

Seven years of microbial community metagenomes from temperate soils affected by an ongoing
coal seam fire

Samuel E. Barnett¹, Ashley Shade^{2#}

¹Department of Microbiology, Genetics, and Immunology, Michigan State University, East
Lansing, MI 48824, USA

² Université Claude Bernard Lyon 1; CNRS, UMR 5557 Écologie Microbienne, 16 rue Dubois
69622 Villeurbanne Cedex, France

Running title: Seven-year metagenome timeseries from Centralia, PA soils

#Address correspondence to Ashley Shade: ashley.shade@cnrs.fr

Abstract

We examined the dynamics of soil microbiomes under heat press disturbance from an underground coal mine fire in Centralia, Pennsylvania. Here we present metagenomic sequencing and assembly data from soil microbiomes across seven consecutive years at repeatedly sampled fire affected sites along with unaffected reference sites.

Announcement

Anthropogenic pressures on the environment fundamentally change ecosystems and their functions, in part by altering microbial communities responsible for important ecosystem services. As many anthropogenic disturbances act long-term such “press” disturbances require committed, interannual observations to observe and understand the resilience of microbial communities and their functional consequences. The six-decade-old coal mine fire in Centralia, Pennsylvania, USA, provides a unique case study to examine the resilience of soil microbial communities to an intense, long-term thermal disturbance (1). We expanded upon previous research at Centralia (2–4) by sequencing metagenomes annually collected from ten sites over seven consecutive years.

Sampling sites in Centralia Pennsylvania USA have been previously discussed (4). For metagenome sequencing, we utilized three reference sites (*i.e.*, unheated soils) and seven fire affected sites (*i.e.*, actively heated soils). 20 cm deep soil cores were collected using a sterilized PVC pipe corer or trowel the first or second week in October yearly between 2015 and 2021, immediately sieved to 4 mm and flash frozen in liquid N₂ in the field and then stored in the lab at -80°C as previously described (4). DNA was extracted from 0.25 g of each soil using phenol-chloroform extraction (5) with BeadBug homogenizer tubes (#Z763764; Benchmark Scientific,

Sayreville USA). Bead beating was performed for 60 s on a Mini-Beadbeater 16 (BioSpec Products Inc., Bartlesville, USA). Illumina shotgun metagenomic DNA libraries were prepared and sequenced by the Research Technology Support Facility at Michigan State University (East Lansing, MI, USA). Libraries were generated, barcoded with unique dual indexes, and pooled using the ThruPLEX DNA-Seq Kit (Takara Bio Inc., Kusatsu, Japan) following standard protocols. At this stage, one sample was removed due to failed library preparation and the remaining 69 other samples were pooled. Metagenomic sequencing was performed in two lanes of a NovaSeq S4 flow cell in 2x150bp paired-end format on a NovaSeq 6000 (Illumina, San Diego, USA).

Metagenomic sequencing reads were processed using a custom workflow based on the Joint Genome Institute SOP (6). Raw reads were first quality trimmed and filtered using bbduk from BBMap version 38.18 (7). Filtered reads were then error corrected using bbcms from BBMap version 38.18 (7). Quality controlled reads were then assembled within individual samples using SPAdes version 3.15.5 in metaSPAdes mode (8) with assembler only. Contigs under 1000 bp were removed and remaining contig taxonomy was determined using Kraken version 2.1.3 (9).

Combined runs generated between 47 million and 172 million quality filtered reads (Table 1). Assemblies produced between 18,000 and 192,000 contigs longer than 1000 bp with N50 values ranging from 1458-11958 (Table 1). Between 58.8 and 79.9% of contigs were classified by Kraken to the phylum level. Of these contigs, the top 10 phyla represented were Pseudomonadota (50.4% on average), Actinomycetota (37.1%), Planctomycetota (3.1%), Acidobacteriota (2.6%), Myxococcota (1.4%), Bacillota (1.2%), Bacteroidota (0.9%), Chloroflexota (0.5%), Thermodesulfobacteriota (0.4%), and Cyanobacteriota (0.4%) (Fig. 1).

Data availability

Raw reads and assembled contigs at least 1000 bp long are available through NCBI in the SRA and genome databases respectively, both under BioProject [PRJNA974462](https://ncbi.nlm.nih.gov/bioproject/PRJNA974462). Raw read and assembly accessions are indicated in Table 1. Code for read processing and assembly are available on GitHub at https://github.com/ShadeLab/Centralia_7year_metagenome_processing_Barnett_2024.

