Supporting Information

Fierer et al. 10.1073/pnas.1215210110

SI Materials and Methods

Sample Collection. The cold desert soils were collected from various sites with the McMurdo Dry Valleys region of Antarctica with the hot desert soils collected from sites in the southwestern United States. The seven "nondesert" soils were collected from tropical forests in Peru and Argentina, an arctic tundra in Alaska, a native tallgrass prairie in Kansas, a temperate deciduous forest in South Carolina, a temperate coniferous forest in North Carolina, and a boreal forest in Alaska (see Table S1 for additional information on each of the collected soils). We did not analyze replicate samples from individual sites as it was not our objective to assess intrasite variability, but rather to assess the general patterns structuring soil microbial communities across major biome types. At each site, the upper 5 cm of mineral soil was collected from 5 to 10 locations within a given plot of ~100 m^2 and composited into a single bulk sample. For those sites with plants, soils were collected during the peak of the plant growing season. All soils were stored at −20 °C immediately after collection.

DNA Extractions. For both the 16S rRNA gene analyses and the shotgun metagenomic analyses DNA was extracted from each soil sample using the approach described in Fierer et al. (1). Briefly, 10 g of soil from each sample was homogenized in a mortar and pestle with liquid N_2, and DNA was extracted from 0.25-g subsamples of each soil using the MoBio PowerSoil DNA extraction kit modified with an additional incubation step at 65 °C for 10 min followed by 2 min of bead beating (2). To obtain sufficient DNA for the shotgun metagenomic analyses, we conducted 4–12 replicate extractions per soil, with the replicates pooled together using the approach described previously (1).

16S rRNA Gene Analyses via Amplicon Sequencing. To determine the diversity and composition of the bacterial communities in each of these soils, we used the protocol described in Caporaso et al. (3). PCR amplifications were conducted in triplicate reactions for each of the 16 soil samples with the 515f/806r primer set that amplifies the V4–V5 region of the 16S rRNA gene. The primer set was selected as it exhibits few biases against individual bacterial taxa and even the 100-bp Illumina reads should yield accurate phylogenetic and taxonomic information (4). The primers contain the appropriate Illumina adapters and the reverse primer contains a 12-bp error-correcting barcode unique to each sample. DNA was amplified in triplicate PCR reactions following the protocol described previously (3), the triplicate reactions were composited, and the amplicons from all samples were pooled together in equimolar concentrations. Sequencing was conducted on an Illumina HiSeq2000 at the University of Colorado Genomics Core Facility following the 2 × 100 bp paired-end protocol (3). Quality filtering of reads and processing of the reads was conducted as described in Caporaso et al. (5). After quality filtering, we obtained 118,000–750,000 forward reads per sample; only the forward reads were used for downstream analyses as it has been shown that including the reverse read adds little additional information (6). For all downstream analyses, we rarefied to 118,000 randomly selected reads per sample to correct for differences in sequencing depth. Reads were assigned to phylotypes at the ≥97% sequence similarity level using the open-reference phylotype picking protocol in QIIME (7). QIIME was used to estimate both taxonomic and phylogenetic metrics of the pairwise distances between communities (Bray–Curtis and unweighted Unifrac distances, respectively). Alpha diversity was determined using both taxonomic metrics (numbers of phylotypes and Shannon index, H′) and a phylogenetic metric (Faith's phylogenetic diversity).

Shotgun Metagenomic Analyses. Shotgun metagenomic analyses were conducted on the same 16 DNA extracts used for the PCR-based 16S rRNA gene analyses described. The laboratory protocol followed that described in the Illumina Paired-End Prep kit protocol. Aliquots of each DNA sample were mechanically sheared and products were size-selected to 170–180 bp and gel purified. Sequencing was performed at Argonne National Laboratory in the Institute for Genomics and Systems Biology Next Generation Sequencing (IGSB-NGS) Core using a 2 × 100 bp sequencing run on the Illumina GAIIx. Sequences were uploaded to MG-RAST (Rapid Annotation using Subsystems Technology for Metagenomes) (8) for downstream analyses. Data accession numbers are provided in Table S2. Sequences were annotated to functional categories against the M5NR database using BLASTX at an e-value cutoff of 1×10^{-2} and the SEED subsystems hierarchy. Reads that could not be annotated by MG-RAST were discarded and subsequent analyses were performed on the metagenomes evenly sampled at random to 688,000 annotated reads per sample. Pairwise distances between each of the 16 metagenomes were determined by calculating Bray–Curtis distances from the relative abundances of reads in each of the level 2 gene categories (of which there were a total of 417 across the 16 rarefied datasets). We used the level 2 abundances instead of the abundances of individual annotated genes (of which there were ~5,000) as this should yield a more conservative estimate of the distances between metagenomes; individual genes should exhibit less overlap between samples than analyses based on more broadly defined functional categories. For taxonomic analyses of the small-subunit (SSU) and large-subunit (LSU) rRNA reads, the shotgun metagenomic data were compared with the Silva SSU and LSU databases available through MG-RAST using a maximum e-value of $1e^{-5}$, a minimum identity of 60%, and a minimum alignment length of 15. For the more detailed analyses of the metagenomic SSU (16S) rRNA reads obtained from bacteria and archaea, we extracted 16S reads from these samples using QIIME by applying a closed-reference phylotype picking process where we search reads against the Greengenes database and discard reads that fail to hit any sequences at greater than or equal to 90% identity. This resulted in 83,486 16S reads across the 16 samples. The taxonomy of each phylotype was assigned as the taxonomy of the best Greengenes hit during the closed-reference phylotype picking process.

