For Sipes et al AEM submission, "Eight metagenome-assembled genomes provide evidence for microbial adaptation in 20,000 to 1,000,000-year-old Siberian permafrost"
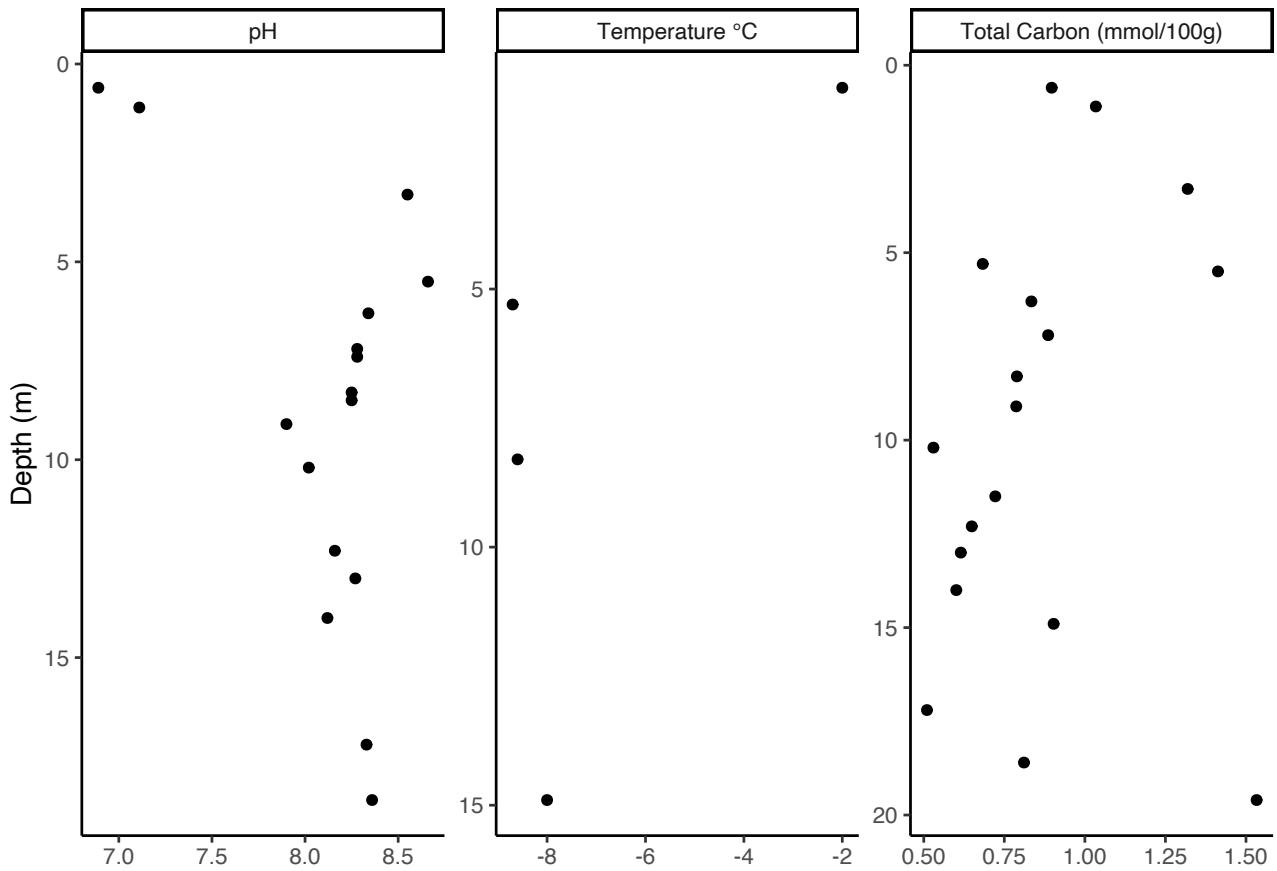