Acknowledgements

This work was supported by the U.S. National Science Foundation CAREER award #1749544 to AS.

References

1. Shade A. 2018. Understanding Microbiome Stability in a Changing World. *mSystems* 3:e00157-17.
2. Lee S-H, Sorensen JW, Grady KL, Tobin TC, Shade A. 2017. Divergent extremes but convergent recovery of bacterial and archaeal soil communities to an ongoing subterranean coal mine fire. *ISME J* 11:1447–1459.
3. Kearns PJ, Shade A. 2018. Trait-based patterns of microbial dynamics in dormancy potential and heterotrophic strategy: case studies of resource-based and post-press succession. *ISME J* 12:2575–2581.
4. Barnett SE, Shade A. 2024. Arrive and wait: Inactive bacterial taxa contribute to perceived soil microbiome resilience after a multidecadal press disturbance. *Ecol Lett*

27:e14393.

5. Griffiths RI, Whiteley AS, O'Donnell AG, Bailey MJ. 2000. Rapid Method for Coextraction of DNA and RNA from Natural Environments for Analysis of Ribosomal DNA- and rRNA-Based Microbial Community Composition. *Appl Environ Microbiol* 66:5488–5491.
6. Clum A, Huntemann M, Bushnell B, Foster B, Foster B, Roux S, Hajek PP, Varghese N, Mukherjee S, Reddy TBK, Daum C, Yoshinaga Y, O'Malley R, Seshadri R, Kyrpides NC, Elie-Fadrosh EA, Chen I-MA, Copeland A, Ivanova NN. 2021. DOE JGI Metagenome Workflow. *mSystems* 6:e00804-20.
7. Bushnell B. 2019. BBMap short-read aligner, and other bioinformatics tools.
8. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome Res* 27:824–834.
9. Wood DE, Lu J, Langmead B. 2019. Improved metagenomic analysis with Kraken 2. *Genome Biol* 20:257.