Statistical Analyses. We conducted principal coordinates analyses in PRIMER (9), using as input the pairwise distances between metagenomes (Bray–Curtis distances) or bacterial communities (Bray–Curtis and UniFrac distances calculated from the 16S rRNA gene amplicon data). To test whether sample categories harbored significantly different metagenomes or microbial communities, we used analysis of similarities (ANOSIM) tests as implemented in PRIMER. Mantel tests were run to assess correlations between metagenomic and bacterial community distance matrices. Mantel tests were also used to compare the Bray–Curtis distances between taxonomic distributions determined via 16S rRNA amplicon sequencing versus shotgun metagenomic sequencing. To determine whether the relative abundances of individual taxa or functional gene categories were significantly different between sample categories (or between the shotgun metagenomic and PCR-based analyses), we conducted pairwise *t* tests with *P* values calculated using a Bonferroni correction for multiple comparisons.

1. Fierer N, et al. (2012) Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients. *ISME J* 6(5): 1007–1017.
2. Lauber C, Knight R, Hamady M, Fierer N (2009) Soil pH as a predictor of soil bacterial community structure at the continental scale: A pyrosequencing-based assessment. *Appl Environ Microbiol* 75:5111–5120.
3. Caporaso JG, et al. (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 6(8):1621–1624.
4. Liu Z, Lozupone C, Hamady M, Bushman FD, Knight R (2007) Short pyrosequencing reads suffice for accurate microbial community analysis. *Nucleic Acids Res* 35(18):e120.
5. Caporaso JG, et al. (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci USA* 108(Suppl 1):4516–4522.
6. Werner JJ, Zhou D, Caporaso JG, Knight R, Angenent LT (2012) Comparison of Illumina paired-end and single-direction sequencing for microbial 16S rRNA gene amplicon surveys. *ISME J* 6(7):1273–1276.
7. Caporaso JG, et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7(5):335–336.
8. Meyer F, et al. (2008) The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9:386.
9. Clarke K, Gorley R (2006) PRIMER (PRIMER-E, Plymouth, UK), version 6.