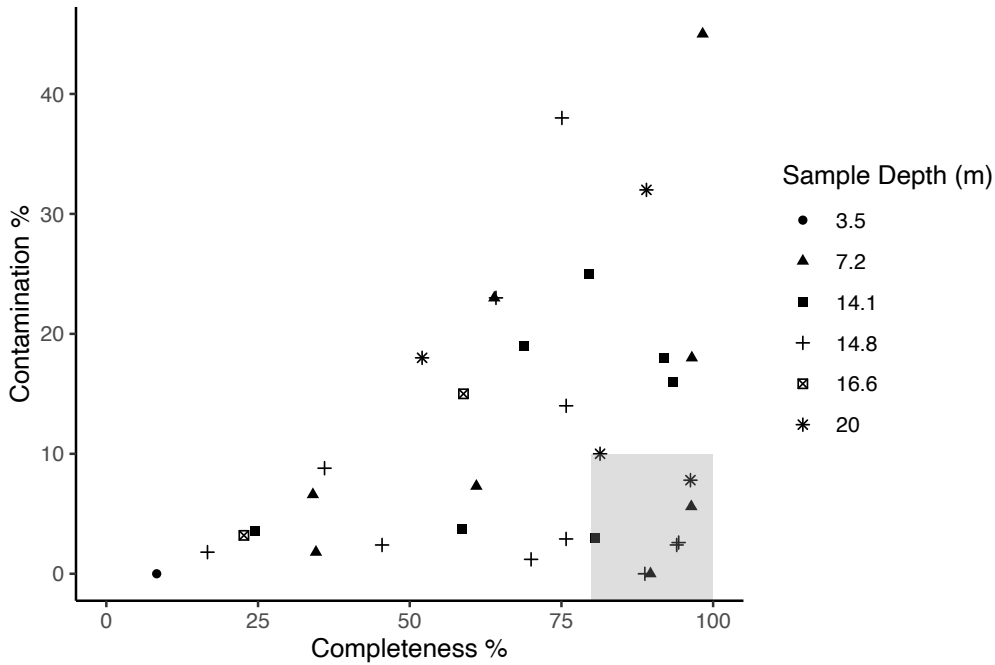**Figure S1:** Extra field-tested parameters from borehole AL3-15.

Figure S2: Siberian permafrost MAG completeness and contamination, shown by the depth of the permafrost layer from which they originated. The shaded region denotes MAGs that were focused on for this study. (≥80% completeness, <10% contamination)
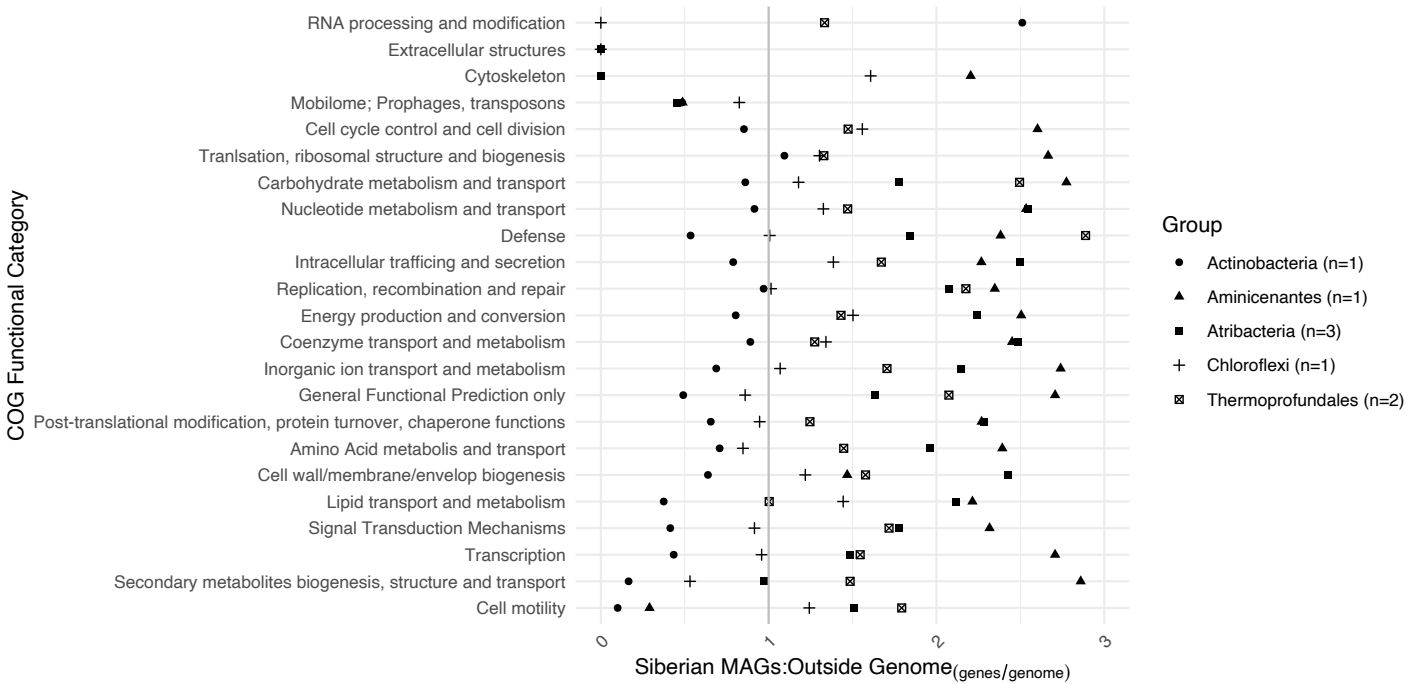

