

# Introduction to microbiome workshop

Dec. 16-17, 2020

<https://mcic-osu.github.io/2020-12-microbiomics-workshop/>

## Instructors:

Dr. Jelmer Poelstra: [poelstra.1@osu.edu](mailto:poelstra.1@osu.edu)

Dr. Antonino Malacrino: [malacrino.1@osu.edu](mailto:malacrino.1@osu.edu)

Dr. M. Soledad Benitez P.: [benitezponce.1@osu.edu](mailto:benitezponce.1@osu.edu)

## Dataset description

**Lab goals:** Contribute to our understanding of winter cover crops effect on soil health, microbial communities and corn-soybean production

### **Field experiment objectives:**

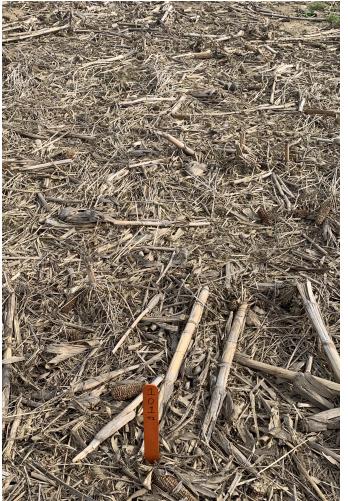
To evaluate the effect of cereal rye cover crop in subsequent soybean production (Fry Farm, Wooster, OH)

To characterize microbial community changes between cover crop establishment, cover crop termination and subsequent soybean crop

To determine the effect of timing of cereal rye termination on subsequent soybean productivity and associated microbiome

# Dataset description

No cover crop    Cover crop (Fall 2018)



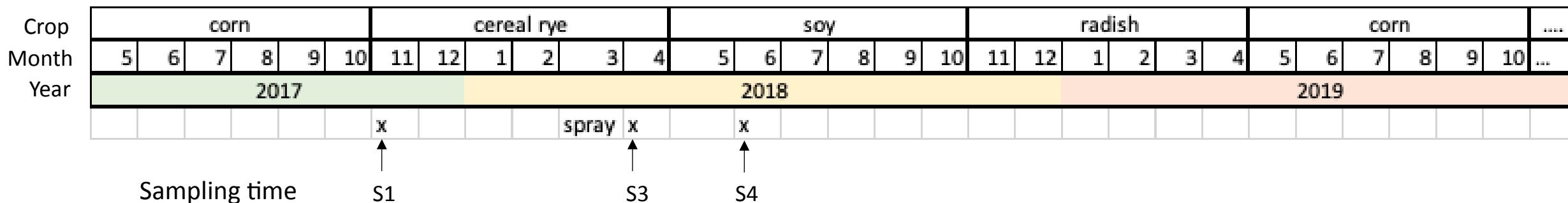
Soybean (2019)



## Treatments

T1	Control - No rye cover crop before soybean
T2	Rye CC terminated 12-14 days before soybean planting
T3	Rye CC terminated 1-0 days prior to soybean planting
T4	Rye CC terminated 5 days after soybean planting

n=6



# Amplicon sequencing data analysis

1. Sequencing quality check (performed at different stages of analyses)
2. Assign sequences to sample (demultiplexing, based on barcodes)
3. Trimming of adapters and primers
  - \* For ITS pipeline – ITS extraction (ITSx)
4. Define SV/OTU (sequence variants / operational taxonomic units)
5. Classify SV/OTUs using appropriate reference database (e.g. SILVA, UNITE)
6. Filter SV/OTUs
7. Generate Sample by OTU/SV table
8. Generate FASTA file with sequence representative of each SV/OTU
9. USE 6 and 7 for downstream analyses

Table 1 List of commonly used tools for metabarcoding data analysis

From: [Mycobiome diversity: high-throughput sequencing and identification of fungi](#)