		Cold Desert						*Hot Desert*			*Tropical Forest*		*Temperate Forest*		*Prairie*	*Boreal Forest*	*Tundra*
Phylum	**Class**	**EB017**	**EB019**	**EB020**	**EB021**	**EB024**	**EB026**	**MD3**	**SF2**	**SV1**	**AR3**	**PE6**	**CL1**	**DF1**	**KP1**	**BZ1**	**TL1**
Crenarchaeota (Archaea)		**0.22**	**0.01**	**0.33**	**0.11**	**0.86**	**0.01**	**6.03**	**6.62**	**6.59**	**4.54**	**0.75**	**0.04**	**0.00**	**0.78**	**0.01**	**0.07**
	Thaumarchaeota	0.22	0.01	0.33	0.11	0.86	0.01	6.03	6.62	6.59	4.54	0.45	0.02	0.00	0.78	0.00	0.01
Euryarchaeota (Archaea)		**0.00**	**0.02**	**0.00**	**0.00**	**0.00**	**0.00**	**0.01**	**0.06**	**0.05**	**0.04**	**0.12**	**0.02**	**0.01**	**0.00**	**0.00**	**0.01**
Acidobacteria		**45.88**	**0.14**	**17.39**	**6.99**	**19.65**	**4.50**	**8.18**	**17.47**	**13.49**	**13.08**	**43.14**	**31.62**	**16.82**	**17.31**	**30.67**	**37.76**
	Acidobacteria	1.88	0.03	2.38	0.81	1.29	0.54	0.96	4.74	2.48	6.67	3.67	7.06	6.36	10.42	5.14	3.26
	Chloracidobacteria	42.36	0.08	13.35	5.83	16.90	3.60	4.44	7.73	5.12	2.51	0.08	1.04	1.29	2.88	0.28	0.29
	Solibacteres	0.77	0.01	0.32	0.04	1.04	0.06	2.16	3.78	4.86	1.16	22.78	16.26	5.80	1.82	16.44	15.17
Actinobacteria		**8.57**	**40.63**	**31.12**	**39.81**	**26.07**	**2.33**	**19.30**	**14.55**	**16.12**	**3.61**	**4.71**	**3.98**	**3.54**	**11.46**	**3.69**	**2.08**
	Actinobacteria	8.57	40.63	31.12	39.81	26.07	2.33	19.30	14.55	16.11	3.61	4.71	3.98	3.54	11.46	3.69	2.08
Bacteroidetes		**12.69**	**40.49**	**13.04**	**9.40**	**15.23**	**24.79**	**13.90**	**11.81**	**15.42**	**2.06**	**1.61**	**2.71**	**3.50**	**6.87**	**5.56**	**4.68**
	Flavobacteria	0.09	27.96	0.16	0.36	0.17	6.91	0.06	0.06	0.08	0.16	0.01	0.05	0.21	0.23	0.02	0.03
	Sphingobacteria	11.19	12.24	12.42	8.20	14.35	17.48	13.50	11.57	15.09	1.76	1.44	2.51	3.15	6.44	5.03	3.76
Chloroflexi		**1.29**	**1.59**	**1.97**	**3.22**	**1.42**	**0.60**	**1.70**	**2.16**	**1.46**	**1.01**	**1.87**	**0.78**	**0.69**	**1.53**	**0.63**	**2.20**
	Anaerolineae	0.18	0.04	0.16	0.29	0.18	0.30	0.09	0.53	0.18	0.25	0.11	0.40	0.25	0.24	0.18	0.51
	Bljii12	0.05	0.00	0.05	0.01	0.02	0.00	0.10	0.14	0.12	0.22	0.25	0.07	0.10	0.39	0.13	0.10
	SOGA31	0.66	0.04	1.26	1.62	0.76	0.05	0.25	0.36	0.31	0.29	0.01	0.10	0.17	0.63	0.16	0.35
	Thermomicrobia	0.14	1.47	0.23	1.19	0.22	0.11	0.11	0.21	0.18	0.01	0.00	0.00	0.00	0.04	0.00	0.00
Cyanobacteria		**0.25**	**0.49**	**0.06**	**0.03**	**1.02**	**25.25**	**14.48**	**9.64**	**2.09**	**0.09**	**0.40**	**0.30**	**0.24**	**0.03**	**0.25**	**0.13**
	Chloroplast	0.01	0.19	0.02	0.00	0.23	0.59	1.32	0.17	0.06	0.03	0.04	0.01	0.03	0.00	0.03	0.00
	Oscillatoriophycideae	0.14	0.17	0.01	0.01	0.48	23.18	12.89	8.99	1.91	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Firmicutes		**0.02**	**0.57**	**0.06**	**0.02**	**0.03**	**0.00**	**1.26**	**0.54**	**0.17**	**3.74**	**0.42**	**0.49**	**0.34**	**0.74**	**0.27**	**0.07**
	Bacilli	0.02	0.41	0.05	0.02	0.03	0.00	1.24	0.44	0.16	3.62	0.39	0.48	0.32	0.73	0.24	0.01
Gemmatimonadetes		**3.32**	**0.14**	**3.96**	**1.62**	**2.35**	**1.06**	**1.31**	**2.12**	**2.37**	**0.55**	**0.84**	**0.77**	**0.45**	**1.40**	**1.81**	**3.24**
Nitrospirae		**0.14**	**0.00**	**0.12**	**0.01**	**0.13**	**0.03**	**0.29**	**0.51**	**0.49**	**1.03**	**0.61**	**0.05**	**0.02**	**0.07**	**0.04**	**0.10**
Planctomycetes		**1.61**	**0.08**	**2.88**	**1.57**	**2.28**	**1.53**	**1.12**	**2.68**	**2.44**	**2.02**	**2.36**	**3.56**	**3.39**	**3.27**	**3.63**	**3.98**
	Phycisphaerae	0.72	0.01	1.71	0.92	1.38	0.46	0.51	0.98	1.16	0.56	0.13	0.57	0.61	1.52	0.98	0.80
	Planctomycea	0.82	0.06	1.14	0.65	0.89	1.07	0.60	1.64	1.27	1.13	2.12	2.76	2.43	1.53	2.42	2.94
Proteobacteria		**11.87**	**5.72**	**9.22**	**5.58**	**10.50**	**23.26**	**15.96**	**19.32**	**24.02**	**19.06**	**30.04**	**30.56**	**28.77**	**23.51**	**31.08**	**28.31**
	Alphaproteobacteria	3.44	3.92	5.52	1.03	6.36	8.63	9.57	9.29	14.46	4.53	10.82	13.24	12.63	11.09	14.83	7.49
	Betaproteobacteria	3.27	0.88	1.86	1.32	1.80	6.33	2.35	2.11	3.45	1.68	3.52	5.27	5.83	5.63	4.26	10.61
	Deltaproteobacteria	1.46	0.11	0.64	0.19	1.02	2.16	2.65	5.18	4.20	10.71	4.25	4.97	5.88	3.38	2.72	4.79
	Gammaproteobacteria	3.18	0.73	0.94	2.27	0.99	5.65	0.91	2.21	1.43	1.14	10.42	6.02	3.34	2.24	8.11	4.76
Thermi		**0.03**	**0.36**	**0.02**	**0.65**	**0.02**	**0.02**	**0.02**	**0.02**	**0.04**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Verrucomicrobia		**4.92**	**0.19**	**12.17**	**4.05**	**13.60**	**10.39**	**2.63**	**4.04**	**3.74**	**40.29**	**5.22**	**17.61**	**34.60**	**28.12**	**14.14**	**9.49**
	Opitutae	0.05	0.04	0.08	0.16	0.27	0.29	0.03	0.12	0.09	0.04	0.05	0.14	0.14	0.10	0.99	0.36
	Spartobacteria	2.94	0.05	11.26	3.56	12.05	1.07	2.06	2.31	2.47	37.46	2.08	14.34	30.73	26.43	8.61	4.79
	Verrucomicrobiae	1.78	0.07	0.61	0.29	1.11	8.92	0.35	1.23	0.82	2.28	2.97	2.54	2.72	1.12	2.87	3.73

Fig. S1. Relative abundances of the major bacterial and archaeal taxa across the 16 soils examined here. Except for the crenarchaeal and euryarcheotal groups, all phyla are bacterial. Numbers in cells indicate percentages, cells colored based on relative abundances (highest relative abundances in red).

