Microascaceae
Cephalothecaceae

Explorative Analysis of Biodiversity in SURF Fungal Species through Sanger Sequencing

Tada Vargas[1], Beth Reman[2], Shane Sarver[2]
Math and Science Department, Oglala Lakota College
Center for Conservation Biological Research, Black Hills State University

Introduction

Located in Lead South Dakota, the formally known Homestake Mine, was established in 1876. It was in operation for approximately 125 years, officially ceasing mining operation in 2002. Due to its depth beneath ground, scientific experiments have been conducted since the mid-1960. Currently, it is now called SURF, Sanford Underground Research Facility and is home to various scientific experiments. This project explores the biodiversity and genetic variability in fungal species collected from SURF. In the spring of 2016, 78 samples were collected from various underground levels. Samples were cultured and isolated, creating 343 isolate samples. Isolates were lysed and diluted, to amplify DNA, using the ThermoScientific Phire Plant Direct PCR kit. Subsequently, 311 isolates were selected for Sanger Sequencing, producing 199 informative genomic sequences. Samples were assembled and identified using genetic softwares: Sequencher and NCBI BLAST. Species were analyzed through CLC Genomics Workbench to create phylogenetic trees and cladograms.

Samples Collection

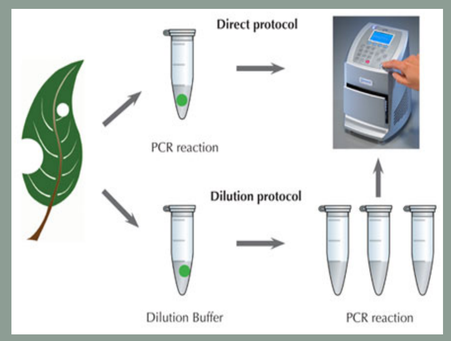
In March and April of 2016, 78 fungal samples were collected on the 800, 1700, 4100, and 485 levels of SURF.
Samples were cultured, isolated, frozen, and stored for further genetic analysis.

PC: Bethany Reman

Methods

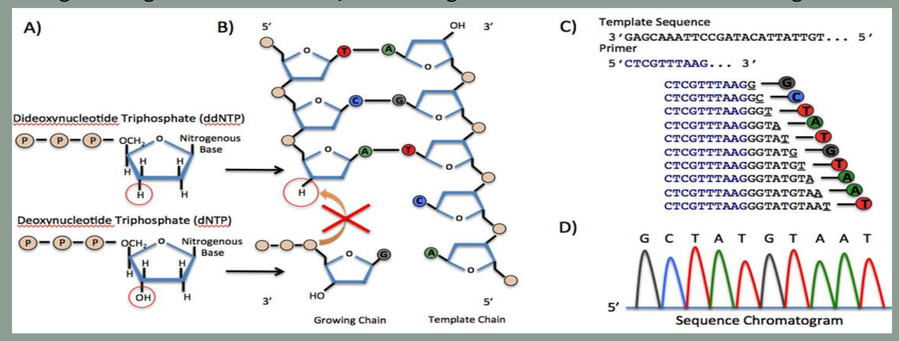
Phire Plant Direct PCR

Direct PCR Amplification of ITS1 and ITS4 primers, were done on 343 isolates, using the Phire protocol parameters.