Figure S3: Ratio of genes per MAG/genome for each group. Genes were summed for each COG Category and divided by the number of MAGs/genomes for each of the eight Siberian MAGs and for the reference genomes. The values shown are the ratio of these genes per MAG/Genome for the eight Siberian MAGs to the ratio of these genes per MAG/Genome for the reference genomes of that group.
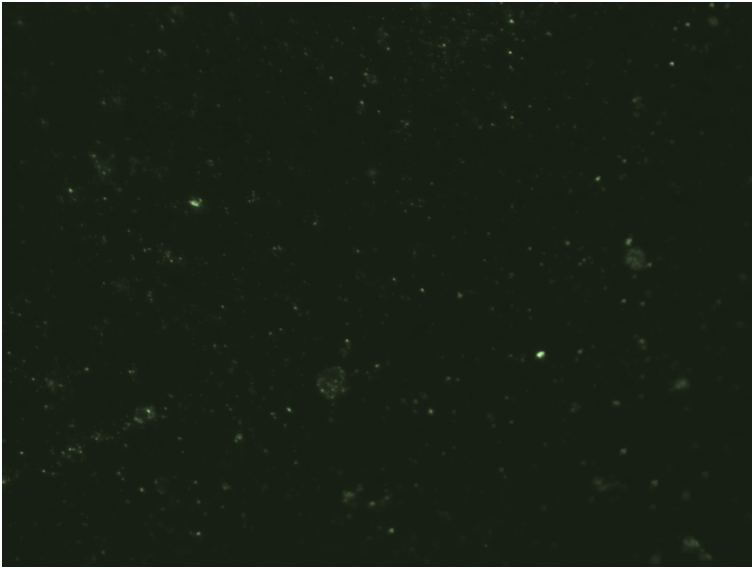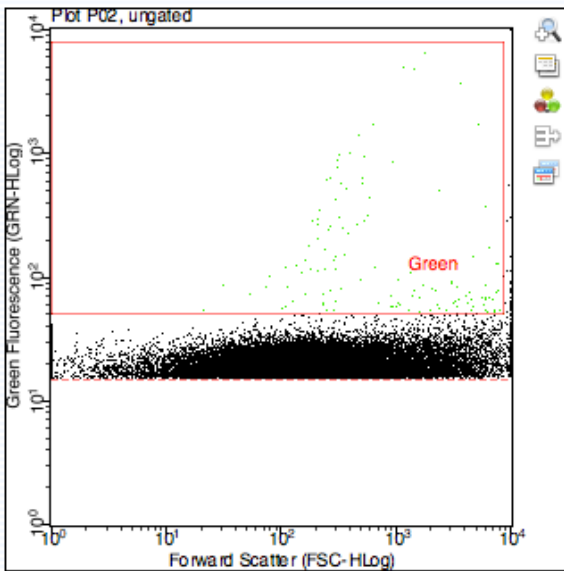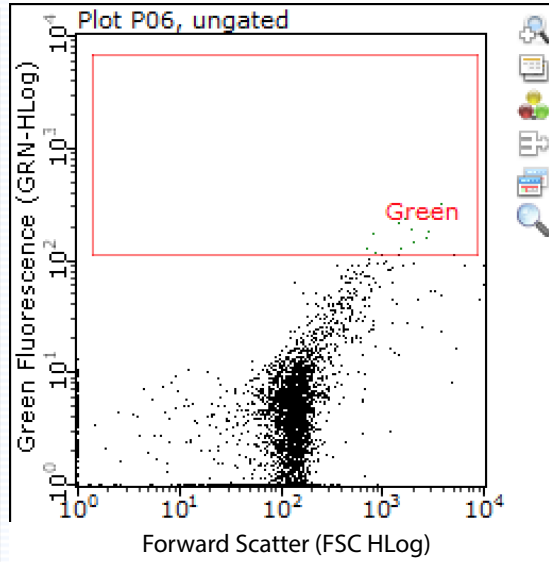
Figure S4: A) Cells from the 5.6m sample, stained with SYBR Gold and viewed under Zeiss Imager M2 1000X magnification. **B)** Guava easyCyte 12HT Benchtop Flow Cytometer output with a gated size and fluorescence (green region, outlined in the red box) compared to the background noise (black) on 3.5m sample. **C)** Similar as panel B but with 7.2m sample.