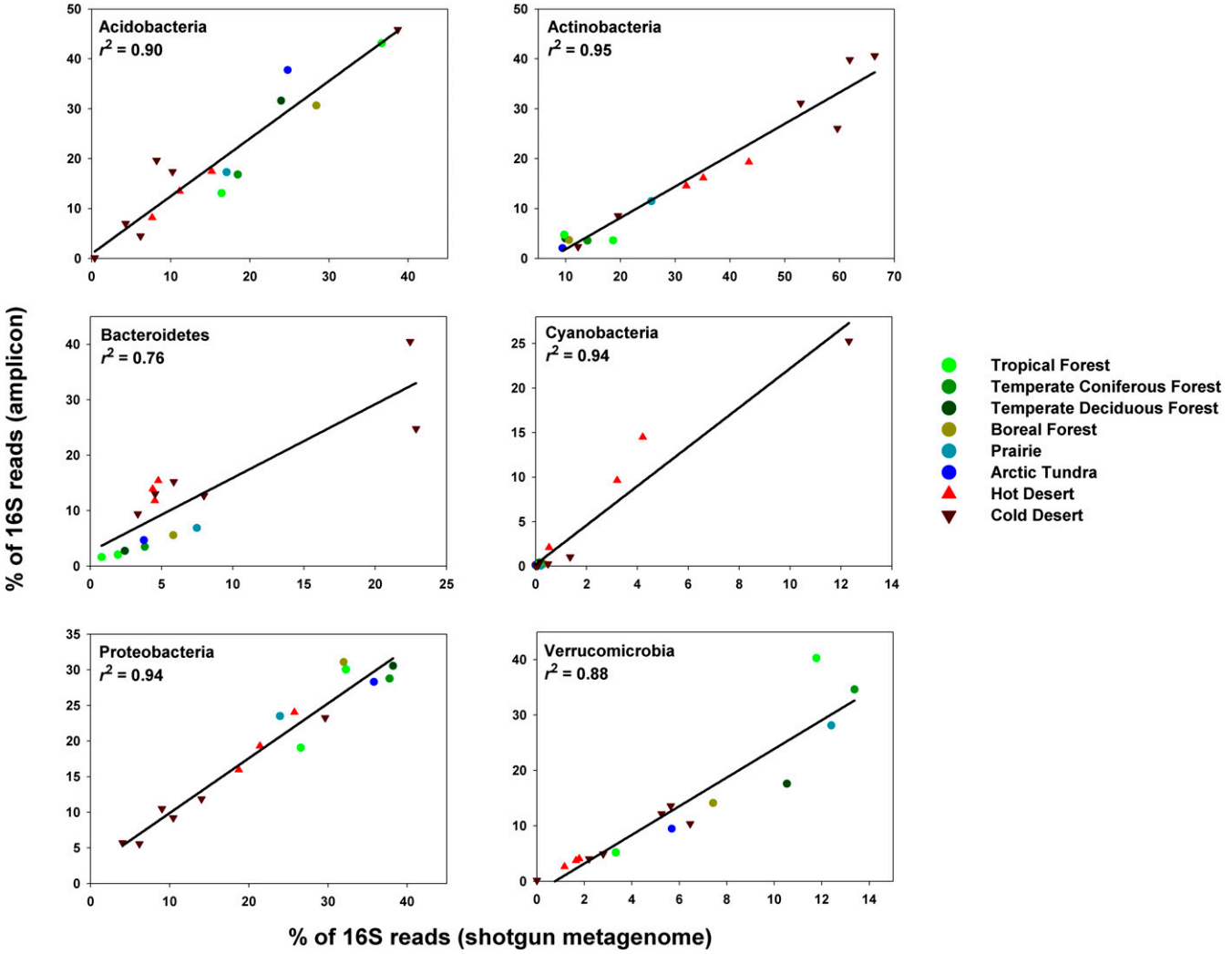

Fig. S2. Comparisons of the relative abundances (percentage of reads) of the dominant soil bacterial phyla determined using 16S rRNA gene data obtained using the shotgun metagenomic approach (*x* axes) versus the 16S rRNA gene data obtained using the amplicon sequencing approach (*y* axes).

