Determining impacts of natural product biosynthetic pathways and environmental conditions on *Nocardiopsis* metabolomes by isotopic ratio outlier analysis (IROA) global metabolomics

Amy L. Lane[1], Elle D. James[1], Felice de Jong[2], Chris Beecher[2,3]

[1] University of North Florida, Jacksonville, FL, USA, [2] IROA Technologies LLC, Bolton, MA, USA , [3] University of Florida, Gainsville, FL, USA

ABSTRACT

The *Nocardiopsis* species actinomycete bacteria harbor biosynthetic pathways for bioactive compounds including toxins, antimicrobial agents, and anticancer substances. Expression of these pathways is dependent upon the diverse conditions in which actinomycetes prevail. It was hypothesized that iron limitation may induce *Nocardiopsis* to upregulate the production of siderophore secondary metabolites (e.g. desferrioxamine or other previously unidentified siderophores), since these types of molecules are known to aid bacteria in sequestering iron.

INTRODUCTION

Over 65% of known microbial secondary metabolites are produced by actinomycetes. Actinomycetes, a group of Gram-positive bacteria, are recognized producers of a wealth of natural products with proven clinical use for treatment of microbial infections, cancer, and other maladies. *Nocardiopsis* sp. prevail in a variety of environments and produce a vast variety of bioactive compounds, including many antibiotics that may aid as signaling molecules to protect their existence under different conditions, stressors and against competing microorganisms.[1] Natural products inspire organic syntheses, are valuable as chemical probes, and offer applications in agriculture and other industries.[2]

Metabolomics-guided approach to investigate metabolites produced by *Nocardiopsis*.
Actinomycete genomes typically harbor over a dozen biosynthetic pathways. However, the majority are not expressed under typical laboratory fermentation conditions. Experimental approaches such as the heterologous expression of biosynthetic genes and fermentation environment manipulation have shown promise for increasing expression of biosynthetic genes. Each of these approaches is reliant upon methods for evaluation of metabolic changes in response to experimental manipulations.

Herein, we implement Isotope Ratio Outlier Analysis (IROA)[3,4] global LC/MS metabolomics to evaluate changes in the metabolome of a model *Nocardiopsis* species resulting from treatment with an iron chelating reagent, bipyridyl. (Fig. 1).

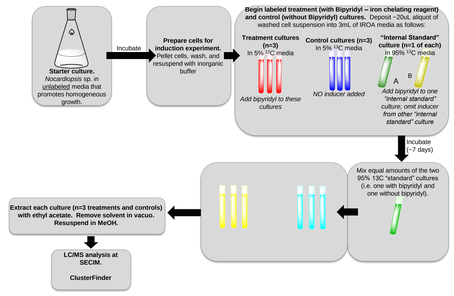


Figure 1. Outline of approach for IROA LC/MS evaluation of metabolic differences between *Nocardiopsis* treatment with iron chelating agent bipyridyl and control cultures without bipyridyl.
SECIM= Southeastern Center of Integrated Metabolomics

Our IROA LC/MS approach (Fig. 1) revealed several metabolites unique to the *Nocardiopsis* sp. The results support IROA as a sensitive technique for metabolite profiling of chemically rich actinomycetes and suggest the broad applicability of this technique toward probing metabolic responses of actinomycetes to genetic and/or environmental manipulation.

THEORY AND DISCUSSION

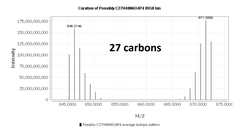
Bacteria utilize siderophores, low molecular weight binding iron-binding complexes, to acquire and transport iron for important cellular processes. The effect of iron limitation on siderophore secondary metabolite production was examined using the IROA Basic Protocol. An "Internal Standard", which mirrored the metabolome of the Control and Experimental samples, was created by mixing equal amounts of two 95% ^{13}C *Nocardiopsis* sp. cultures; with and without bipyridyl iron-chelation treatment. Control and Experimental samples, labeled with 5% ^{13}C, were individually mixed with Internal Standard and analyzed as a pooled sample so that variability and ion suppression were accounted for, and fewer samples were required to be analyzed.

The approach enables software algorithms to remove false data such as noise or artifactual peaks, identified by the absence of isotopic signatures, allowing for a very dramatic reduction in data size. Peaks of biological origin are perfectly paired: each IROA envelope is half Internal Standard and half Control or Experimental. The ratio of these halves quantitates the response relative to the Internal Standard. For example, the IROA envelope for an unknown 27 carbon molecule has two halves separated by a 27 carbon mass difference (Fig. 2). The sum of the base peaks and their associated isotopic peaks are used to accurately calculate their respective areas and ratios. The number of carbons calculated for each IROA peak together with accurate mass for either of the monoisotopic peaks makes it possible to unambiguously determine the empirical formula for masses below 400.

