Temperature and soil moisture control microbial community composition in an arctic-alpine ecosystem along elevational and micro-topographic gradients

K. Frindte[1], K. Werner[1,3], R. Pape[2], J. Löffler[2], C. Knief[1]

Temperature dependency of alpine microbes

Corresponding author: Katharina Frindte, kfrindte@uni-bonn.de, Tel.: 0049228732979

[1] Institute of Crop Science and Resource Conservation (INRES), Molecular Biology of the Rhizosphere, Nussallee 13, 53115 Bonn, Germany

[2] Department of Geography, University of Bonn, Meckenheimer Allee 166, 53115 Bonn, Germany

[3] Beuth Hochschule für Technik, Seestraße 64, 13347 Berlin, Germany

Detailed information about Illumina sequence analysis

Sequencing resulted in 5.36 million paired end reads. The reads were assembled with PANDAseq [1], and demultiplexed with a Python program (kindly provided by Kurt Stüber, Max Planck-Genome-centre Cologne). Quality check revealed that all sequences had a phred score > 3, as recommended by Bokulich *et al.* [2]. Primer sequences were removed via Cutadapt [3]. After assembly and quality check, 3.06 million reads remained. Sequence alignment was done with PyNAST [4]. Afterwards sequences were clustered into OTUs using UCLUST and the SILVA v.119 SSU reference database (97 % consensus; [5]). Chimeras were removed with USEARCH 6.1 [5] using a Greengenes database (gg_97_otus_4feb2011) as reference. Moreover, singletons, reads assigned to unclassified groups of organisms, chloroplasts and mitochondria were removed from the data set, finally resulting in 2.81 million reads (accounting for roughly 52 % of all generated sequence reads). BIOM files were converted into readable tables for further analyses.

Detailed information about PLSR

PLSR, which is also known as 'projection on latent structures' [6], combines elements from PCA and multiple linear regression (MLR). As a limited information approach, introduced by Wold as a soft modeling technique to handle various modeling problems in situations where it is difficult or impossible to meet the hard assumptions of more traditional statistics, PLSR has the advantage that it works without distributional assumptions and with nominal, ordinal, and interval scaled variables [7–10]. Moreover, PLSR deals efficiently with unreliability and heteroscedasticity issues [11]. The method is well suited for the analysis of datasets in which the number of explanatory variables exceeds the number of observations [12] and/or in case of highly correlated predictors [13, 14]. Instead of using all independent variables simultaneously as predictors, the PLSR extracts just a few components (latent factors) from the independent variables that maximize not only the explained variance in the independent variables (like PCA), but also in the dependent variable. Thus, PLSR uses fewer explanatory variables than were in the original data. Further, PLSR was described as more reliable than PCA-regression and multiple regression when identifying relevant variables, especially in cases of small sample size, and strongly shielded against both type I and type II errors [14]. Thus, PLSR is very useful where emphasis is on theory development instead of testing, in a confirmatory sense, how well a theoretical model fits observed data [15]. Our motivation to apply PLSR was the need to identify – out of all explanatory variables – a subset of relevant variables that is responsible for explaining the variation in the response [16]. In this context of variable selection, a PLSR model is built and its output is solely used to assess the (relative) importance of each explanatory variable, i.e., the focus is to determine the influential explanatory variables rather than the response [17].

References for methodological details

1. Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD. PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics* 2012; **13**: 31.
2. Bokulich NA, Subramanian S, Faith JJ, Gevers D, Gordon JI, Knight R, et al. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat Methods* 2013; **10**: 57–9.
3. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 2011; **17**: 10–12.
4. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010; **7**: 335–336.
5. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010; **26**: 2460–2461.
6. Abdi H. Partial least squares regression and projection on latent structure regression (PLS Regression). *Wiley Interdiscip Rev Comput Stat* 2010; **2**: 97–106.
7. Wold H. Model construction and evaluation when theoretical knowledge is scarce: Theory and application of partial least squares. In: Kmenta J, Ramsey JB (eds). *Evaluation of econometric models*. 1980. Academic Press, New York, pp 47–74.
8. Vinzi VE, Chin WW, Henseler J, Wang H. Editorial: Perspectives on partial least squares. In: Vinzi VE, Chin WW, Henseler J, Wang H (eds). *Handbook of partial least squares – Concepts, methods and applications*. 2010. Springer, Berlin, pp 1–20.
9. Dijkstra T. Some comments on maximum likelihood and partial least squares methods. *J Econom* 1983; **22**: 67–90.
10. Michael Haenlein AMK. A beginner’s guide to partial least squares analysis, Understanding Statistics’. Statistical Issues in Psychology and Social Sciences, Volume 3. *Underst Stat* 2004; **3**: 283–297.
11. Martens H, Naes T. Multivariate calibration. 1989. John Wiley & Sons, New York.
12. Boulesteix A-L, Strimmer K. Partial least squares: a versatile tool for the analysis of high-dimensional genomic data. *Brief Bioinform* 2006; **8**: 32–44.
13. Wehrens R, Mevik B-H. The pls package: principal component and partial least squares regression in R. *J Stat Softw* 2007; **18**: 1–23.
14. Carrascal LM, Galván I, Gordo O. Partial least squares regression as an alternative to current regression methods used in ecology. *Oikos* 2009; **118**: 681–690.
15. Sosik JJ, Kahai SS, Piovoso MJ. Silver bullet or voodoo statistics? A primer for using the partial