Name	Description and link	Refs
DADA2	Amplicon sequence variant analysis pipeline	38
	<a href="https://benjneb.github.io/dada2/">https://benjneb.github.io/dada2/</a>	
Galaxy	Web-based platform, including various analytical tools	183
	<a href="https://usegalaxy.org/">https://usegalaxy.org/</a>	
LotuS	Full pipeline for amplicon data	47
	<a href="http://psbweb05.psb.ugent.be/lotus/index.html">http://psbweb05.psb.ugent.be/lotus/index.html</a>	
mothur	Versatile software suite (designed mostly for 16S rRNA)	35
	<a href="https://www.mothur.org">https://www.mothur.org</a>	
AMPTk	Full pipeline for amplicon data	27
	<a href="http://amptk.readthedocs.io">http://amptk.readthedocs.io</a>	
OBITools	Versatile software package	184
	<a href="https://git.metabarcoding.org/obitools">https://git.metabarcoding.org/obitools</a>	
PipeCraft	Full pipeline for amplicon data (with graphical user interface)	46
	<a href="https://plutof.ut.ee/#/datacite/10.15156%2FBIO%2F587450">https://plutof.ut.ee/#/datacite/10.15156%2FBIO%2F587450</a>	
PIPITS	Full pipeline for fungal ITS amplicon data (only for Illumina data)	48
	<a href="https://github.com/hsgweon/pipits">https://github.com/hsgweon/pipits</a>	
QIIME	Full pipeline for amplicon data (designed mostly for 16S rRNA)	185
	<a href="https://qiime2.org">https://qiime2.org</a>	
SEED2	Full pipeline for amplicon data (with graphical user interface; on Windows)	186
	<a href="http://www.biomed.cas.cz/mbu/lbwrf/seed">http://www.biomed.cas.cz/mbu/lbwrf/seed</a>	
Microbiology.se	Tools, including ITSx and Metaxa2, for processing ITS, SSU and LSU data	32, 187
	<a href="http://microbiology.se">http://microbiology.se</a>	
USEARCH	Versatile software package	33
	<a href="https://www.drive5.com/usearch">https://www.drive5.com/usearch</a>	
VSEARCH	Versatile software package	34
	<a href="https://github.com/torognes/vsearch">https://github.com/torognes/vsearch</a>	

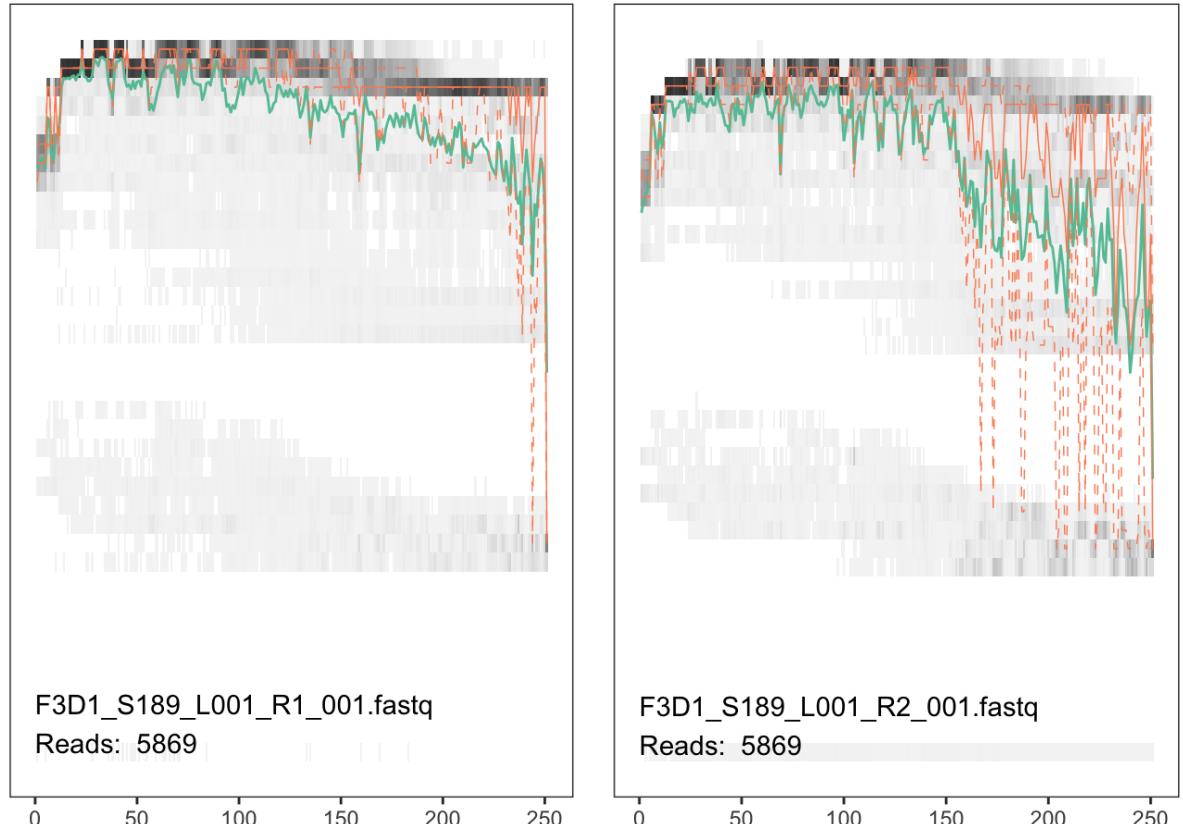
ITS, internal transcribed spacer; LSU, large subunit; rRNA, ribosomal RNA; SSU, small subunit.