A
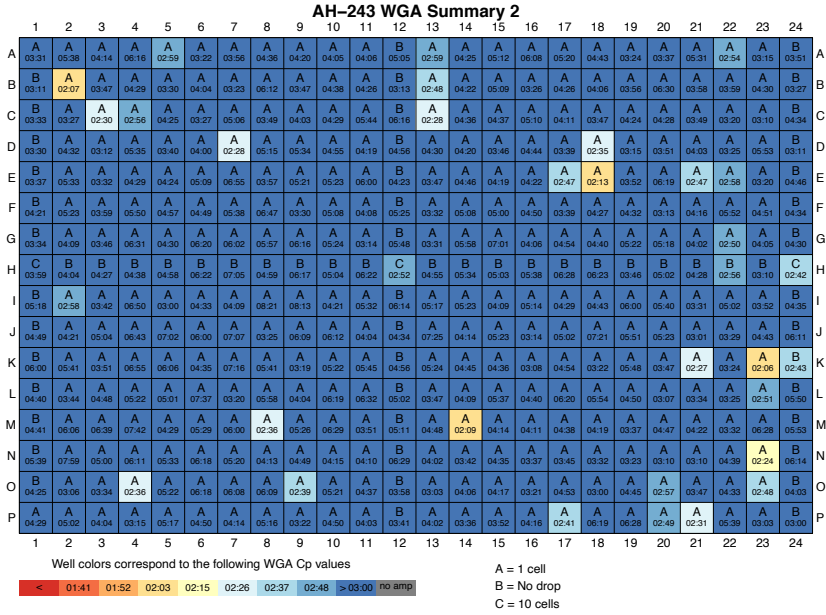
B
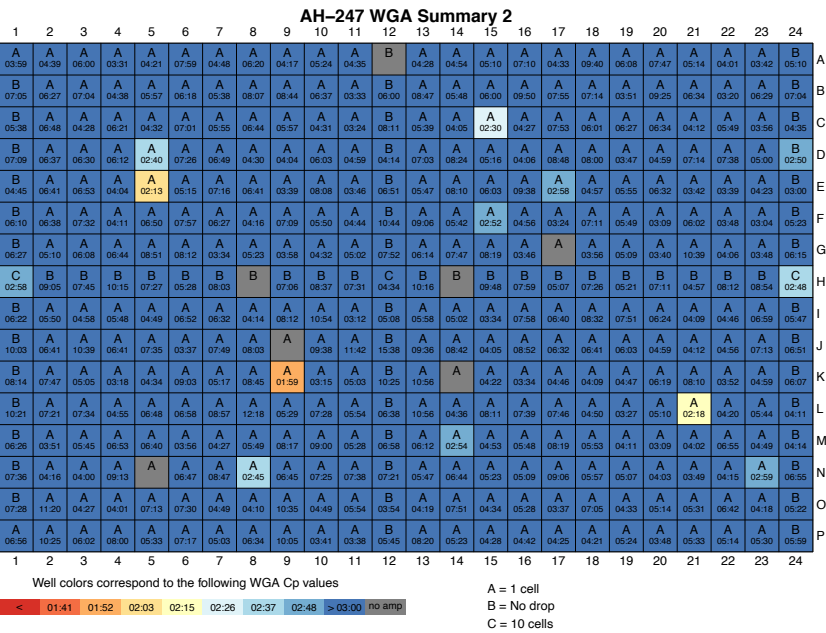
Figure S5: Soil from Borehole AL3-15 underwent WGA-X from Bigelow with SAG Generation 2 single cell sorting servicing. AH 243 plate (A) was created from soil 5.6m below the surface, the wells that were taken for further library preparation were B2, E18, K23, M14, and N23. AH247 (B) was from 7.2m below the surface of borehole AL3-15, the wells that were used for library prep were E5, K9, L21, N8, and C15.
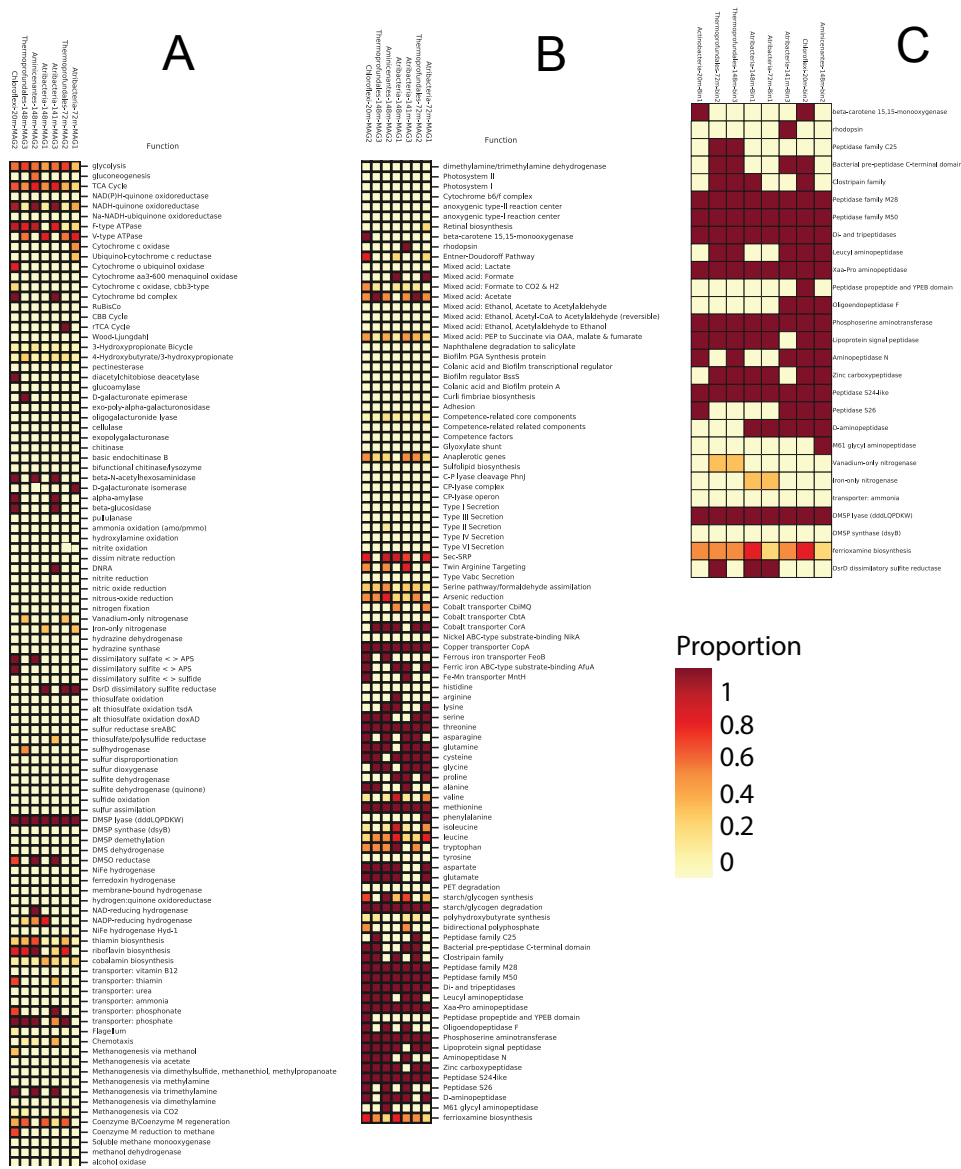
Figure S6: Functional pathways for each of the representative MAGs. Code used to make this can be found on https://github.com/bjtully/BioData/tree/master/KEGGDecoder. A-B) Standard KEGG-Decoder output cut in half for readability. C) Output from KEGG-Expander code. Color saturation indicates the percentage of genes present from that metabolic pathway. Where single genes are listed instead of pathways (e.g., Leucyl aminopeptidase) the possible values are either 0% or 100%.