least squares data analytic technique in group and organization research. *Gr Organ Manag* 2009; **34**: 5–36.

16. Mehmood T, Liland KH, Snipen L, Sæbø S. A review of variable selection methods in partial least squares regression. *Chemom Intell Lab Syst* 2012; **118**: 62–69.

17. Mehmood T, Ahmed B. The diversity in the applications of partial least squares: an overview. *J Chemom* 2016; **30**: 4–17.

Table S1: Environmental variables used for PLSR analysis

Exploratory variable and measurement system	Variable used for PLSR
5 years soil temperature ONSET's HOBO loggers (type H21-002) with temperature sensors of type S-TMB-002 (±0.2 °C accuracy) equipped with passively ventilated radiation shields. 1 minute interval	5 years soil temperature amplitude
	5 years soil temperature maximum
	5 years soil temperature minimum
	5 years soil temperature mean
	5 years soil temperature -16.9 °C - 22.4 °C (0.1 °C steps; resulting in potential individual temperature thresholds)
2 years soil temperature ONSET's HOBO loggers (type H21-002) with temperature sensors of type S-TMB-002 (±0.2 °C accuracy) equipped with passively ventilated radiation shields. 1 minute interval	2 years soil temperature amplitude
	2 years soil temperature maximum
	2 years soil temperature minimum
	2 years soil temperature mean
	2 years soil temperature -16.9 °C - 18.8 °C (0.1 °C steps; resulting in potential individual temperature thresholds)
1 year soil temperature ONSET's HOBO loggers (type H21-002) with temperature sensors of type S-TMB-002 (±0.2 °C accuracy) equipped with passively ventilated radiation shields. 1 minute interval	1 year soil temperature amplitude
	1 year soil temperature maximum
	1 year soil temperature minimum
	1 year soil temperature mean
	1 year soil temperature -16.6°C - 18.8 °C (0.1 °C steps; resulting in potential individual temperature thresholds)
5 years air temperature ONSET's HOBO loggers (type H21-002) with temperature sensors of type S-TMB-002 (±0.2 °C accuracy) equipped with passively ventilated radiation shields. 1 minute interval	5 years air temperature amplitude
	5 years air temperature maximum
	5 years air temperature minimum
	5 years air temperature mean
	5 years air temperature -25.0 °C - 25 °C (0.1 °C steps; resulting in potential individual temperature thresholds)
2 years air temperature ONSET's HOBO loggers (type H21-002) with temperature sensors of type S-TMB-002 (±0.2 °C accuracy) equipped with passively ventilated radiation shields. 1 minute interval	2 years air temperature amplitude
	2 years air temperature maximum
	2 years air temperature minimum
	2 years air temperature mean
	2 years air temperature -25.0 °C - 25 °C (0.1 °C steps; resulting in potential individual temperature thresholds)
1 year air temperature ONSET's HOBO loggers (type H21-002) with temperature sensors of type S-TMB-002 (±0.2 °C accuracy) equipped with passively ventilated radiation shields. 1 minute interval	1 year air temperature amplitude
	1 year air temperature maximum
	1 year air temperature minimum
	1 year air temperature mean
	1 year air temperature -25.0 °C - 25.0 °C (0.1 °C steps; resulting in potential individual temperature thresholds)
5 years soil moisture ONSET's HOBO loggers with soil moisture sensors of type S-SMD-M005 (± 3 % accuracy) 1 minute interval	5 years soil moisture amplitude
	5 years soil moisture maximum
	5 years soil moisture minimum
	5 years soil moisture mean
	5 years soil moisture 0.5 - 0 m^3 water per m^3 soil (0.01 steps; resulting in potential individual temperature thresholds)
2 years soil moisture ONSET's HOBO loggers with soil moisture sensors of type S-SMD-M005 (± 3 % accuracy) 1 minute interval	2 years soil moisture amplitude
	2 years soil moisture maximum
	2 years soil moisture minimum
	2 years soil moisture mean
	2 years soil moisture 0.5 - 0 m^3 water per m^3 soil (0.01 steps; resulting in potential individual temperature thresholds)