Nilsson et al 2019

# Things to look for in a pipeline

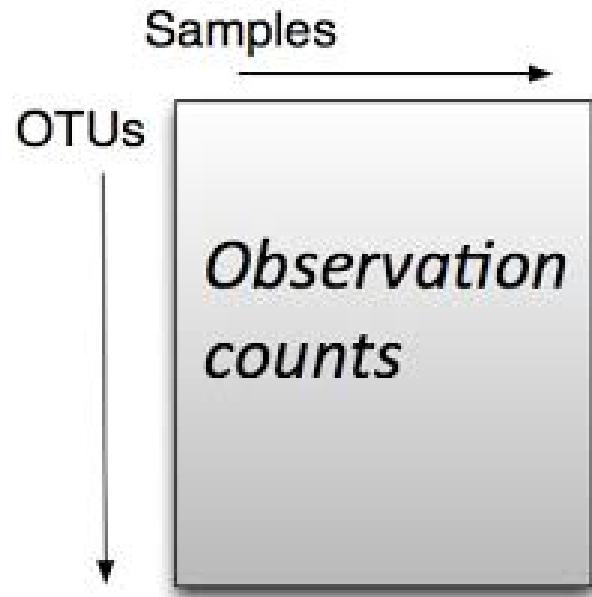
- Different quality control steps:
  - Sequence length (trimming)
  - Low quality reads (filtering)
  - Deal with sequencing errors (denoising/error correction)
  - Homopolymers
  - Chimeras
- OTU vs SV (podcast:  
<https://bioinformatics.chat/amplicon-sequence-variants>)
- Deal with controls
- Documentation
- Format/compatibility with downstream analysis

FASTQ – quality profile plot



[https://benjjneb.github.io/dada2/tutorial\\_1\\_6.html](https://benjjneb.github.io/dada2/tutorial_1_6.html)