	Cold Desert Soils						Hot Desert Soils			Non-Desert Soils						
	EB017	EB019	EB020	EB021	EB024	EB026	MD3	SF2	SV1	AR3	BZ1	CL1	DF1	KP1	PE6	TL1
Amino Acids and Derivatives	15.68	15.52	15.44	15.65	15.24	16.48	15.47	15.43	15.70	15.30	15.58	15.54	15.87	15.88	16.13	15.52
Carbohydrates	15.90	15.96	15.66	15.75	15.26	14.63	17.78	17.72	16.89	16.08	15.76	16.48	16.07	17.89	17.12	16.16
Cell Division and Cell Cycle	2.68	2.58	2.95	3.07	2.78	2.62	2.47	2.50	2.30	2.24	1.78	1.71	1.80	2.02	1.83	1.85
Cell Wall and Capsule	3.37	2.73	2.72	2.56	2.60	3.30	2.78	2.98	2.92	3.08	3.19	3.58	3.41	2.80	3.11	3.91
Clustering-based subsystems	17.53	19.09	18.88	19.88	19.81	18.41	18.48	17.92	17.97	16.92	16.24	15.62	16.11	16.57	15.65	16.41
Cofactors, Vitamins, Prosthetic Groups, Pigments	4.73	6.47	5.57	5.23	5.36	5.64	5.57	5.89	5.63	4.84	5.41	5.80	5.96	5.76	5.42	5.39
DNA Metabolism	4.42	5.18	4.64	4.73	4.57	4.78	4.30	3.62	4.45	2.96	3.88	3.59	3.68	4.02	3.27	3.71
Dormancy and Sporulation	0.03	0.05	0.06	0.08	0.06	0.10	0.07	0.03	0.06	0.01	0.01	0.01	0.03	0.02	0.01	0.02
Fatty Acids, Lipids, and Isoprenoids	1.19	1.48	1.29	1.32	1.32	2.01	1.44	1.35	1.38	1.24	1.57	1.53	1.39	1.32	1.22	1.54
Iron acquisition and metabolism	0.11	0.71	0.18	0.15	0.13	0.17	0.20	0.26	0.29	0.28	0.18	0.22	0.19	0.22	0.14	0.14
Membrane Transport	1.95	1.46	1.65	1.45	1.86	1.63	1.72	1.69	1.88	2.49	3.02	2.65	3.04	2.15	1.92	2.64
Metabolism of Aromatic Compounds	0.63	0.48	0.47	0.39	0.42	0.66	0.82	0.73	0.92	1.37	1.22	1.44	1.39	1.20	1.21	1.18
Miscellaneous	3.41	3.38	3.15	2.96	3.36	3.93	4.14	3.89	4.28	3.94	3.89	4.14	3.95	4.16	4.39	3.76
Motility and Chemotaxis	0.03	0.09	0.08	0.10	0.09	0.12	0.19	0.06	0.26	0.06	0.38	0.34	0.37	0.19	0.17	0.29
Nitrogen Metabolism	0.43	0.40	0.45	0.31	0.77	0.62	0.71	0.63	0.69	0.98	0.78	0.94	0.98	0.94	0.96	0.99
Nucleosides and Nucleotides	3.10	2.85	3.32	3.15	3.11	2.91	2.88	3.17	2.92	2.90	2.59	2.42	2.50	2.64	2.43	2.35
Phages, Prophages, Transposable elements, Plasmids	2.10	2.33	2.27	1.88	1.92	1.87	2.17	2.01	1.87	1.53	1.82	1.64	1.62	1.52	1.20	1.98
Phosphorus Metabolism	1.15	0.54	1.01	1.01	0.94	0.63	0.90	1.10	0.97	1.08	0.88	0.82	0.88	0.83	0.89	0.78
Photosynthesis	0.01	0.01	0.00	0.00	0.02	0.18	0.05	0.06	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Potassium metabolism	0.15	0.06	0.14	0.01	0.15	0.21	0.10	0.17	0.23	0.62	0.56	0.62	0.66	0.41	0.59	0.90
Protein Metabolism	10.55	9.60	10.87	11.54	10.76	9.36	9.64	9.89	9.88	8.86	7.73	7.02	7.02	8.15	7.90	7.22
Regulation and Cell signaling	0.08	0.19	0.10	0.10	0.12	0.19	0.14	0.15	0.15	0.13	0.20	0.17	0.20	0.14	0.19	0.23
Respiration	3.75	3.74	3.78	4.06	3.82	2.91	3.58	3.80	3.51	4.20	3.61	3.80	3.89	3.69	3.99	3.88
RNA Metabolism	2.82	2.25	2.77	2.70	2.88	2.86	2.47	2.60	2.66	2.18	2.27	2.10	1.91	2.19	2.41	2.28
Secondary Metabolism	0.20	0.14	0.18	0.15	0.28	0.29	0.19	0.23	0.20	0.25	0.22	0.28	0.24	0.24	0.22	0.19
Stress Response	1.16	0.85	0.96	1.11	0.99	0.86	0.93	1.14	1.00	1.39	1.36	1.42	1.36	1.28	0.93	1.35
Sulfur Metabolism	0.24	0.25	0.30	0.28	0.32	0.26	0.44	0.38	0.47	0.56	0.65	0.62	0.71	0.67	0.58	0.55
Virulence, Disease and Defense	2.61	1.62	1.11	0.36	1.05	2.39	0.39	0.60	0.50	4.51	5.20	5.49	4.77	3.10	6.09	4.80

Fig. S3. The relative abundances of functional gene categories as determined from the shotgun metagenomic data with gene categories defined at the lowest level of resolution. Values in each cell indicate percent abundance with the color of each cell indicating the *z* score for that particular gene category (blue colors, negative *z* scores; red colors, positive *z* scores).

Fig. S4. Specific gene categories that are relatively more or less abundant in desert soils than in soils from other, nondesert biome types. Only those gene categories that are significantly different between the two categories of samples (Bonferroni corrected *P* values <0.05) are shown here; *x* axis shows the relative percentage difference in abundance compared with the nondesert soils. Dark red symbols indicate cold desert soils; light red symbols indicate hot desert soils.