Exploratory variable and measurement	Variable used for PLSR
1 year soil moisture ONSET's HOBO loggers with soil moisture sensors of type S-SMD-M005 (± 3 % accuracy) 1 minute interval	1 year soil moisture amplitude
	1 year soil moisture maximum
	1 year soil moisture minimum
	1 year soil moisture mean
	1 year soil moisture 0.5 - 0.0 m³ water per m³ soil (0.01 steps; resulting in potential individual temperature thresholds)
Soil texture Corn size was measured using a Particle Size Analyzer, type L950 (Horiba)	Texture, only mineral components % clay
	Texture, only mineral components % fine silt
	Texture, only mineral components % medium silt
	Texture, only mineral components % corse silt
	Texture, only mineral components % fine sand
	Texture, only mineral components % medium sand
	Texture, only mineral components % coarse sand
	Substrate organic + mineral components % loss on ignition
	Substrate organic + mineral components % clay
	Substrate organic + mineral components % fine silt
	Substrate organic + mineral components % medium silt
	Substrate organic + mineral components % coarse silt
	Substrate organic + mineral components % fine sand
	Substrate organic + mineral components % medium sand
	Substrate organic + mineral components % coarse sand
	% total particles < µm 0.172
	% total particles < µm 0.2
	% total particles < µm 0.226
	% total particles < µm 0.259
	% total particles < µm 0.296
	% total particles < µm 0.339
	% total particles < µm 0.389
	% total particles < µm 0.445
	% total particles < µm 0.51
	% total particles < µm 0.584
	% total particles < µm 0.877
	% total particles < µm 1.005
	% total particles < µm 1.151
	% total particles < µm 1.318
	% total particles < µm 1.51
	% total particles < µm 1.729
	% total particles < µm 2
	% total particles < µm 2.269
	% total particles < µm 2.599
	% total particles < µm 2.976
	% total particles < µm 3.409
	% total particles < µm 3.905
	% total particles < µm 4.472
	% total particles < µm 5.122
	% total particles < µm 5.867
	% total particles < µm 6.3
	% total particles < µm 6.72
	% total particles < µm 7.697
	% total particles < µm 8.816

Exploratory variable	Variable used for PLSR
Soil texture	% total particles < µm 10.097
	% total particles < µm 11.565
	% total particles < µm 13.246
	% total particles < µm 15.172
	% total particles < µm 17.377
	% total particles < µm 20
	% total particles < µm 22.797
	% total particles < µm 26.111
	% total particles < µm 29.907
	% total particles < µm 34.255
	% total particles < µm 39.234
	% total particles < µm 44.938
	% total particles < µm 51.471
	% total particles < µm 58.953
	% total particles < µm 63
	% total particles < µm 67.523
	% total particles < µm 77.34
	% total particles < µm 88.583
	% total particles < µm 101.46
	% total particles < µm 116.21
	% total particles < µm 125
	% total particles < µm 133.103
	% total particles < µm 152.453
	% total particles < µm 174.616
	% total particles < µm 200
	% total particles < µm 229.075
	% total particles < µm 262.376
	% total particles < µm 300.518
	% total particles < µm 344.206
	% total particles < µm 394.244
	% total particles < µm 451.556
	% total particles < µm 517.2
	% total particles < µm 592.387
	% total particles < µm 630
	% total particles < µm 777.141
	% total particles < µm 890.116
	% total particles < µm 1019.51
	% total particles < µm 1167.72
	% total particles < µm 1337.48
	% total particles < µm 1531.91
	% total particles < µm 1754.61
	% total particles < µm 2000
Nitrogen (C:N analyzer)	Nitrogen content %
Carbon (C:N analyzer)	Carbon content %
C:N (C:N analyzer)	Carbon to nitrogen ratio
pH (pH electode)	pH value

Table S2: Sequences of barcoded forward primers targeting the 16S rRNA gene. Primers include a stagger sequence (3 – 5 bp per primer), a barcode (8 bp) and the primer sequence itself. The reverse primer was not modified.