# Analysing and interpreting your data

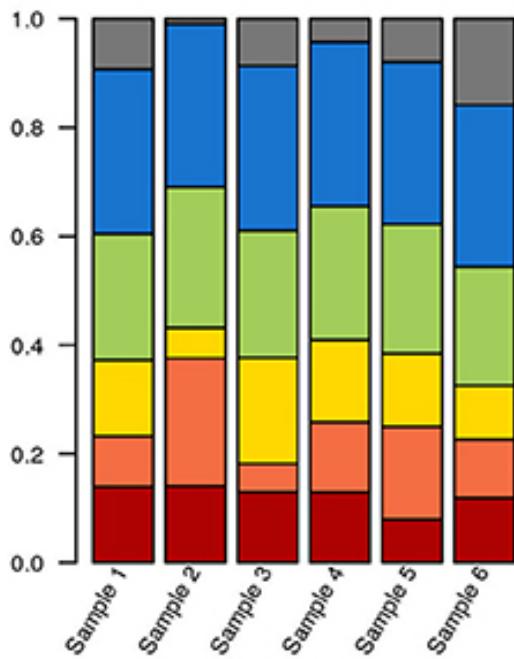
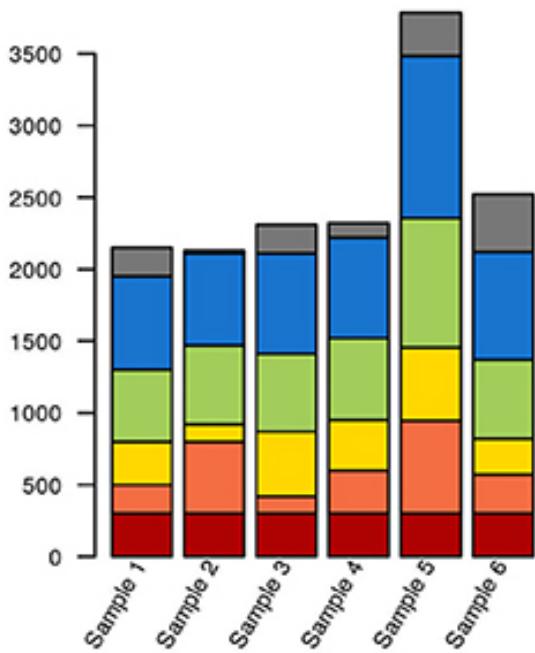


	A	B	C	D	E	F
1		sample01	sample01	sample02	sample02	sample03
2	OTU0001	528	68	1755	773	2167
3	OTU0002	138	68	559	2588	1198
4	OTU0003	36	533	673	351	815
5	OTU0004	1	2618	5	17	19
6	OTU0005	224	237	81	271	313

**Metadata**  
**Taxonomy**

# To normalize or not?

What is your question and hypothesis?

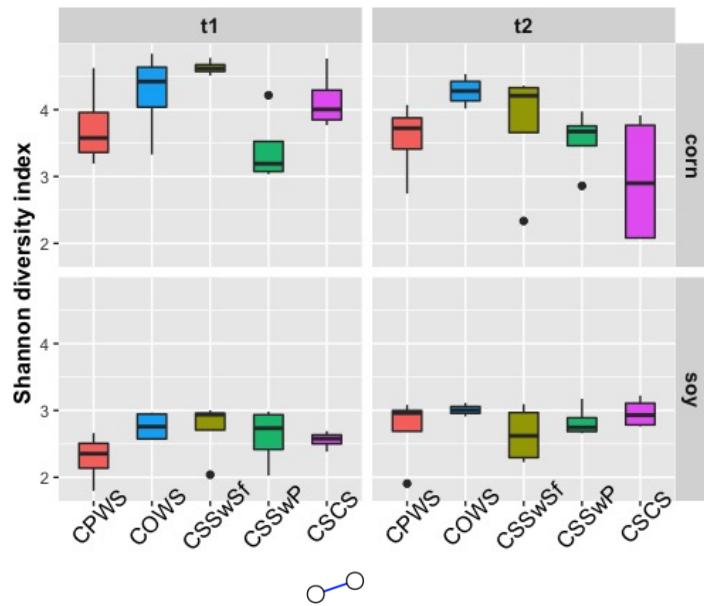


McMurdie, Paul J., and Susan Holmes. "Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible." *PLoS Comput Biol* 10, no. 4 (2014): e1003531.