Bioinformatic Methods

Metagenomic curating

Dneasy PowerSoil Kit- DNA extraction
(5g permafrost x 10 technical reps)

Nextera XT DNA Library Prep Kit @ University of Tennessee Genomics Core. Illumina MiSeq 2 x 250 cycle protocol

Kbase.us:

Trimmomactic v 0.36
Sliding window size: 10
Min quality: 28

MetaSPAdes v3.11.1 for Single Sample Assembly
Min contig length 2000bp

MaxBin2 Contig Binning v2.2.3
probability threshold 0.8
min contig length 2000bp

GTDB-TK v 1.4.0 and maximum likelihood tree w/139 conserved protein & 230 reference genome on CLC Workbench 6

CheckM v 1.0.18 for MAG completeness and contamination

Bowtie2 v 2.3.3 for read mapping MAGs <-> Metagenomes

JGI Open Access Genomes

Download genomes from non-permafrost environments
Atribacteria (n=66)
Actinobacteria (n = 57)
Thermoprofundales (n = 9)
Chloroflexi (n = 42)
Aminicenantes (n = 54)

anvi'o v6.1: anvi-pan-genome analysis

anvi-pan-genome:
For each above phylum group
--minbit 0.5 for 50% similarity
--mcl-inflation 2 for distantly related genomes

anvi-import-misc-data: to compare the Siberian MAG to the reference genomes from JGI.