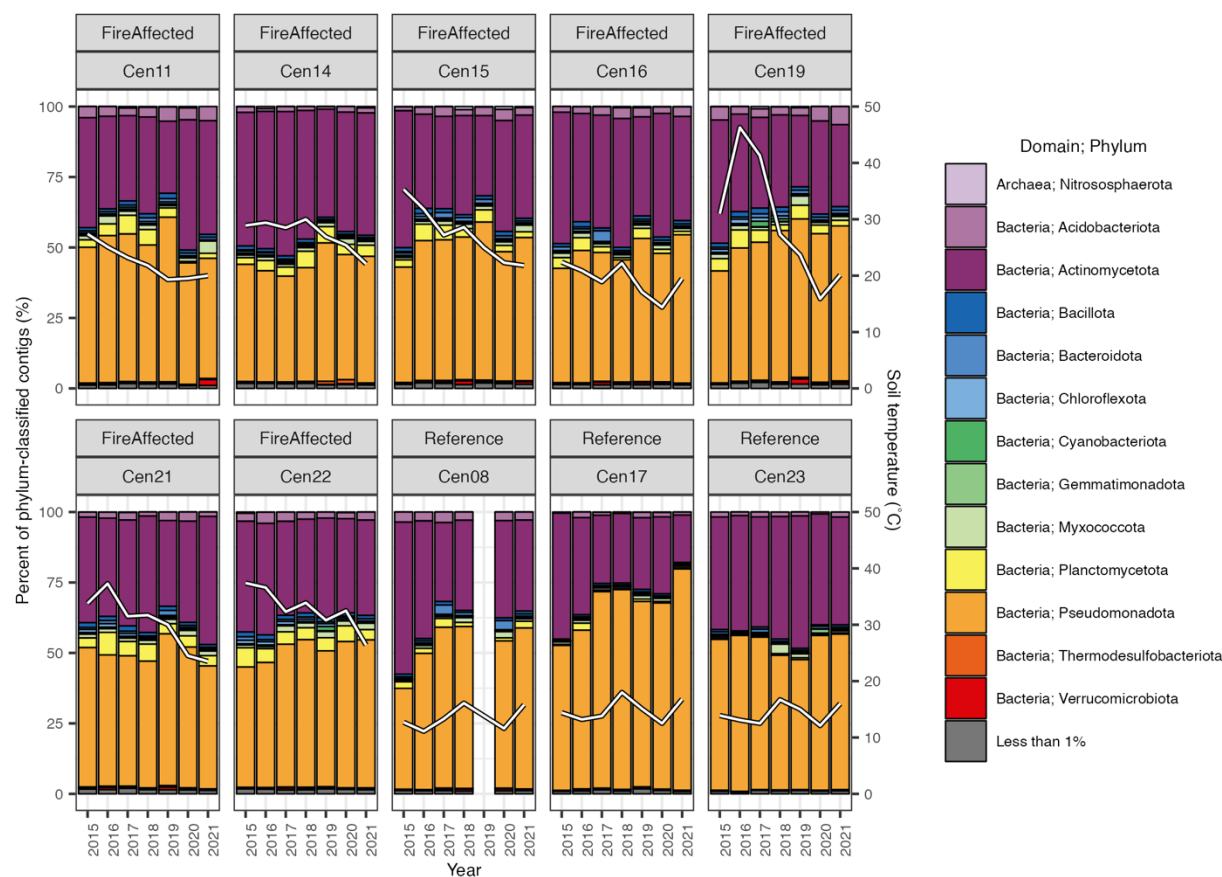
Sample ID	Site	Collection date	Soil Temp. (°C)	Fire classification	Raw read accession	No. filtered reads	Assembly accession	No. contigs	Assembly length (bp)	N50 (bp)	Longest Contig (bp)
Cen08_13102015	Cen08	13-Oct-2015	12.7	Reference	SRR24677618	124327406	JBAIFX000000000	160124	351137115	2279	658717
Cen08_12102016	Cen08	12-Oct-2016	11.1	Reference	SRR24677617	86715826	JBAIFY000000000	101237	243866855	2683	1053386
Cen08_21102017	Cen08	21-Oct-2017	13.3	Reference	SRR24677606	75877554	JBAIFZ000000000	74881	186130345	2816	155567
Cen08_04102018	Cen08	4-Oct-2018	16.1	Reference	SRR24677595	87148206	JBAIGA000000000	69059	177362065	3067	303328
Cen08_15102020	Cen08	15-Oct-2020	11.6	Reference	SRR24677584	80638542	JBAIGB000000000	72853	195537628	3389	325882
Cen08_06102021	Cen08	6-Oct-2021	15.8	Reference	SRR24677573	105866020	JBAIGC000000000	82965	234730043	3720	631307
Cen11_12102015	Cen11	12-Oct-2015	27.4	FireAffected	SRR24677562	111788990	JBAIGD000000000	138367	288073727	2024	278524
Cen11_11102016	Cen11	11-Oct-2016	25.1	FireAffected	SRR24677552	98625612	JBAIGE000000000	129881	281691484	2221	647046
Cen11_20102017	Cen11	20-Oct-2017	23.2	FireAffected	SRR24677551	114203656	JBAIGF000000000	163879	374766051	2358	895580
Cen11_03102018	Cen11	3-Oct-2018	21.8	FireAffected	SRR24677550	119434650	JBAIGG000000000	144335	345423017	2626	987027
Cen11_06102019	Cen11	6-Oct-2019	19.3	FireAffected	SRR24677616	68277302	JBAIGH000000000	82766	201039944	2701	855103
Cen11_13102020	Cen11	13-Oct-2020	19.5	FireAffected	SRR24677615	108943344	JBAIGI000000000	118956	238107306	2047	71568
Cen11_06102021	Cen11	6-Oct-2021	20	FireAffected	SRR24677614	93275172	JBAIGJ000000000	71500	151498571	2182	303453
Cen14_12102015	Cen14	12-Oct-2015	28.9	FireAffected	SRR24677613	123668420	JBAIGK000000000	117428	374847644	4394	901286
Cen14_11102016	Cen14	11-Oct-2016	29.4	FireAffected	SRR24677612	126966694	JBAIGL000000000	110577	403592710	6604	937175
Cen14_19102017	Cen14	19-Oct-2017	28.5	FireAffected	SRR24677611	122282866	JBAIGM000000000	121111	318405198	3155	717853
Cen14_02102018	Cen14	2-Oct-2018	30	FireAffected	SRR24677610	105414742	JBAIGN000000000	118460	387227442	4877	901280
Cen14_06102019	Cen14	6-Oct-2019	26.9	FireAffected	SRR24677609	101355352	JBAIGO000000000	123103	358796064	3771	883014
Cen14_13102020	Cen14	13-Oct-2020	25.4	FireAffected	SRR24677608	92846146	JBAIGP000000000	106684	262291366	2670	1293213
Cen14_05102021	Cen14	5-Oct-2021	22.1	FireAffected	SRR24677607	134497346	JBAIGQ000000000	97150	239260764	2798	1265966
Cen15_12102015	Cen15	12-Oct-2015	35.3	FireAffected	SRR24677605	125281422	JBAIGR000000000	128596	440403073	5157	1037399
Cen15_11102016	Cen15	11-Oct-2016	31.8	FireAffected	SRR24677604	153739606	JBAIGS000000000	120958	459008620	7193	960662
Cen15_19102017	Cen15	19-Oct-2017	27.2	FireAffected	SRR24677603	123055626	JBAIGT000000000	145031	425783382	3797	813692
Cen15_02102018	Cen15	2-Oct-2018	28.6	FireAffected	SRR24677602	137128144	JBAIGU000000000	144158	397881104	3459	643680