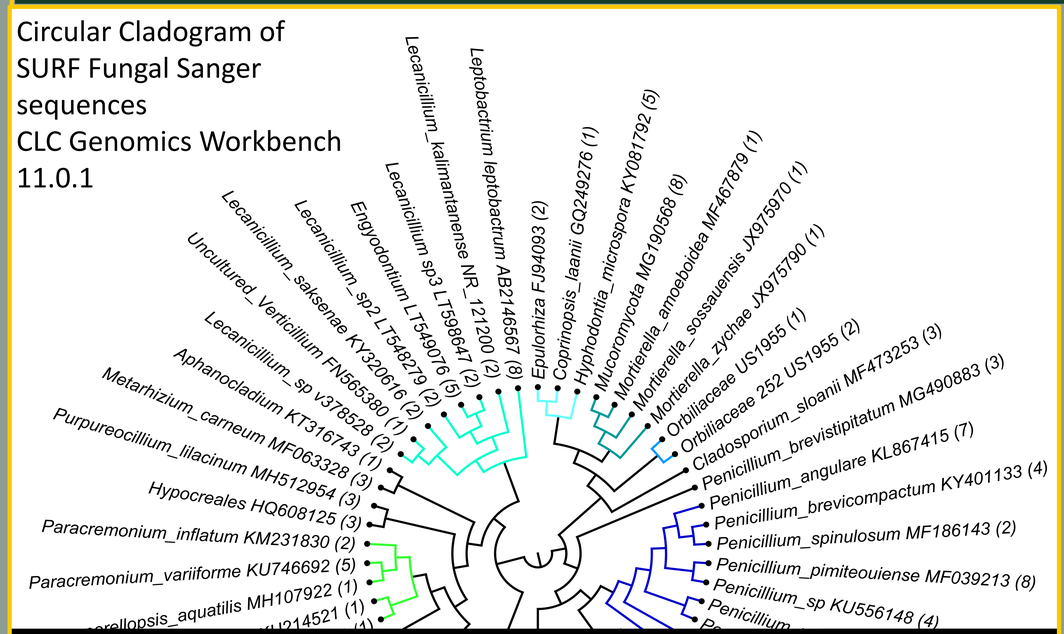
Sanger Sequencing

Applied Biosystems Sanger sequencing, using BigDye terminator v1.1, was done on 311 Direct PCR products. Successful sample sequences were assembled, using Sequencher DNA Sequnce Analysis Software and identified using NIH NCBI BLAST. A total of 199 contigs were generated and analyzed through CLC Genomics to create a cladogram.



Results

Circular Cladogram of SURF Fungal Sanger sequences
CLC Genomics Workbench 11.0.1



From 343 SURF fungal isolates, 311 produced Phire Direct PCR products. Of 311 samples, 199 generated successful Sanger sequences. With an identity parameter of minimum match percentage at greater than 95% and an overlap of 20 basepairs, Sequencher analysis assembled the Sanger sequences into 54 contigs. The circular cladogram represents three phyla: Ascomycota, Basidiomycota, and Mucoromycota and 34 genera within those phyla. According to NCBI, all samples matched at 97% or higher, with an exception of two, Orbiliaceae 252 and Orbiliaceae which were identified at 94 - 95% accession identity. In addition, 7 species contigs could not be identified greater than genus and three were only identified by phylum with a possible species identification.

Genus	Sample Size
Trichoderma	37
Penicillium	33
Scedosporium	18
Aspergillus	11
Lecanicillium	10

Class	Sample Size
Agaricomycetes	8
Mortierellomycetes	11
Dothideomycetes	5
Eurotiomycetes	7
Leotiomycetes	12
Orbiliomycetes	4

Family	Sample Size
Clavicipitaceae	3
Cordycipitaceae	13
Hypocreomycetidae	8
Nectriaceae	10
Ophiocordycipitaceae	3
Microascaceae	2
Cephalothecaceae	1

Conclusion

Phire Plant Direct PCR produced 311 quality product for sequencing. Samples that did not amplify could be a result of low quality DNA during the mechanical cell lysis stage. Sanger sequencing generated 199 usable genomic data for Sequencher analysis. A majority of successful PCR was single stranded for ITS1. ITS4 appeared to provide low quality sequencing information. Incomplete removal of EtOH could have affected amplification resulting in dirty base pair calls. Trichoderma and Aspergillus contigs displayed genetic variability, which will be investigated with an identity percentage of greater than 99%. Seven branches on the cladogram didn't appeared align correctly with their respective taxonomy. This issue could be a result of inaccurate accession identification in NCBI and needs to be further investigation.

Future Work

Samples that were unsuccessful for both ITS1 and ITS4 will be undergo Phire Direct PCR again to remove possible technician error. Samples that were unsuccessfully amplified for ITS1 will be re-sequenced under more careful conditions. Samples that displayed genetic variation in sequencher will be re-evaluated using an identity match of greater than 97%. Furthermore, samples that displayed vague taxonomy matches through NCBI BLAST will be further investigated using additional databases for more accurate identification. Also an outsource reference will be added to FASTA files so address alignment issues discovered in phylogentic trees. Next Gene. Sequencing will be done on samples that were only 94-95% identity matched.

Work Cited

Lesko, K. T. "The Sanford Underground Research Facility at Homestake (SURF)." *Physics Procedia*. Vol. 61. Elsevier B.V., 2015. 542–551. *Physics Procedia*. Web.
Waddell, E. J. et al. "Phylogenetic Evidence of Noteworthy Microflora from the Subsurface of the Former Homestake Gold Mine, Lead, South Dakota." *Environmental Technology* 31.8–9 (2010): 979–991. *Environmental Technology*. Web.
Wang, Qiong, and Jim R Cole. "Comparison of Three Fugal ITS Reference Sets." *Rdp Technical Report*, 2014 (2014): n. pag. Print

Acknowledgement

This project has been funded by Black Hills State University Center for Conservation Biological Resources and South Dakota EpScor Award # IIA – 1355423. Perspectives expressed in this material may not reflect the views of the National Science Foundation. A sincere thank you to Oxana Gorbatenko for sitting with me through hours of painstaking sequencher analysis. A heart felt thank you to Shawn Garnette, who has emotionally supported me through a summer of sequencing analysis and to my children for putting their summer on hold so their mother is closer to achieving her dreams.