anvi-get-enriched-functions-per-pan-group:
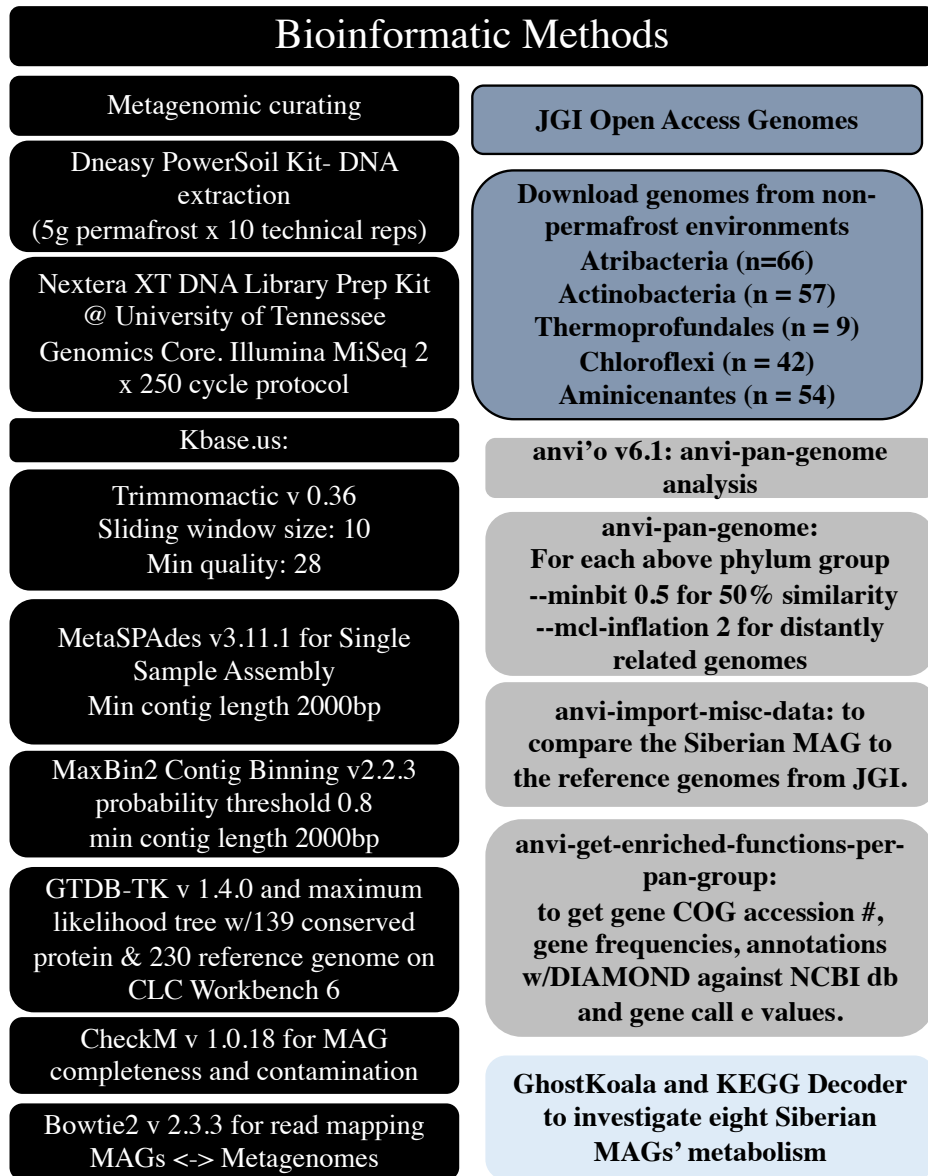**to get gene COG accession #, gene frequencies, annotations w/DIAMOND against NCBI db and gene call e values.**

GhostKoala and KEGG Decoder to investigate eight Siberian MAGs' metabolism

Figure S7: Pictorial representation of the methods used to conduct this study with special attention to the parameters used for each step. Black coloration highlights metagenome processing, open access genomes taken from JGI Database is highlighted in blue. Grey color highlights pan genomic analysis done with Anvi'o v 6.1. Light blue highlights KEGG Decoder for metabolism. 'db' stands for database.

List of Supplementary Datafiles

Supplemental Info For	File Number
Thermoprofundales	1
Atribacteria	2
Chloroflexi	3

Actinobacteria	4
Aminicenantes	5
Genes	6
Read Mapping	7

Table S1: Age determination of inorganic and organic fractions using ^{14}C dating from upper Yedoma Siberian permafrost samples. The inorganic carbon fraction was released via reaction with phosphoric acid. Two different fractions of organic carbon (presumably "younger" and "older") were sequentially released by combustion at 400°C and 800˚C, respectively.