Name	**Stagger + barcode + Primer 515f**	**Name**	**Stagger + barcode + Primer 515f**
515f-BS1	GGGCTAACACCTA **GTGCCAGCMGCCGCGGTAA**	515f-BS25	TCCGTTACGAATC **GTGCCAGCMGCCGCGGTAA**
515f-BS2	TTTAGACGTAGCT **GTGCCAGCMGCCGCGGTAA**	515f-BS26	AGGTCCTACTA **GTGCCAGCMGCCGCGGTAA**
515f-BS3	TGGAGATATAGGA **GTGCCAGCMGCCGCGGTAA**	515f-BS27	GATTGGTCTTC **GTGCCAGCMGCCGCGGTAA**
515f-BS4	GAACACAGTTG **GTGCCAGCMGCCGCGGTAA**	515f-BS28	CAAAAACCGTGT **GTGCCAGCMGCCGCGGTAA**
515f-BS5	CCACCTACAAC **GTGCCAGCMGCCGCGGTAA**	515f-BS29	CCCGAGGTCCTTG **GTGCCAGCMGCCGCGGTAA**
515f-BS6	AAACGTCGGCT **GTGCCAGCMGCCGCGGTAA**	515f-BS30	TGGGTTGTCCC **GTGCCAGCMGCCGCGGTAA**
515f-BS7	TAACGACGTCAA **GTGCCAGCMGCCGCGGTAA**	515f-BS31	GTTTTCATTAGG **GTGCCAGCMGCCGCGGTAA**
515f-BS8	ACCCGCGTTTCG **GTGCCAGCMGCCGCGGTAA**	515f-BS32	TTTATTGATCCGA **GTGCCAGCMGCCGCGGTAA**
515f-BS9	TGGAAGGTCTGAC **GTGCCAGCMGCCGCGGTAA**	515f-BS33	ATTATCGCCAG **GTGCCAGCMGCCGCGGTAA**
515f-BS10	CATGTTTCACT **GTGCCAGCMGCCGCGGTAA**	515f-BS34	CAACCAGGAGGC **GTGCCAGCMGCCGCGGTAA**
515f-BS11	CCCGCTCCAGCCT **GTGCCAGCMGCCGCGGTAA**	515f-BS35	GGCCGAACTGT **GTGCCAGCMGCCGCGGTAA**
515f-BS12	GCCGGTGCGGTTA **GTGCCAGCMGCCGCGGTAA**	515f-BS36	AAAACTAGTCAT **GTGCCAGCMGCCGCGGTAA**
515f-BS13	CAATTGCAGCCTC **GTGCCAGCMGCCGCGGTAA**	515f-BS37	ATTATGAGTTAAC **GTGCCAGCMGCCGCGGTAA**
515f-BS14	CCCGGGGCGAGGA **GTGCCAGCMGCCGCGGTAA**	515f-BS38	TCCCGCTGGCGA **GTGCCAGCMGCCGCGGTAA**
515f-BS15	CTAGTGGGATA **GTGCCAGCMGCCGCGGTAA**	515f-BS39	GTTGTAGAGCT **GTGCCAGCMGCCGCGGTAA**
515f-BS16	GAATTTATCTCCG **GTGCCAGCMGCCGCGGTAA**	515f-BS40	GAATCTACTGCGC **GTGCCAGCMGCCGCGGTAA**
515f-BS17	CAATGACTAACTG **GTGCCAGCMGCCGCGGTAA**	515f-BS41	GCCGTCGCGTAC **GTGCCAGCMGCCGCGGTAA**
515f-BS18	GAAAATCCTATT **GTGCCAGCMGCCGCGGTAA**	515f-BS42	AGGTTGTAGGTC **GTGCCAGCMGCCGCGGTAA**
515f-BS19	CTTCACGTGTT **GTGCCAGCMGCCGCGGTAA**	515f-BS43	TTTAAAGCGGTC **GTGCCAGCMGCCGCGGTAA**
515f-BS20	TCACCTTTACA **GTGCCAGCMGCCGCGGTAA**	515f-BS44	GGGCCACTCTAAG **GTGCCAGCMGCCGCGGTAA**
515f-BS21	TCCGGCTAGATTC **GTGCCAGCMGCCGCGGTAA**	515f-BS45	AGCTGAGAGTG **GTGCCAGCMGCCGCGGTAA**
515f-BS22	GGGCAGAGAACTC **GTGCCAGCMGCCGCGGTAA**	515f-BS46	CAATCTTCTGATG **GTGCCAGCMGCCGCGGTAA**
515f-BS23	GTTAGCTCAGTT **GTGCCAGCMGCCGCGGTAA**	515f-BS47	AGGCCACAGTGCA **GTGCCAGCMGCCGCGGTAA**
515f-BS24	GAATGGTACTTGC **GTGCCAGCMGCCGCGGTAA**	515f-BS48	TATAGTAGTGG **GTGCCAGCMGCCGCGGTAA**

Table S3: Bacterial families showing significant changes in relative abundance in dependence on elevation or micro-topography. Only families with > 1 % read abundance are shown in this table. Responsive families were identified using Kruskal-Wallis tests in the STAMP program.