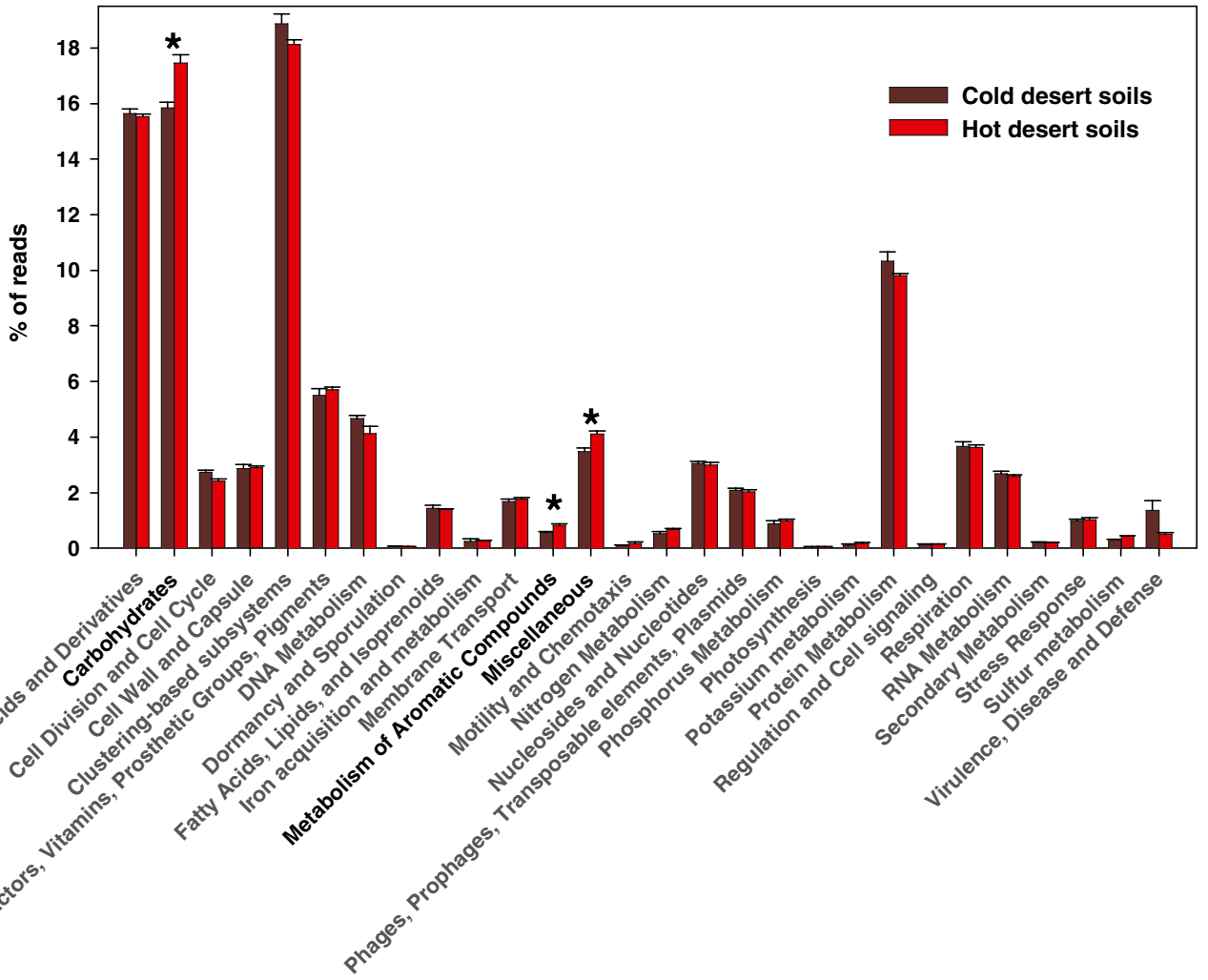

Fig. S5. The relative abundances of major categories of functional genes in metagenomes from cold desert soils from the McMurdo Dry Valley sites versus those soils collected from hot deserts in North America. Asterisks and bold type indicate those categories with significantly different relative abundances between the two desert types (Bonferroni corrected *P* values <0.05, uncorrected *P* values <0.002).