Sample Depth (m)	Treatment	C yield (mg)	δ^{13}C‰ VPDB	^{14}C age BP (yr)
2.9-3	Inorganic C	0.73	-6.0	21,760±120
2.9-3	400˚C	1.01	-26.7	41,700±1,400
2.9-3	800˚C	1.41	-28.8	38,590±980
3.5	Inorganic C	1.64	-3.7	26,260±210
3.5	400˚C	1.21	-26.9	20,158±99
3.5	800˚C	2.34	-27.5	18,228±78
5.6	Inorganic C	2.36	-4.6	33,900±550
5.6	400˚C	0.97	-27.8	38,020±910
5.6	800˚C	0.27	-24.3	29,040±420

Table S2: All metagenome assembled genomes (MAG's) completeness and contamination scores as reported by CheckM. GTDB-Tk taxonomy was used. Missing information is due to unavailable assignments. MAGs analyzed in the main text were those with ≥80% completeness and ≤10% contamination.

Sample Depth and MAG	Completeness %	Contamination %	Domain	Phylum	Class
03.5m MAG 01	8	0	Bacteria	Firmicutes	Bacilli
07.2m MAG 01	90	0	Bacteria	Caldiatribacteriota	JS1
07.2m MAG 02	96	6	Archaea	Thermoplasmatota	E2

07.2m MAG 03	96	18	Bacteria	Planctomycetota	
07.2m MAG 04	64	23	Bacteria	Actinobacteriota	UBA1414
07.2m MAG 05	35	2	Bacteria	Actinobacteriota	UBA1414
07.2m MAG 06	98	44	Bacteria	Aminicenantes	Aminicenantia
07.2m MAG 07	61	7	Bacteria	Firmicutes	Mahellia
07.2m MAG 08	34	7	Archaea	Bathyarchaeota	
14.1m MAG 01	25	4	Bacteria	Actinobacteriota	UBA1414
14.1m MAG 02	80	25	Bacteria	Caldiatribacteriota	JS1
14.1m MAG 03	81	3	Bacteria	Caldiatribacteriota	
14.1m MAG 04	59	4	Bacteria	Actinobacteriota	UBA1414
14.1m MAG 05	93	16	Archaea	Thermoplasmatota	E2
14.1m MAG 06	92	18	Bacteria	Aminicenantes	Aminicenantia
14.1m MAG 07	69	19	Bacteria	Aminicenantes	Luteitaleia
14.8m MAG 01	89	0	Bacteria	Caldiatribacteriota	JS1
14.8m MAG 02	94	3	Bacteria	Aminicenantes	
14.8m MAG 03	94	2	Archaea	Thermoplasmatota	E2
14.8m MAG 04	70	1	Bacteria	Actinobacteriota	UBA1414
14.8m MAG 05	36	9	Bacteria	Planctomycetota	Phycisphaerae
14.8m MAG 06	17	2	Bacteria	Actinobacteriota	UBA1414
14.8m MAG 07	76	14	Bacteria	Chloroflexota	Anaerolineae
14.8m MAG 08	45	2	Bacteria	Actinobacteriota	UBA1414
14.8m MAG 09	75	38	Archaea	Crenarchaeota	Bathyarchaeia
14.8m MAG 10	75	3	Bacteria	Planctomycetota	FEN-1346
14.8m MAG 11	64	23	Bacteria	Acidobacteriota	Vicinamibacteria
16.6m MAG 01	59	15	Bacteria	Caldiatribacteriota	JS1
16.6m MAG 2	23	3	Bacteria	Chloroflexota	UBA2235
20m MAG 01	96	8	Bacteria	Actinobacteriota	Actinobacteria
20m MAG 02	81	10	Bacteria	Chloroflexota	Anaerolineae
20m MAG 03	89	32	Bacteria	Actinobacteriota	Thermoleophilia
20m MAG 04	52	18	Bacteria	Actinobacteriota	Actinobacteria

Table S3: List of genes that made up the created COG categories in Figure 3B.

Transporters	GeneName	Taxa	COG Letter	COG Category
	ABC-type transport system involved in Fe-S cluster assembly, permease and ATPase components	Actinobacteria	O	Post-translational modification, protein turnover, chaperone functions
	Na+-translocating ferredoxin:NAD+ oxidoreductase RNF, RnfG subunit	Actinobacteria	C	Energy production and conversion
	Succinate-acetate transporter protein	Aminicenantes	S	Function Unknown
	ABC-type amino acid transport/signal transduction system, periplasmic component/domain	Atribacteria	T	Signal Transduction Mechanisms