Taxa		Alpine belt		Micro-topography		
Family	**Phyla**	Higher in	*p*	Higher in	Compared to	*p*
Coxiellaceae	*Proteobacteria*			all	depression	<0.001
Pedosphaeraceae	*Verrucomicrobia*					
Hyphomicrobiaceae	*Proteobacteria*					
Thermogemmatisporaceae	*Chloroflexi*	middle	< 0.001			
Ktedonobacteraceae	*Chloroflexi*	middle	<0.001			
JG37 AG 4	*AD3*	middle	<0.001			
DS18	*Acidobacteria*	middle	<0.001			
Unclassified *Acidomicrobiales*	*Acidobacteria*	low	<0.001			
Mycobacteriaceae	*Actinobacteria*	low	<0.001			
Gemmataceae	*Planctomycetes*	middle	<0.001	all	depression	<0.001
Nitrosophaeraceae	*Crenarchaeota*	middle	<0.001			
Betaproteobacteria - EB1003	*Proteobacteria*	middle	<0.001	slopes	ridge	<0.001
Isosphaeraceae	*Planctomycetes*	low	<0.001	all	depression	<0.001
auto67_4W	*Verrucomicrobia*	low	0.0015			
iii1-15	*Acidobacteria*	middle	0.0025			
TM7-1	*Candidatus* Saccharibacteria	low	0.0025			
Methylocystaceae	*Proteobacteria*	low	< 0.001			
Solibacterales	*Actinobacteria*	middle	0.003			
Acidimicrobiales	*Actinobacteria*	low	0.003			
Unclassified *Solirubrobacterales*	*Actinobacteria*	low	0.004			
Acetobacteraceae	*Proteobacteriaceae*	low	0.0047	slopes	depression	<0.001
Conexibacteraceae	*Actinobacteria*	middle	0.048	ridge	depression / south slope	<0.001

Taxa		Alpine belt		Micro-topography		
Family	**Phyla**	Higher in	*p*	Higher in	Compared to	*p*
Unclassified Actinomycetales	*Actinobacteria*	low	0.008			
WD2101	*Planctomycetes*			ridge	all	<0.001
Ellin6529	*Chloroflexi*			depression	all	<0.001
Unclassified *Methylacidiphilales*	*Verrucomicrobia*	low	0.017	ridge	all	<0.001
Unclassified *Methylacidiphilales*	*Verrucomicrobia*			north slopes	depression	<0.001
Geobacteraceae	*Proteobacteria*			depression	all	<0.001
Unclassified *Actinomyceteales*	*Actinobacteria*			all	depression	<0.001
Sinobacteraceae	*Proteobacteria*			all	depression	<0.001
Comamonadaceae	*Proteobacteria*			depression	All	<0.001
unclassified WPS-2	*WPS-2*	middle	0.020	all	depression	<0.001
Burkholderiaceae	*Proteobacteria*			ridge	depression	<0.001
Ellin6513	*Acidobacteria*			all	depression	<0.001
Unclassified Acidobacteriaceae	*Acidobacteria*			ridge/ north slope	depression	<0.001
Pirellulaceae	*Planctomycetes*			depression	slopes	<0.001
Alphaproteobacteria - Ellin329	*Proteobacteria*			all	ridge	<0.001
Unclassified *Myxococcales*	*Proteobacteria*			depression	slopes	<0.001
Sphingobacteriaceae	*Proteobacteria*			ridge	depression / south slope	<0.001
Bradyrhizobiaceae	*Proteobacteria*			slopes	depression	<0.001
Synthrophaceae	*Proteobacteria*			Slopes	depression	<0.001
Koribacteraceae	*Acidobacteria*			south slope	ridge/ depression	<0.001

