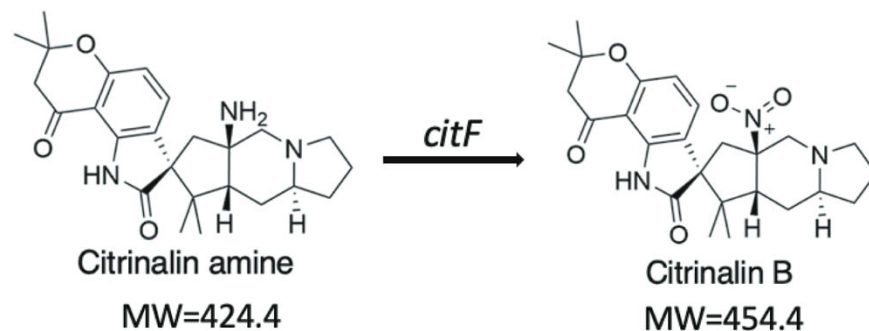
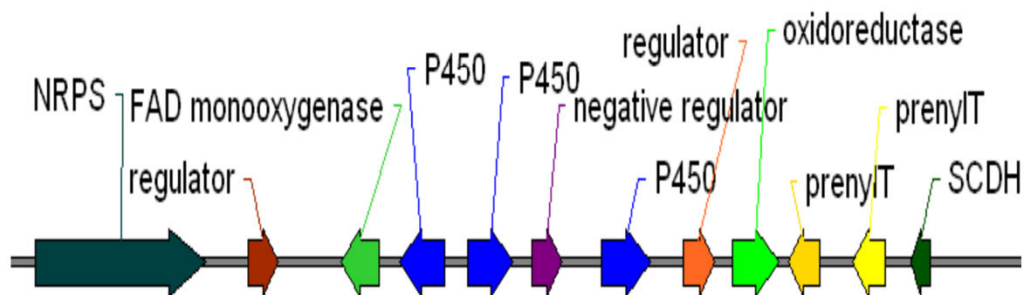
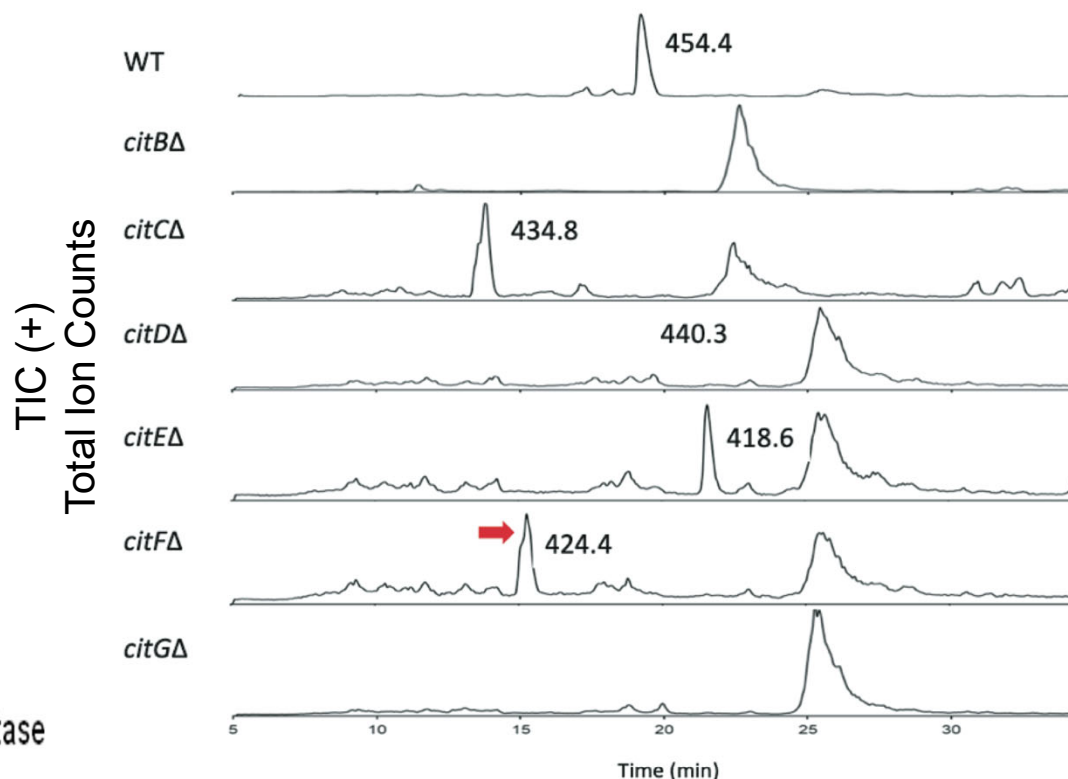