RESULTS

Quantitation of changes in actinomycetes metabolomes
This study describes the use of the IROA LC/MS, an analytical and bioinformatic metabolomics approach (Fig. 1), applied to the understanding of metabolic differences between *Nocardiopsis* sp. with and without bipyridyl treatment used to induce iron limitation. The 95% U-^{13}C IROA peaks in the labeled IROA Internal Standard were used by ClusterFinder™ software to identify and quantitate associated 5% U-^{13}C-labeled peaks in both *Nocardiopsis* sp. Control and Experimental samples (Fig. 2). 685 IROA peaks were identified; 400 named and 385 formula determined between positive mode (344 compounds) and negative mode (341). 260 compounds were considered significant, including metabolites that were found in either the iron-limited samples or untreated samples but not both (Fig. 3).

Figure 2. Representative IROA peaks.
27 carbon molecule curated using ClusterFinder.



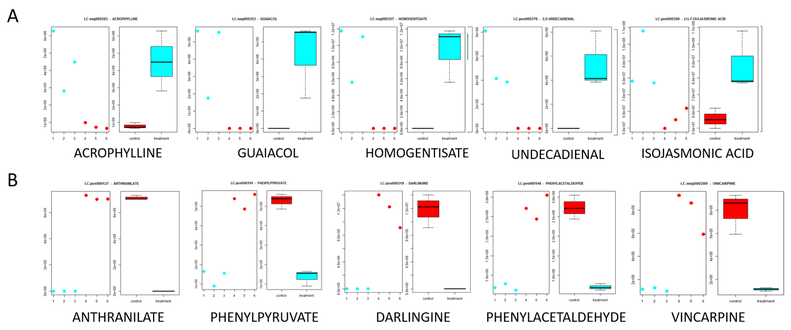


Figure 3. Box plots of statistically significant compounds. (A) Metabolites upregulated in iron-limited *Nocardiopsis* sp. but not in Control samples; (B) Metabolites upregulated in Control samples relative to iron-limited *Nocardiopsis* sp.

Separation of metabolic response by Random Forest (RF), Self-organizing Maps (SOM), and Non-negative Matrix Factorization (NMF)
Bioinformatics techniques including: Principal Component Analysis (PCA), Random Forest (RF), Self-organizing Maps (SOM), Non-negative Matrix Factorization (NMF) all yielded clean separations of each group. The confusion table (Fig. 4A) shows that RF is able to perfectly discriminate between the bipyridyl treated and Control samples using both the full data set (685 compounds) and pruned dataset (281 compounds). SOM (Fig. 4B), an unsupervised algorithm mapping similar compounds, separated the top compounds into 23 nodes. NMF assumes that every dataset (matrix of compounds by samples) may be seen as the sum of k submatrices. NMF was run with k values from 2 to 6. k values 1-5 achieved statistically significant fits, and all had at least one plot similar to that shown in Fig. 4C (NMF2; p value of $1.7X10^{-3}$).

A

Confusion table [full dataset]			685 compounds
	control	treatment	class.error
control	3	0	0
treatment	0	3	0
Confusion table [pruned dataset]			281 compounds
control	3	0	0
treatment	0	3	0

B C

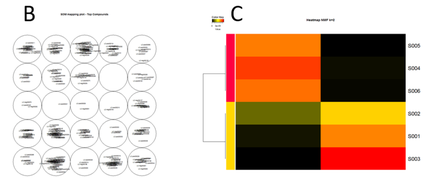


Figure 4. Random Forest (RF), Self-organizing Maps (SOM), and Non-negative Matrix Factorization (NMF) analysis of statistically significant compounds. (A) RF confusion table; (B) SOM map plot of 24 nodes; (C) NMF plot of k=2 value.

Interrogation of up- and down-regulated Compounds
Results from the impact of iron limitation on the transcriptome of *Caulobacter crescentus*, a freshwater α-proteobacterium, revealed altered expression of genes involved in energy metabolism, oxidative stress, and other regulatory processes.[5] One such gene, 4-hydroxyphenylpyruvate dioxygenase (HPPD) is upregulated. HPPD catalyzes the conversion of 4-hydroxyphenylpyruvate into homogentisate, and phenylpyruvate to 2-hydroxyphenylacetate, consistent with our findings of increased levels of homogentisate (Fig. 3A) and decreased levels of phenylpyruvate (Fig. 3B) in treated cultures.

A SOM/NMF correlation map of compounds upregulated in iron-limited *Nocardiopsis* sp. relative to untreated Controls was created (Fig. 5) to determine similarity. Upregulated compounds clearly partitioned into NMF2.1 and the majority were members of two SOM nodes, depicted in blue and green. NMR and structure elucidation will be performed on NMF2.1 and SOM compound members to confirm their identify as compounds of interest.