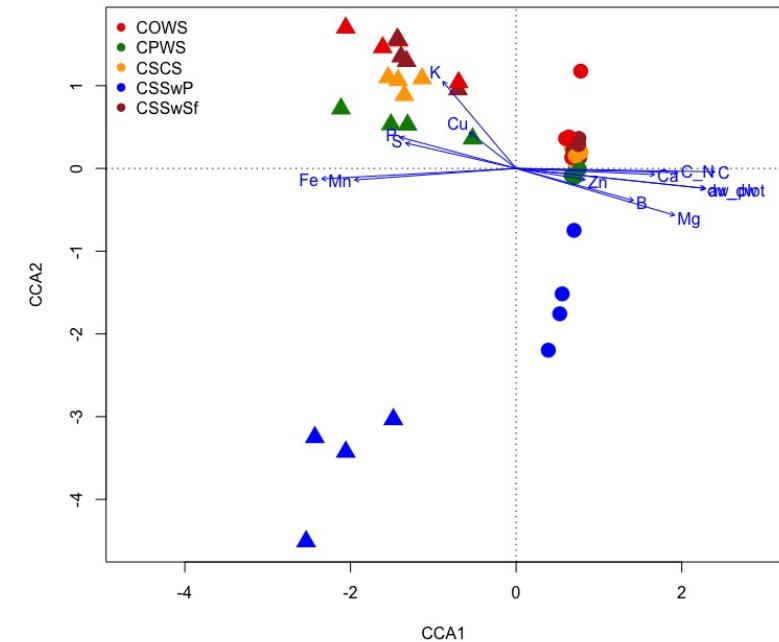
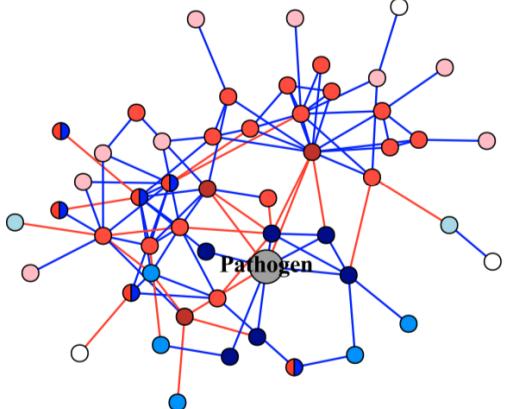
Weiss, Sophie, Zhenjiang Zech Xu, Shyamal Peddada, Amnon Amir, Kyle Bittinger, Antonio Gonzalez, Catherine Lozupone, et al. "Normalization and Microbial Differential Abundance Strategies Depend upon Data Characteristics." *Microbiome* 5, no. 1 (March 3, 2017): 27.

<https://doi.org/10.1186/s40168-017-0237-y>.

# Which analyses to apply?



D



- **Different methods, packages, visualization tools**  
phyloseq, ampvis2, vegan (R)
- **Dealing with controls, replicates and contaminants**  
decontam

# Submit to databases!!!

Phytobiomes  
Journal



A Transdisciplinary Journal of Sustainable Plant Productivity

## PERSPECTIVE

### Community-Driven Metadata Standards for Agricultural Microbiome Research

J. P. Dundore-Arias,<sup>1,†</sup> E. A. Eloë-Fadrosh,<sup>2</sup> L. M. Schriml,<sup>3</sup> G. A. Beattie,<sup>4</sup> F. P. Brennan,<sup>5</sup> P. E. Busby,<sup>6</sup> R. B. Calderon,<sup>7</sup> S. C. Castle,<sup>8</sup> J. B. Emerson,<sup>9</sup> S. E. Everhart,<sup>10</sup> K. Eversole,<sup>11</sup> K. E. Frost,<sup>12</sup> J. R. Herr,<sup>13</sup> A. I. Huerta,<sup>14</sup> A. S. Iyer-Pascuzzi,<sup>15</sup> A. K. Kalil,<sup>16</sup> J. E. Leach,<sup>17</sup> J. Leonard,<sup>18</sup> J. E. Maul,<sup>19</sup> B. Prithiviraj,<sup>20</sup> M. Potrykus,<sup>21</sup> N. R. Redekar,<sup>22</sup> J. A. Rojas,<sup>23</sup> K. A. T. Silverstein,<sup>24</sup> D. J. Tomso,<sup>25</sup> S. G. Tringe,<sup>26</sup> B. A. Vinatzer,<sup>27</sup> and L. L. Kinkel<sup>28</sup>