- Determined correct protoplasting conditions and osmotic stabilizer for *P. citrinum*
- Determined effective concentration of antibiotic for *P. citrinum* protoplasts
- Successfully transformed Cas9 plasmid into *P. citrinum*

Target 2: Citranalin

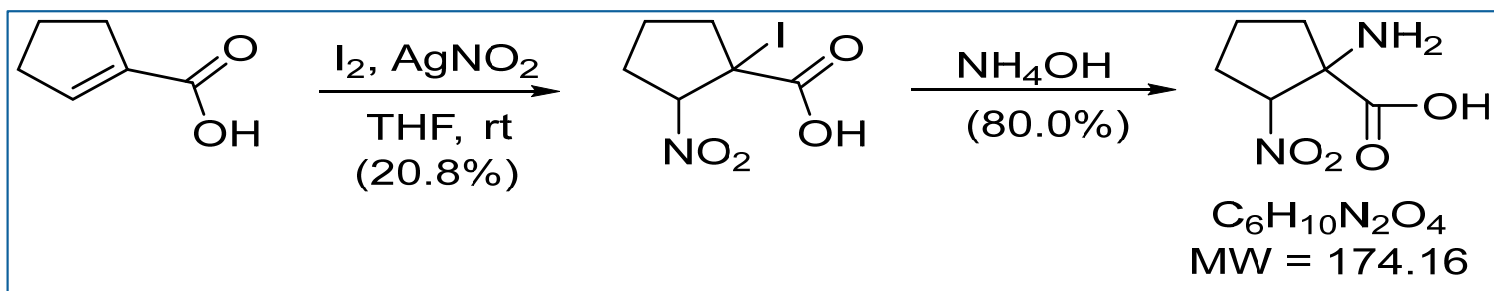
Identification of “nitration” gene

- The *citF* gene identified by deletion analysis
- Encodes the enzyme that catalyzes the oxidation of an amine to a nitro group



Target 3: 1-amino-2-nitrocyclopentanecarboxylic acid (1A2NC) from *Aspergillus wentii*

- Obtained *Aspergillus wentii* from an English fungal culture collection that was reported in the literature to produce the nitroorganic **1-amino-2-nitrocyclopentanecarboxylic acid**



- Cultured *Aspergillus wentii* under the literature conditions and investigated some variations on those conditions, but no product was detected

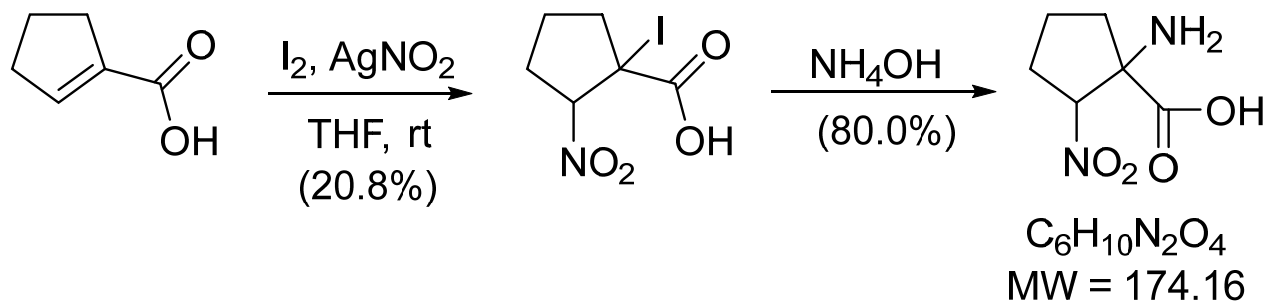
Target 3: 1A2NC from *Aspergillus wentii*

Possibilities:

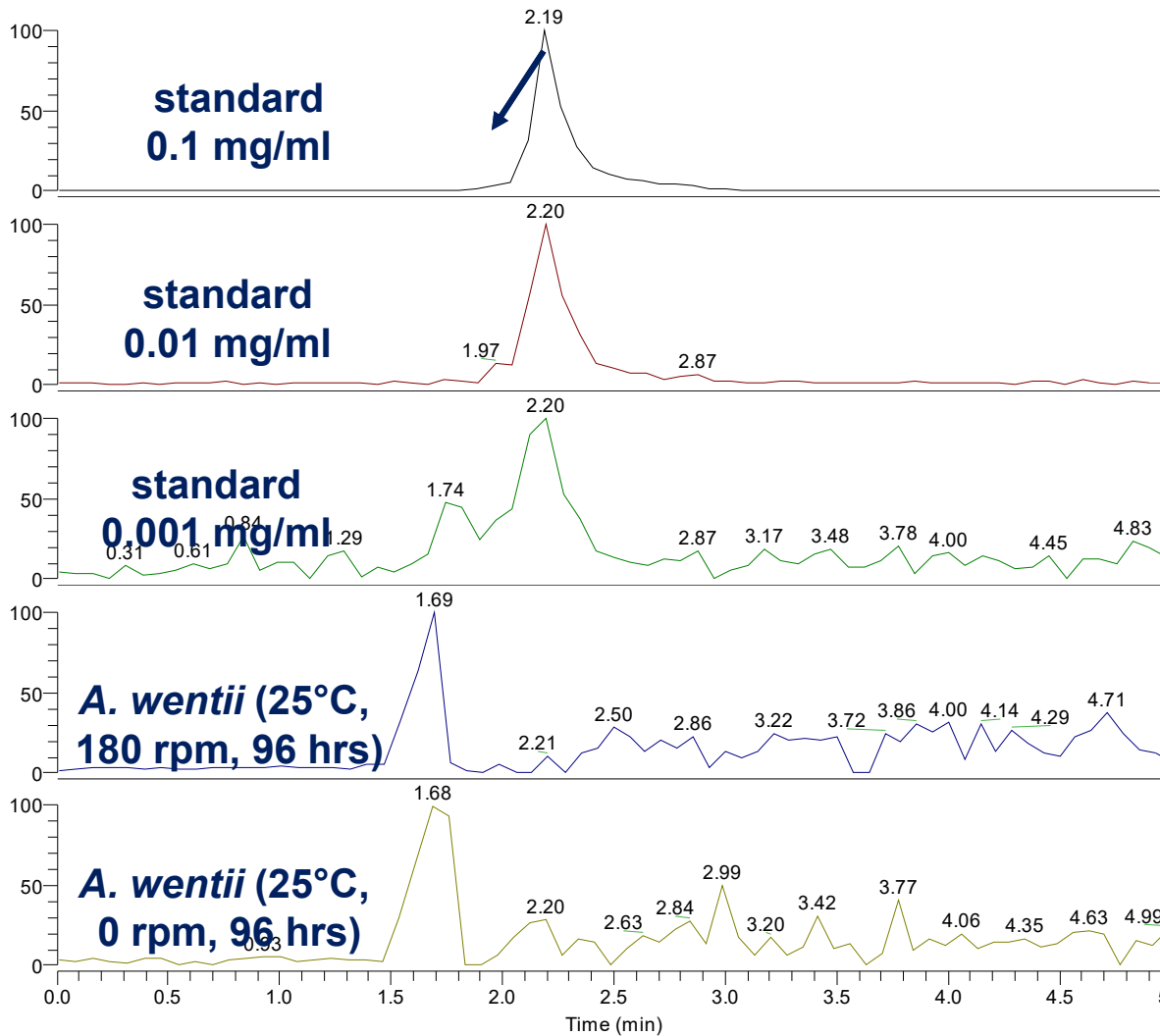
- **1-amino-2-nitrocyclopentanecarboxylic acid** is not stable during ESI ionization.
- **1-amino-2-nitrocyclopentanecarboxylic acid** do not appear in the window of our HPLC analytical condition.

Solutions:

- Chemical synthesis of **1-amino-2-nitrocyclopentanecarboxylic acid** as a compound standard.