Cen15_06102019	Cen15	6-Oct-2019	25	FireAffected	SRR24677601	65267684	JBAIGV000000000	45047	112422555	2918	470205
Cen15_13102020	Cen15	13-Oct-2020	22.3	FireAffected	SRR24677600	115695690	JBAIGW000000000	113636	312812197	3463	2060246
Cen15_05102021	Cen15	5-Oct-2021	21.8	FireAffected	SRR24677599	108315540	JBAIGX000000000	99775	258547356	2858	915079
Cen16_12102015	Cen16	12-Oct-2015	22.4	FireAffected	SRR24677598	107303616	JBAIGY000000000	142748	347970221	2696	828568
Cen16_11102016	Cen16	11-Oct-2016	20.9	FireAffected	SRR24677597	113594820	JBAIGZ000000000	143271	351210329	2784	2423317
Cen16_19102017	Cen16	19-Oct-2017	18.9	FireAffected	SRR24677596	89626840	JBAIHA000000000	71367	203652565	3771	3243675
Cen16_02102018	Cen16	2-Oct-2018	22.2	FireAffected	SRR24677594	103256354	JBAIHB000000000	115981	279324043	2708	832243
Cen16_06102019	Cen16	6-Oct-2019	17.1	FireAffected	SRR24677593	98261786	JBAIHC000000000	147813	377442901	2928	2264022
Cen16_15102020	Cen16	15-Oct-2020	14.4	FireAffected	SRR24677592	99322676	JBAIHD000000000	107954	282993089	3109	998537
Cen16_05102021	Cen16	5-Oct-2021	19.5	FireAffected	SRR24677591	124001218	JBAIHE000000000	142040	365981807	2963	2177591
Cen17_12102015	Cen17	12-Oct-2015	14.4	Reference	SRR24677590	75573920	JBAIHF000000000	32434	49409001	1458	43291
Cen17_12102016	Cen17	12-Oct-2016	13.2	Reference	SRR24677589	98667852	JBAIHG000000000	110909	208504520	1821	312323
Cen17_21102017	Cen17	21-Oct-2017	13.8	Reference	SRR24677588	86602822	JBAIHH000000000	31726	51312983	1478	243625
Cen17_03102018	Cen17	3-Oct-2018	18	Reference	SRR24677587	96476092	JBAIHI000000000	25206	42231953	1526	134148
Cen17_06102019	Cen17	6-Oct-2019	15.1	Reference	SRR24677586	96267612	JBAIHH000000000	52161	82649230	1517	53302
Cen17_14102020	Cen17	14-Oct-2020	12.5	Reference	SRR24677585	88243710	JBAIHK000000000	38196	63932648	1575	121107
Cen17_05102021	Cen17	5-Oct-2021	16.8	Reference	SRR24677583	50766062	JBAIHL000000000	18915	30147416	1502	66439
Cen19_12102015	Cen19	12-Oct-2015	31	FireAffected	SRR24677582	131714862	JBAIHM000000000	183930	558956168	4188	1562083
Cen19_11102016	Cen19	11-Oct-2016	46.3	FireAffected	SRR24677581	139131824	JBAIHN000000000	98344	387090322	7537	2346858
Cen19_19102017	Cen19	19-Oct-2017	41.3	FireAffected	SRR24677580	118895810	JBAIHO000000000	89093	324744972	6338	1496731
Cen19_02102018	Cen19	2-Oct-2018	27.3	FireAffected	SRR24677579	112603898	JBAIHP000000000	117000	378578881	5051	1690910
Cen19_06102019	Cen19	6-Oct-2019	23.7	FireAffected	SRR24677578	118770438	JBAIHQ000000000	118285	301729243	2814	1263664
Cen19_13102020	Cen19	13-Oct-2020	15.9	FireAffected	SRR24677577	92357358	JBAIHR000000000	102505	291470536	3685	1174055
Cen19_05102021	Cen19	5-Oct-2021	20.1	FireAffected	SRR24677576	103309696	JBAIHS000000000	120270	306633097	2938	797129
Cen21_12102015	Cen21	12-Oct-2015	33.8	FireAffected	SRR24677575	138522504	JBAIHT000000000	107426	419858096	7997	3014071
Cen21_11102016	Cen21	11-Oct-2016	37.3	FireAffected	SRR24677574	144438616	JBAIHU000000000	138350	531085627	7301	1234492
Cen21_19102017	Cen21	19-Oct-2017	31.5	FireAffected	SRR24677572	107447060	JBAIHV000000000	107230	383560699	6302	992864
Cen21_02102018	Cen21	2-Oct-2018	31.7	FireAffected	SRR24677571	123745008	JBAIHW000000000	131392	281481763	2161	1013129