	ABC-type amino acid transport/signal transduction system, periplasmic component/domain	Atribacteria	E	Amino Acid metabolis and transport
	ABC-type amino acid transport/signal transduction system, periplasmic component/domain	Atribacteria	T	Signal Transduction Mechanisms
	ABC-type amino acid transport/signal transduction system, periplasmic component/domain	Atribacteria	E	Amino Acid metabolis and transport
	Cation transport ATPase	Atribacteria	P	Inorganic ion transport and metabolism
	Lipopolysaccharide export system protein LptC	Atribacteria	M	Cell wall/membrane/envelop biogenesis
	Na+/H+ antiporter NhaD or related arsenite permease	Atribacteria	P	Inorganic ion transport and metabolism
	Predicted ABC-type sugar transport system, permease component	Atribacteria	R	General Functional Prediction only
	Na+/alanine symporter	Chloroflexi	E	Amino Acid metabolis and transport
	ABC-type glycerol-3-phosphate transport system, periplasmic component	Thermoprofundales	G	Carbohydrate metabolism and transport
	ABC-type glycerol-3-phosphate transport system, permease component	Thermoprofundales	G	Carbohydrate metabolism and transport
	ABC-type sugar transport system, ATPase component	Thermoprofundales	G	Carbohydrate metabolism and transport
	ABC-type sugar transport system, permease component	Thermoprofundales	G	Carbohydrate metabolism and transport
	Ca2+/H+ antiporter	Thermoprofundales	P	Inorganic ion transport and metabolism
	Na+/H+ antiporter NhaD or related arsenite permease	Thermoprofundales	P	Inorganic ion transport and metabolism
	Putative Ca2+/H+ antiporter, TMEM165/GDT1 family	Thermoprofundales	R	General Functional Prediction only
	Ribose/xylose/arabinose/galactoside ABC-type transport system, permease component	Thermoprofundales	G	Carbohydrate metabolism and transport
	TctA family transporter	Thermoprofundales	R	General Functional Prediction only
Osmoregulation				
	Uncharacterized membrane protein YagU, involved in acid resistance, DUF1440 family	Actinobacteria	S	Function Unknown
	Outer membrane lipoprotein SlyB	Actinobacteria	M	Cell wall/membrane/envelop biogenesis
	Plasmid replication initiator protein	Actinobacteria	X	Mobilome; Prophages, transposons
	Na+-translocating ferredoxin:NAD+ oxidoreductase RNF, RnfG subunit	Actinobacteria	C	Energy production and conversion
	Cation transport ATPase	Atribacteria	P	Inorganic ion transport and metabolism
	Na+/H+ antiporter NhaD or related arsenite permease	Atribacteria	P	Inorganic ion transport and metabolism
	Na+/alanine symporter	Chloroflexi	E	Amino Acid metabolis and transport
	Uncharacterized membrane protein DedA, SNARE-associated domain	Thermoprofundales	S	Function Unknown
	Predicted signal transduction protein containing a membrane domain, an EAL and a GGDEF domain	Thermoprofundales	T	Signal Transduction Mechanisms
	Na+/H+ antiporter NhaD or related arsenite permease	Thermoprofundales	P	Inorganic ion transport and metabolism
	Putative Ca2+/H+ antiporter, TMEM165/GDT1 family	Thermoprofundales	R	General Functional Prediction only
	Ca2+/H+ antiporter	Thermoprofundales	C	Energy production and conversion
	Ca2+/H+ antiporter	Thermoprofundales	P	Inorganic ion transport and metabolism
Sugar Utilization				
	Phosphotransferase system cellobiose-specific component IIA	Actinobacteria	G	Carbohydrate metabolism and transport

	Phosphotransferase system cellobiose-specific component IIC	Actinobacteria	G	Carbohydrate metabolism and transport
	Cellobiose phosphorylase	Atribacteria	G	Carbohydrate metabolism and transport
	D-lyxose ketol-isomerase	Atribacteria	G	Carbohydrate metabolism and transport
	Rhamnose utilisation protein RhaD, predicted bifunctional aldolase and dehydrogenase	Atribacteria	G	Carbohydrate metabolism and transport
	Trehalose-6-phosphate synthase	Atribacteria	G	Carbohydrate metabolism and transport
	Predicted ABC-type sugar transport system, permease component	Atribacteria	R	General Functional Prediction only
	Hydroxyethylthiazole kinase, sugar kinase family	Chloroflexi	H	Coenzyme transport and metabolism
	Cellobiose phosphorylase	Thermoprofundales	G	Carbohydrate metabolism and transport
	Ribose/xylose/arabinose/galactoside ABC-type transport system, permease component	Thermoprofundales	G	Carbohydrate metabolism and transport
	ABC-type sugar transport system, permease component	Thermoprofundales	G	Carbohydrate metabolism and transport
	ABC-type sugar transport system, ATPase component	Thermoprofundales	G	Carbohydrate metabolism and transport