Iron-limited upregulated compounds	compound ID	NMF2.1
2,4-HEPTADIENAL	LC-pos000367	0.9172
9-OXONONANOIC ACID	LC-pos000373	0.9263
UROCANATE	LC-pos000068	0.9634
4-HYDROXYBENZOATE	LC-neg000354	1.0000
GUAIACOL	LC-neg000353	1.0000
ALPHA-METHYLSTYRENE	LC-neg000343	1.0000
Possibly C16H30N2O5 (329.2 @9.9)	LC-neg000365	1.0000
Possibly C17H31NO3 (298.2 @10.7)	LC-pos000382	1.0000
HOMOGENTISATE	LC-neg000357	1.0000
O-DECANOYL-L-CARNITINE	LC-neg000364	1.0000
N-HEXADECANOYL-SERINE	LC-neg000367	1.0000
ETHYL 2-METHYLBUT-2-ENOATE	LC-pos000125	1.0000

Figure 5. NMF and SOM correlation map for cyn upregulated compounds. SOM nodes, blue and green.

METHODS

Experimental design: *Nocardiopsis* sp. were maintained in media that promoted homogenous growth. Cells were washed, and resuspended with inorganic butter. 20uL aliquots of washed cell suspension were added to 3 mL of isotopically defined (95% or 5% U-^{13}C), until fully transformed to the isotopic balance of the media. Bipyridyl was added to Treatment cultures (5% ^{13}C ; n=3) and "Standard A" culture (95% ^{13}C ; n=1) but not Control Cultures (5% ^{13}C ; n=3) or "Standard B" culture (95% ^{13}C ; n=1) and incubated for 7 days (Fig. 1). Aliquots of 5% ^{13}C liquid cultures (Control and Experimental) were individually pooled with an Internal Standard Control made from a 1:1 mix of 95% ^{13}C Standard A: Standard B, and cultures were extracted with ethyl acetate. Samples were stored at -20 °C .
LC-MS: Pooled (95%/5%) sample extracts were analyzed by reverse phase UHPLC-HRMS using an Ace Excel PFP-C 18 2.1x100 mm, 2.0 µm column. The UHPLC pump was a Dionex Ultimate 3000, and the mass spectrometer was a Thermo Fisher Q-Exactive Orbitrap. Solvent A was 0.1% formic acid in water, and solvent B was acetonitrile. The collected data were initially processed by Xcalibur Workstation software (version 3.0).
Data processing and analysis: The Thermo RAW datafiles were converted to mzXML data files which were imported directly into the IROA ClusterFinder Software Program for peak picking and compound identification and the IROA portal for statistical analysis.

REFERENCES

1. Bennur T, Ravi Kumar A, Zinjarde SS, Javdekar V, " *Nocardiopsis* species: a potential source of bioactive compounds", *J Appl Microbiol* 2016, 120(1):1-16. doi:10.1111/jam.12950.
2. Walsh CT, Fischbach MA, "Natural Products Version 2.0: Connecting Genes to Molecules", *J Am Chem Soc* 2010, 132(8): 2469-2493. doi: 10.1021/ja909118a
3. de Jong F, Beecher C, "Addressing the current bottlenecks of metabolomics: Isotopic Ratio Outlier Analysis (IROA®), an isotopic-labeling technique for accurate biochemical profiling", *Bioanalysis* 2012, 4(18), 2303-14.
4. Stupp GS, Clendinen CS, Ajredini R, Szewc MA, Garrett T, Menger RF, Yost RA, Beecher C, Edison AS. "Isotopic Ratio Outlier Analysis Global Metabolomics of Caenorhabditis elegans" *Analytical Chemistry* 2013, 85(24), 11858-11865. doi: 10.1021/ac4025413.
5. da Silva Neto JF, Lourenco RF, Marques MV, "Global transcriptional response of *Caulobacter crescentus* to iron availability", *BMC Genomics* 2013, 14:549. DOI: 10.1186/1471-2164-14-549.

SUMMARY

The IROA analytical and bioinformatic metabolomics approach was used to interrogate samples treated with and without iron-chelation to look for potential *Nocardiopsis* sp. natural products. Figure 4 shows siderospore or stress response secondary metabolites candidate compounds that were upregulated in response to iron-limitation (bipyridyl treatment). Compounds present in Control samples (without bipyridyl treatment) but absent in Experimental iron-chelated samples suggest that these molecules are assembled via enzymes that require Fe as a cofactor and/or are downregulated during iron limitation to conserve resources.

Acknowledgements
We are thankful for support from a Southeast Center for Integrated Metabolomics (SECIM) Pilot and Feasibility grant for this project. 5% and 95% U-^{13}C labeled metabolites were produced especially for IROA Technologies by Cambridge Isotope Laboratories, Inc. (CIL)