Target 3: 1A2NC

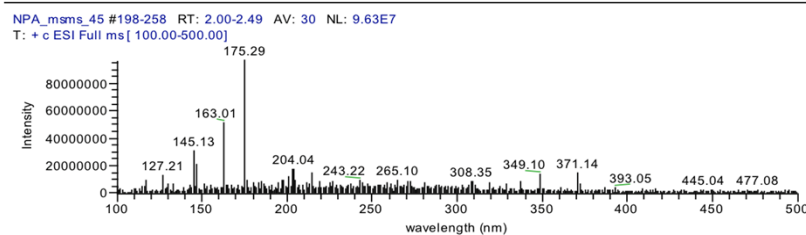


- The detection limit of **1-amino-2-nitrocyclopentanecarboxylic acid** is around 1 ppm.
- Therefore, *if A. wentii* produced **1-amino-2-nitrocyclopentanecarboxylic acid** in the above growth conditions, the yield is much less than 1 ppm.

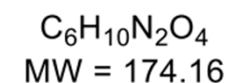
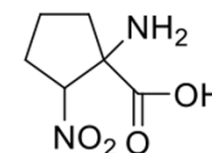
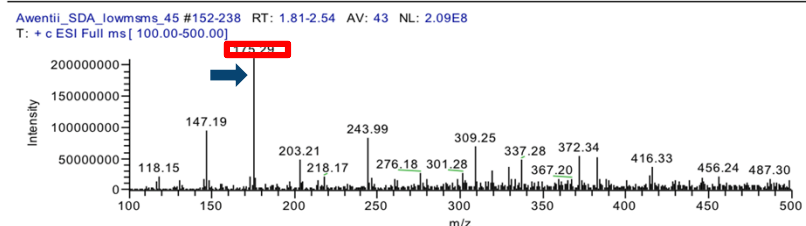
Target 3: 1A2NC

- Determined better sporulation conditions
- Screened many additional growth conditions
- Found that SDA media (Sabouraud Dextrose Agar) promoted the production of 1A2NC
- Sequenced the genome of *A. wentii* by Illumina

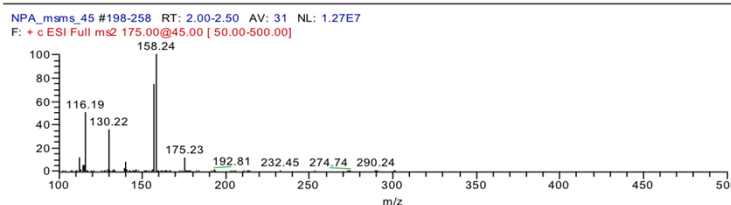
Synthesized Standard
MS



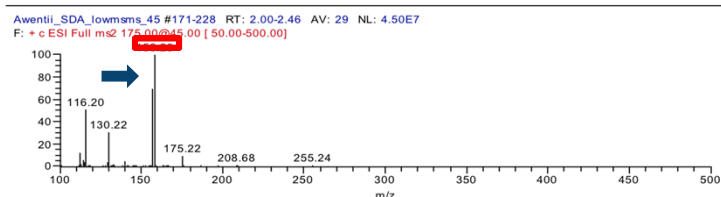
SDA Media
MS



Synthesized Standard
MS-MS



SDA Media
MS-MS

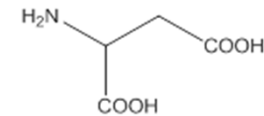


Target 3: 1A2NC

- Implement CRISPR system in *A. wentii*
- Perform deletion analysis of candidate nitroorganic biosynthetic genes
- Over-express genes identified as necessary for 1A2NC biosynthesis in *A. niger* to demonstrate function
- More than one gene may be necessary

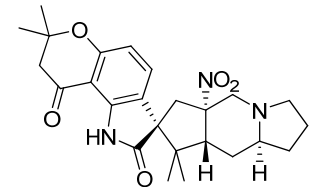
Summary

Target 1: 3NPA



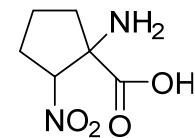
- ❖ Developed large list of candidate genes from comparative transcriptomics
- ❖ Continue working on gene deletion conditions in *A. oryzae* to identify all genes necessary for 3NPA biosynthesis
- ❖ Expression of candidate genes in *A. niger* but no successful 3NPA production by single gene deletion to date

Target 2: Citranalin



- ❖ Citranalin amine intermediate has been isolated and characterized by NMR and mass spectrometry
- ❖ Developed transformation system and CRISPR for *P. citrinum*
- ❖ Knocked out genes in the Citranalin NRPS gene cluster and identified the oxidoreductase involved in nitration
- ❖ Mutant studies demonstrate the citF gene encodes an enzyme that catalyzes the production of nitro groups in fungi.