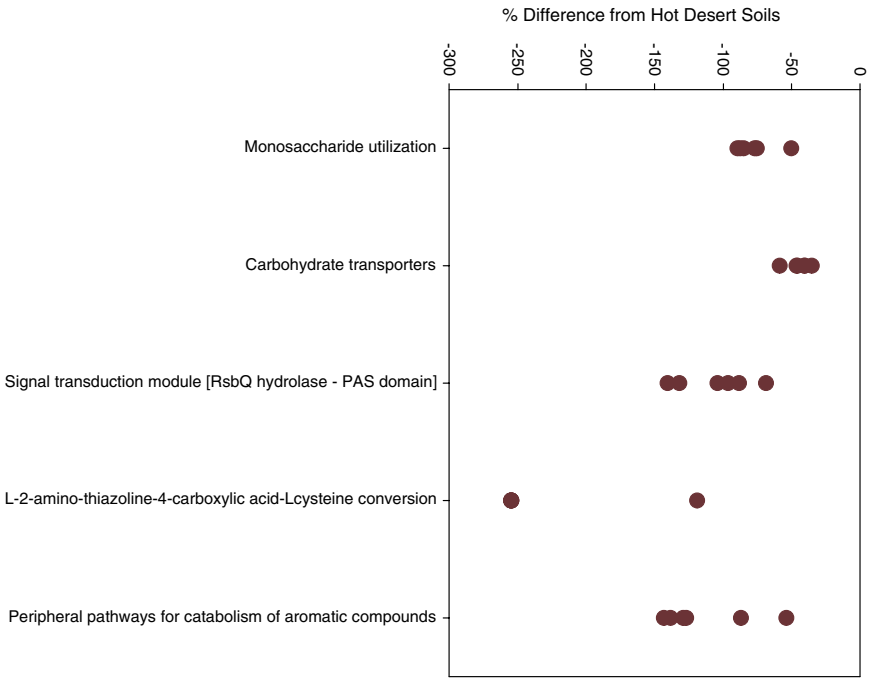

Fig. S6. Specific gene categories that are relatively less abundant in cold desert soils than in soils from the hot deserts. Only those gene categories that are significantly different between the two categories of samples (Bonferroni corrected *P* values <0.05) are shown here. At this level of resolution, no individual gene categories were significantly more abundant in the cold desert soils than in the hot desert soils.

Table S1. Characteristics of the 16 soils included in this study

Biome type	Sample ID	Location description	Latitude, °	Longitude, °	Elevation, MASL	Dominant plant species	MAT	MAP	% organic C	% N	C:N ratio	% silt + clay	pH
Polar desert	EB017	Garwood Valley, Antarctica	−78.03	163.87	353	No plants	−19	100	0.070	0.006	12.2	~10	9.50
Polar desert	EB019	Lake Bonney Valley, Antarctica	−77.73	162.31	109	No plants	−19	100	0.040	0.007	5.6	~10	8.73
Polar desert	EB020	Lake Fryxell Valley, Antarctica	−77.61	163.25	11	No plants	−19	100	0.050	0.005	10.8	~10	9.26
Polar desert	EB021	Lake Hoare Valley, Antarctica	−77.64	162.89	187	No plants	−19	100	0.020	0.003	6.7	~10	9.95
Polar desert	EB024	Wright Valley, Antarctica	−77.53	161.70	109	No plants	−19	100	0.080	0.013	6.1	~10	8.15
Polar desert	EB026	Lake Bonney Valley, Antarctica	−77.73	162.31	114	No plants	−19	100	0.160	0.010	15.4	~10	8.87
Hot desert	MD3	Mojave Desert, California	34.90	−115.65	967	*Larrea tridentata, Ambrosia dumosa*	21.0	150	0.12	0.018	6.8	19	7.90
Hot desert	SF2	Chihuahuan Desert, Galisteo, New Mexico	35.38	−105.93	1500	*Atriplex canescens, Sporobolus airoides*	13	250	0.42	0.052	8.14	58	8.38
Hot desert	SV1	Chihuahuan Desert, Sevilleta LTER, New Mexico	34.33	−106.73	1480	*Larrea tridentata*	13.5	210	0.30	0.060	5.0	22	8.31
Tropical forest	AR3	Misiones, Argentina	−26.73	−54.68	20	*Tabebuia* sp., *Luehea* sp.	23	1,400	2.44	0.248	9.844	82	5.90
Boreal forest	BZ1	Bonanza Creek LTER, Alaska	64.80	−148.25	300	*Picea glauca*	−2.9	260	3.03	0.156	19.4	80	5.12
Temperate deciduous forest	CL1	Calhoun Experimental Forest, South Carolina	34.62	−81.67	150	*Quercus* spp., *Carya* spp., *Acer rubrum*	15.9	1,250	2.33	0.119	19.6	35	5.68
Temperate coniferous forest	DF1	Duke Forest, North Carolina	35.97	−79.08	163	*Pinus taeda, Liquidambar styraciflua*	14.6	1,100	2.78	0.074	37.7	43	5.37
Temperate grassland	KP1	Konza Prairie LTER, Kansas	39.10	−96.60	100	*Andropogon gerardii, Sorghastrum nutans*	12.5	835	6.12	0.450	13.6	78	6.37
Tropical forest	PE6	Manu National Park, Peru	−12.63	−71.23	440	*Parkia* sp., *Virola* sp.	25	4,000	3.34	0.330	10.12	65	4.12
Arctic tundra	TL1	Toolik Lake LTER, Alaska	68.63	−149.58	894	*Eriophorum vaginatum*	−9.3	400	7.02	0.375	18.7	57	4.58

MAP, mean annual precipitation (mm H_2O); MASL, meters above sea level; MAT, mean annual temperature (°C). Soil organic carbon and nitrogen concentrations were measured on a CE Elantech model NC2100 elemental analyzer (ThermoQuest Italia, Milan). Soil pH was measured after shaking a soil/water (1:1, wt/vol) suspension for 30 min. With the exception of the soils from Antarctica, where soil texture was estimated from values reported in Hunt et al. (1), particle size analyses were conducted at the Division of Agriculture and Natural Resources Analytical Laboratory (Davis, CA) using standard methods.

1. Hunt HW, Fountain AG, Doran PT, Basagic H (2010) A dynamic physical model for soil temperature and water in Taylor Valley, Antarctica. *Antarct Sci* 22:419–434.