Cen21_06102019	Cen21	6-Oct-2019	30	FireAffected	SRR24677570	95991016	JBAIHX000000000	136882	358586810	2982	1789082
Cen21_13102020	Cen21	13-Oct-2020	24.5	FireAffected	SRR24677569	95674248	JBAIHY000000000	117940	299851161	2876	525405
Cen21_05102021	Cen21	5-Oct-2021	23.5	FireAffected	SRR24677568	149012682	JBAIHZ000000000	163041	414815233	2882	521355
Cen22_12102015	Cen22	12-Oct-2015	37.4	FireAffected	SRR24677567	172424074	JBAIIA000000000	85782	404428333	11958	1907411
Cen22_11102016	Cen22	11-Oct-2016	36.6	FireAffected	SRR24677566	134402064	JBAIIB000000000	108876	432565722	7259	2072426
Cen22_20102017	Cen22	20-Oct-2017	32.3	FireAffected	SRR24677565	143726066	JBAIIC000000000	163716	606439961	6325	2160330
Cen22_03102018	Cen22	3-Oct-2018	33.9	FireAffected	SRR24677564	127906794	JBAIID000000000	191865	585602205	4031	1469064
Cen22_06102019	Cen22	6-Oct-2019	30.8	FireAffected	SRR24677563	127426364	JBAIIE000000000	142582	488030302	5126	1690897
Cen22_13102020	Cen22	13-Oct-2020	32.5	FireAffected	SRR24677561	112687806	JBAIIF000000000	152990	503652415	5174	2160791
Cen22_07102021	Cen22	7-Oct-2021	26.4	FireAffected	SRR24677560	109576490	JBAIIG000000000	152720	424686625	3493	1577924
Cen23_13102015	Cen23	13-Oct-2015	13.9	Reference	SRR24677559	97457822	JBAIIH000000000	71816	141042993	1858	352202
Cen23_12102016	Cen23	12-Oct-2016	13.1	Reference	SRR24677558	93880664	JBAIII000000000	116985	219169519	1842	1138925
Cen23_21102017	Cen23	21-Oct-2017	12.5	Reference	SRR24677557	94190168	JBAIIJ000000000	83044	156188699	1837	103517
Cen23_04102018	Cen23	4-Oct-2018	16.7	Reference	SRR24677556	47821322	JBAIIK000000000	19974	36773187	1825	44173
Cen23_07102019	Cen23	7-Oct-2019	15	Reference	SRR24677555	102127352	JBAIIL000000000	83503	153966893	1774	97317
Cen23_15102020	Cen23	15-Oct-2020	12.1	Reference	SRR24677554	74528594	JBAIIM000000000	52468	105979068	2003	89100
Cen23_06102021	Cen23	6-Oct-2021	16	Reference	SRR24677553	70046314	JBAIIN000000000	48122	88671983	1795	57462

100

101 **Table 1:** Metagenome info for all samples including read and assembly summaries and accessions. Number of reads passing quality
102 filtering combines both forward and reversed reads. Contigs only include those at least 1000 bp long.



105 **Figure 1:** Contributions of the topmost represented phyla to the pool of contigs classified to the
106 phylum level. Contribution is defined as the percent of the classified contigs. Phyla not making
107 up at least 1% of the contigs in any sample are combined as “less than 1%”. White lines indicate
108 soil temperature at sampling (secondary y axis).