Target 3: 1A2NC



- ❖ Identified culture conditions to produce the compound in *A. wentii*
- ❖ Sequenced the genome of the fungus

Future Plan

- ❖ Utilize protein/DNA homology searches BLAST to identify related genes
- ❖ Expression libraries and high throughput screening of substrates
- ❖ Need equally high throughput derivatization GC-MS and LC-MS metabolomics to identify products
- ❖ Evolution of enzymes (see Nobel Prize winner Francis Arnold, for example)
- ❖ Ultimately as pathways to energetic materials near feasibility, need a Design-Build-Test-Learn pipeline to design and implement EM pathways and develop the processes
- ❖ Department of Energy, BioEnergy Technologies Office (DOE-BETO) funded Agile BioFoundry is one example of an open (accessible) BioFoundry

List of Scientific Publications & Presentations

- Deng S, KR Pomraning, P Bohuskyi, and JK Magnuson. 2018. “Draft Genome Sequence of *Aspergillus oryzae* ATCC 12892”. *Genome Announcements*. 6:e00251-18; doi:10.1128/genomeA.00251-18
- Establishing CRISPR in *P. citrinium* and using it to identify the nitration enzyme in citrinalin synthesis, manuscript in preparation
- The third publication we anticipate will be identification of the 3-NPA production genes by deletion analysis in *A. oryzae* and over-expression in *A. niger*. Target journal: *Applied Microbiology and Biotechnology*, *Applied and Environmental Microbiology* or equivalent IF
- Deng, S., Wang, C., and Magnuson J.K. 2019 Utilizing CRISPR technique to identify nitroorganics biosynthesis enzymes in fungi. SERDP & ESTCP Symposium. Washington, DC
- Magnuson J.K. 2018 Synthetic Biology and Energetic Materials - Technical Interchange Meeting. The Hague, Netherlands
- Deng, S., Dijk, J.v., Chiang, Y., Wang, C., and Magnuson J.K. 2017 Approaches to identify the biosynthetic routes to nitroorganics. SERDP & ESTCP Symposium. Washington, DC

Acknowledgements

PNNL

- Jon Magnuson
- Shuang Deng
- Marie Swita
- Kyle Pomraning
- Ziyu Dai
- Pavlo Bohutskyi

USC

- Clay Wang, co-PI
- Yi-ming Chiang
- Jan VanDijk (Zymergen)
- David Sherman (U. Michigan, Clay's collaborator)