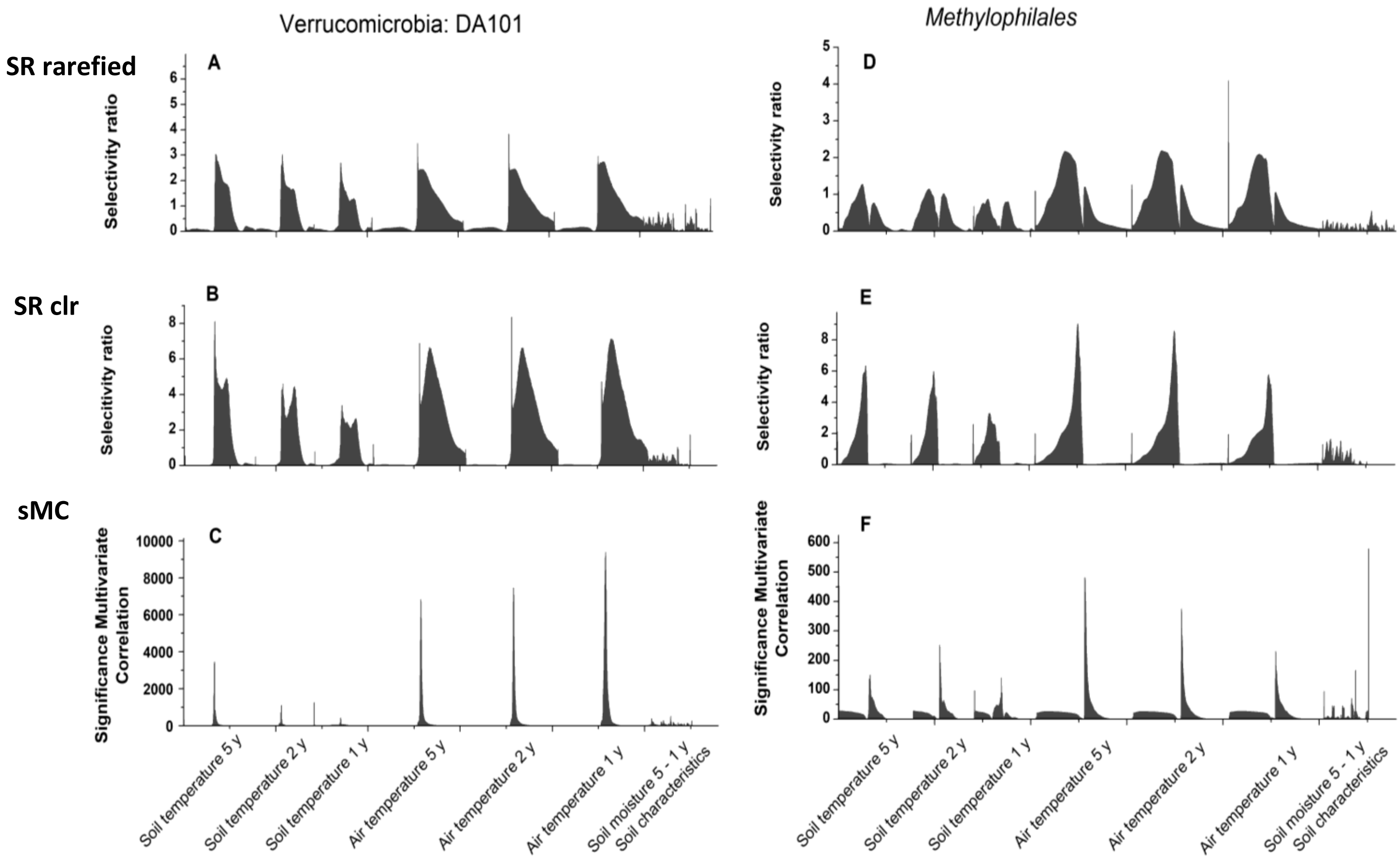

Figure S1: Ecological response graphs showing exemplarily the ecological response patterns of the verrucomicrobial genus DA101 (left) and unclassified taxa of the order *Methylophilales* (right). Plots display the selectivity ratios (SRs) of PLSR calculated using a rarefied table of taxonomic groups (TG) (A + D), SRs calculated with a centered log-ratio (clr) transformed table (B + E) and significance Multivariate Correlation values (sMC) for a PLSR calculated using the rarefied TG table (C + F). Plots display SR and sMC values of all environmental variables arranged along the x-axis. The first three sections represent soil temperature measures over 1, 2 and 5 years, respectively. Within each temperature section, stepwise increasing threshold values ranging from -16.9 °C to +22.4 °C are shown, followed by single values for amplitude, maximum, minimum and mean. The same pattern is used to present air temperature in a range from -25 °C to +25 °C and soil moisture ranging from 0 to 0.5 m^3 H_2O m^{-3} soil. The last section named soil characteristics includes soil texture, C, N, C:N ratio and pH values, according to table S1.