Table S2. Diversity metrics, MG-RAST IDs, and percent of metagenomic reads annotated

Biome type	Sample ID	MG-RAST ID	% of quality reads annotated	Metagenomic richness (S)	Metagenomic diversity (H′)	Bacterial 16S richness (S)	Bacterial 16S diversity (H′)	Bacterial 16S phylogenetic diversity (PD)
Polar desert	EB017	4477900.3	14.5	1,535	6.39	4,527	5.31	300.0
Polar desert	EB019	4477901.3	23.6	1,663	6.52	2,796	3.60	261.1
Polar desert	EB020	4477902.3	17.3	1,376	6.33	4,936	5.79	305.6
Polar desert	EB021	4477903.3	15.9	1,228	6.17	2,845	4.57	195.3
Polar desert	EB024	4477904.3	17.2	1,386	6.34	4,124	5.56	270.0
Polar desert	EB026	4477803.3	20.5	2,231	6.78	2,935	4.92	232.9
Hot desert	MD3	4477805.3	16.4	1,948	6.60	8,895	6.72	485.8
Hot desert	SF2	4477872.3	14.4	1,850	6.56	10,078	6.93	554.4
Hot desert	SV1	4477873.3	17.3	1,981	6.68	9,929	7.14	527.4
Tropical forest	AR3	4477875.3	13.3	1,814	6.51	9,264	5.72	537.1
Boreal forest	BZ1	4477876.3	17.5	2,270	6.79	9,002	6.54	512.9
Temperate deciduous forest	CL1	4477877.3	18.2	2,393	6.81	12,352	7.06	675.0
Temperate coniferous forest	DF1	4477899.3	18.3	2,414	6.81	12,150	6.68	664.6
Temperate grassland	KP1	4477804.3	17.2	2,193	6.72	10,253	6.60	557.4
Tropical forest	PE6	4477807.3	15.6	2,317	6.70	8,772	6.66	476.8
Arctic tundra	TL1	4477874.3	18.8	2,375	6.84	6,965	6.27	437.6

Percentages of quality-filtered shotgun metagenomic reads that could be annotated to functional gene categories and diversity indices calculated from both the shotgun metagenomic data and the 16S rRNA gene amplicon data.

Table S3. Overall structure of the soil microbial communities

		Cold desert						Hot desert			Other nondesert biomes						
		EB017	EB019	EB020	EB021	EB024	EB026	MD3	SF2	SV1	AR3	BZ1	CL1	DF1	KP1	PE6	TL1
SSU	Archaea	0.30	0.04	0.20	0.16	0.31	0.09	1.85	3.06	3.08	3.37	0.14	0.00	0.00	1.28	0.53	0.00
	Bacteria	80.24	94.34	94.80	96.60	95.24	95.30	90.62	85.47	87.97	85.56	73.64	76.63	76.99	86.84	87.76	89.81
	Fungi (all)	7.54	0.15	0.85	0.21	1.09	1.21	2.36	3.93	3.26	2.56	11.21	10.33	9.32	3.54	3.63	3.53
	Fungi (Basidiomycota)	0.30	0.00	0.12	0.05	0.07	0.09	0.32	1.02	0.28	0.54	6.45	5.48	3.01	0.10	0.30	0.71
	Fungi (Ascomycota)	7.24	0.15	0.73	0.16	0.99	1.12	1.99	2.75	2.90	2.02	4.69	4.53	5.89	3.24	3.10	2.52
	Streptophyta	1.48	4.46	1.30	0.63	1.06	1.52	1.15	1.78	0.89	2.83	1.83	1.27	2.05	3.63	1.89	1.51
	Other eukaryotes	2.91	0.86	1.99	2.20	1.23	0.67	1.71	1.99	1.61	3.10	2.04	1.75	2.74	1.38	2.79	1.92
LSU	Archaea	0.06	0.00	0.04	0.00	0.03	0.03	0.11	0.35	0.42	1.24	0.00	0.00	0.11	0.24	0.18	0.07
	Bacteria	91.82	98.84	96.47	97.33	95.91	94.03	92.18	89.49	91.19	89.74	74.78	75.39	81.72	92.82	90.21	89.75
	Fungi (all)	2.38	0.12	0.63	0.30	0.95	0.80	2.75	3.42	2.87	2.22	11.01	10.50	7.88	2.02	3.02	4.07
	Fungi (Basidiomycota)	0.31	0.00	0.00	0.00	0.00	0.03	0.07	0.39	0.42	0.62	7.29	6.66	3.47	0.32	0.67	1.38
	Fungi (Ascomycota)	1.53	0.12	0.60	0.30	0.66	0.62	2.28	2.86	2.38	1.48	3.41	3.42	3.68	1.37	1.51	2.69
	Streptophyta	2.44	0.45	1.26	1.39	1.84	3.45	1.99	2.77	2.19	3.83	2.25	3.06	1.79	2.02	3.08	1.45
	Other eukaryotes	1.47	0.47	1.00	0.69	0.60	1.05	0.62	0.74	0.53	0.87	1.26	0.96	1.37	1.21	1.33	0.58

Percentages of small-subunit (SSU) and large-subunit (LSU) rRNA gene reads from the shotgun metagenomic data that could be assigned to major taxonomic groups. Category "Other eukaryotes" includes Mollusca, Nematoda, Arthropoda, Rotifera, and Annelida.