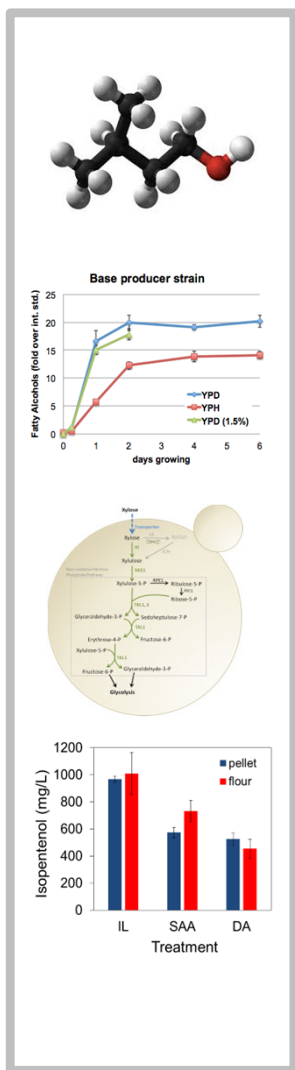
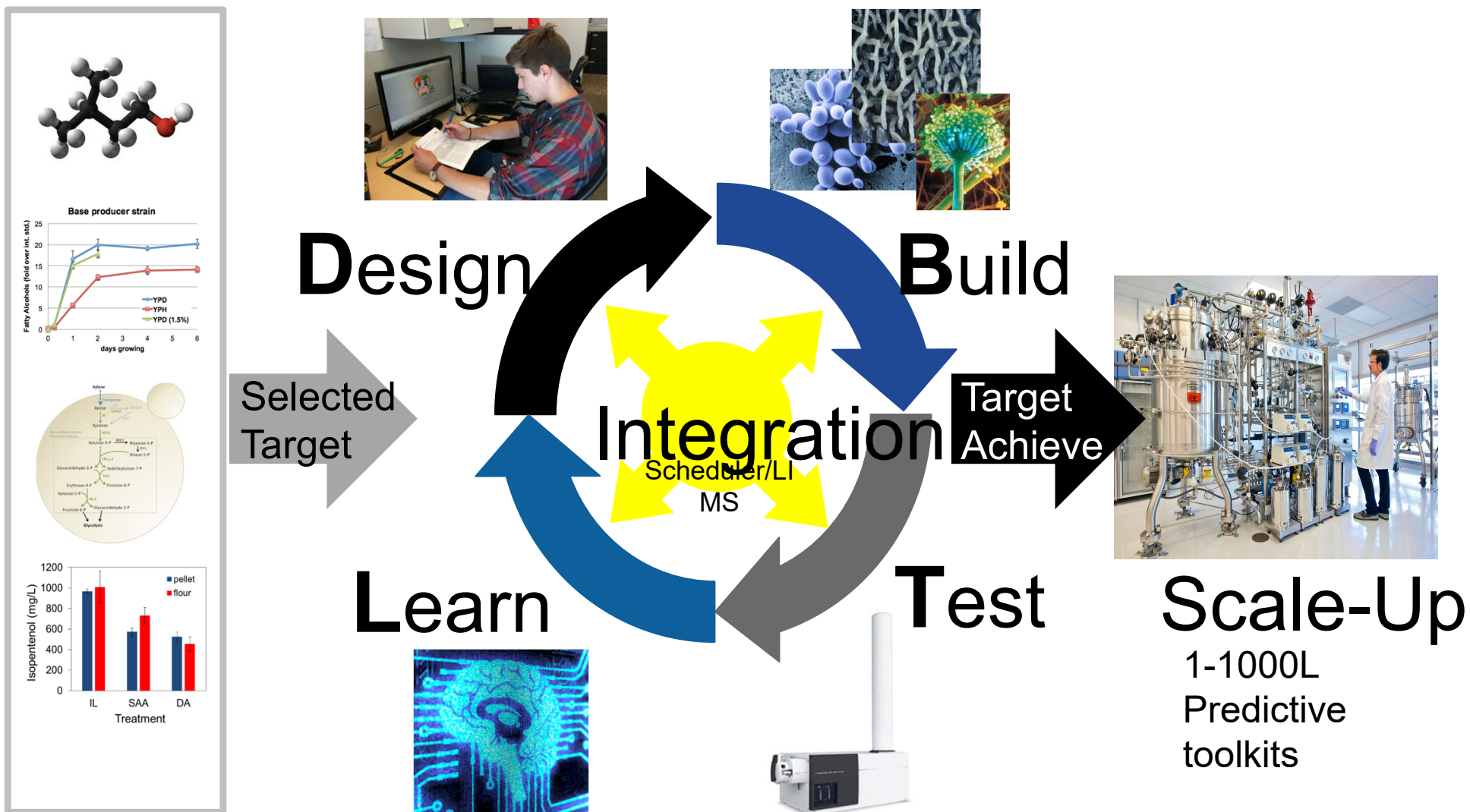
SERDP

- Robin Nissan
- Braxton Lewis
- Kristen Lau
- WP team
- Reviewers

Extra Slides on Relevant Current Project

The Agile BioFoundry Approach

Funded by DOE, EERE, BioEnergy Technologies Office (BETO)



ABF capabilities

- **Targets:**

- Commodity chemicals and sustainable aviation fuel (SAF) precursors that can reduce reliance on petroleum
- Current target classes: organic acids, SAF lipid pathway molecules

- **Primary hosts:**

- Three primary host organisms within the ABF
 1. *Pseudomonas putida*, *Calcicellulosiruptor* sp.: makes organic acids, polyhydroxyalkanoates, ethanol
 2. *Rhodospiridium toruloides*, oleaginous yeast, terpenes, lipids, fatty alcohols
 3. *Aspergillus niger*, *A. pseudoterreus*, *Lipomyces starkeyi*, acid-tolerant, pH 1 to 3, organic acids, e.g., 3-hydroxypropionic acid, aconitic, itaconic, malic
- Host Onboarding: currently have 20+ species in our repertoire

- **Capabilities**

- Broad DESIGN & BUILD capabilities; synthetic biology skills in diverse organisms
- Deep and broad TEST capabilities: transcriptomics, proteomics, metabolomics, lipidomics
- Broad LEARN capabilities: ^{13}C -MFA, FBA, machine learning (ART, ANN), kinetic models (Bayesian MCA), TEA/LCA