Exemplarily, the plots for DA101 show a TG with consistent results for all PLSR analyses, while the plots for *Methylophilales* are less reproducible based on the different algorithms applied to assess the impact of the environmental factors. In the majority of TGs, the patterns were in general well reproducible, especially the category of the most important driver (i.e. temperature, soil moisture or soil characteristics) remained identical for approx. 70% of all taxa (table S4). However, the specific most relevant driver (e.g., 2-year versus 5-year moisture regime) was not necessarily identical when applying different approaches, due to the fact that SR and sMC values were often in a similar range within a category. Taken together the findings for all analyzed TGs, overall findings about the most important drivers for the whole microbial community were the in good agreement (Figure 4), while the results for individual TGs differ to some extent based on the applied algorithm for environmental parameter identification.

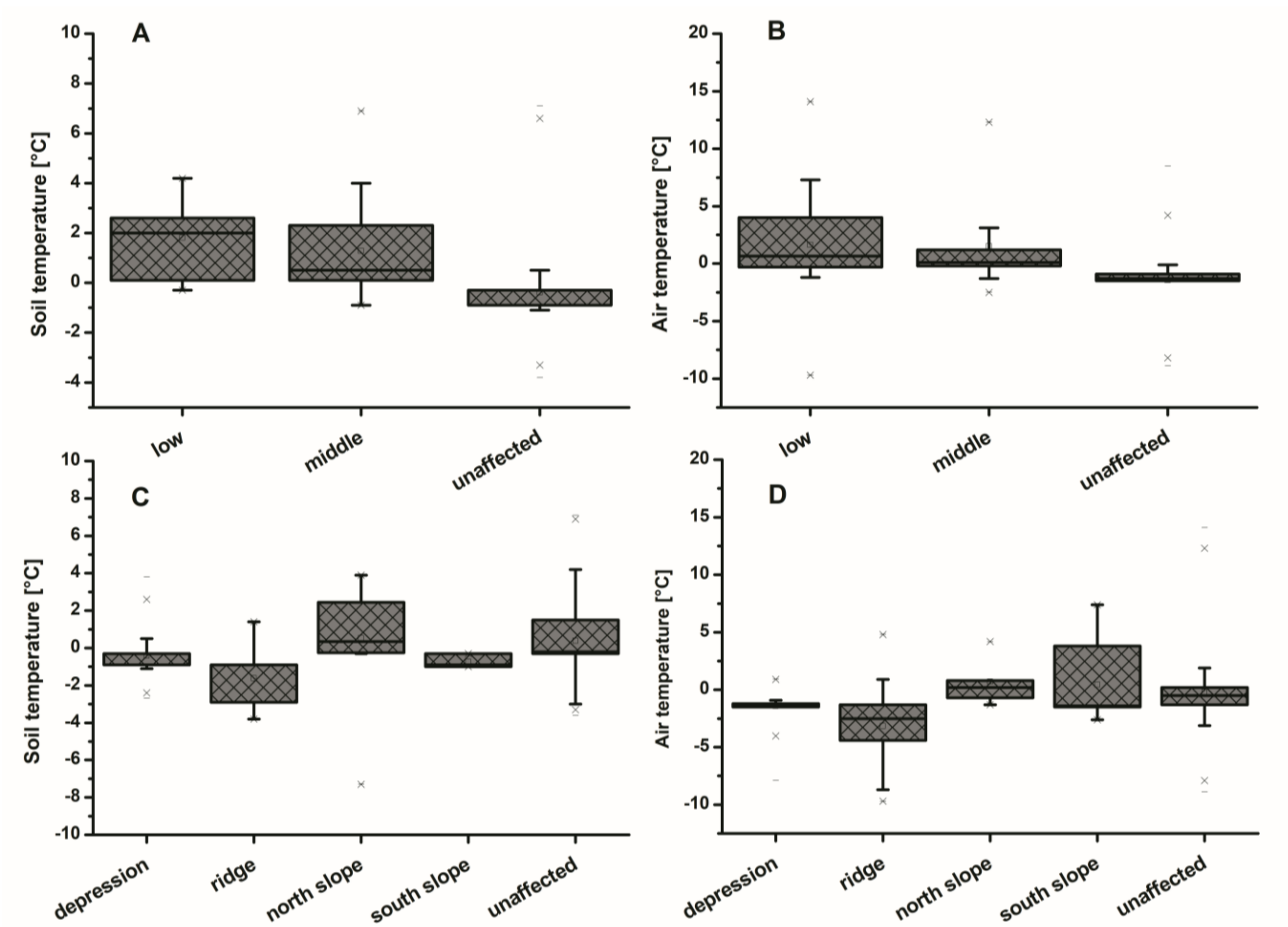

Figure S2: Boxplots displaying soil (A + C) and air temperature (B + D) selectivity value thresholds > 2 for taxa being affected by elevation (A + B), micro-topography (C + D), or being unaffected by elevation or micro-topography, respectively. Displayed are the median and the 25th and 75th percentile as middle line and grey box. Whiskers indicate minimum and maximum.

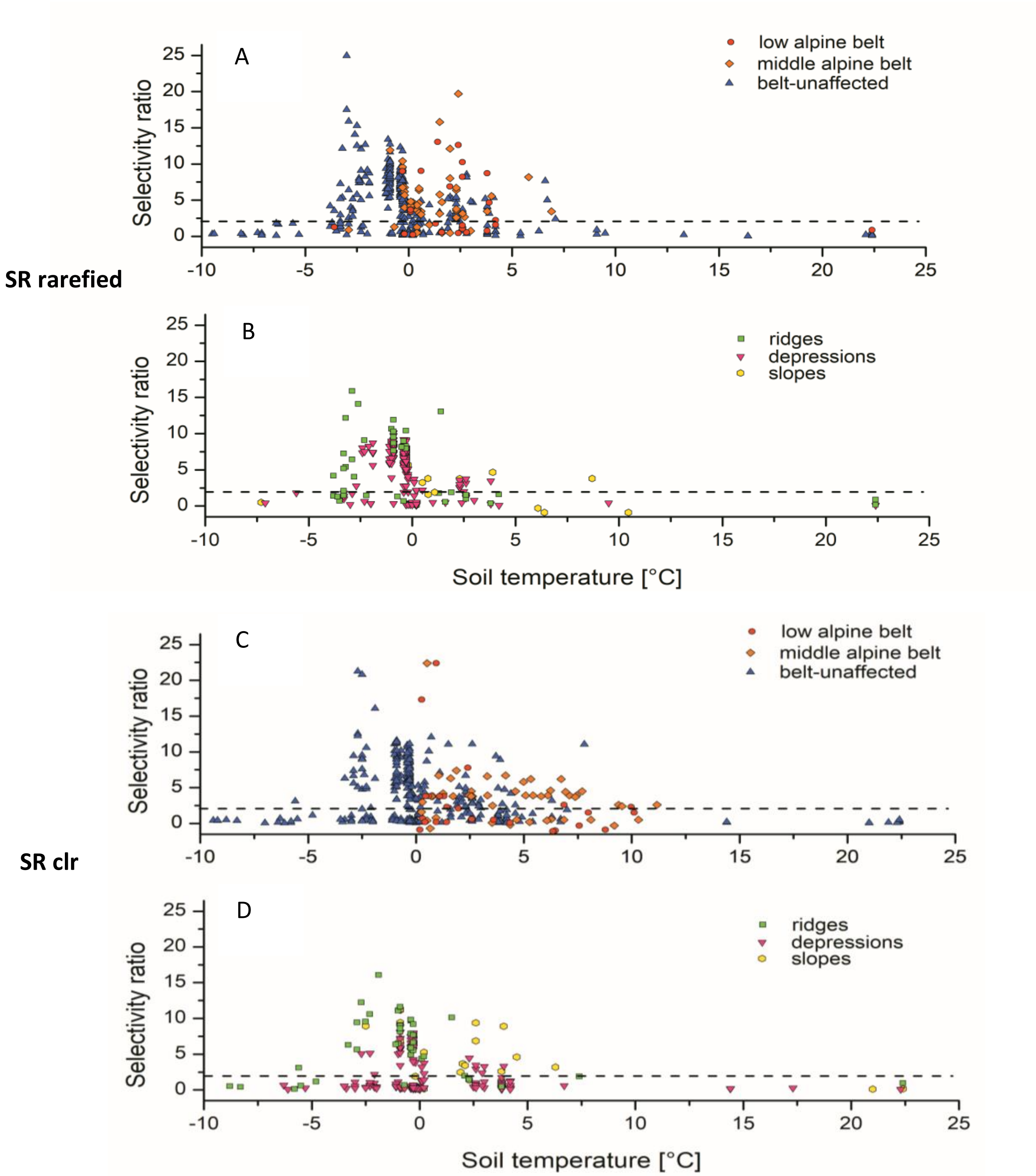

Figure S3: Plots displaying the maximum selectivity ratio (SR) in dependence on the 5-year soil temperature threshold values of each taxonomic group (TG) evaluated based on a rarefied TG dataset (A + B) or a centered log-ratio transformed dataset (C + D). Different colors indicate TGs that were significantly affected or unaffected by elevation (A + C) or affected by different micro-topographic expositions (B + D) according to STAMP. The dotted line indicates the SR cut-off at the value of 2, below which variables were considered to have no explanatory effect. The comparison of SR values derived from a rarefied dataset are largely in agreement with those derived from a clr-transformed dataset.

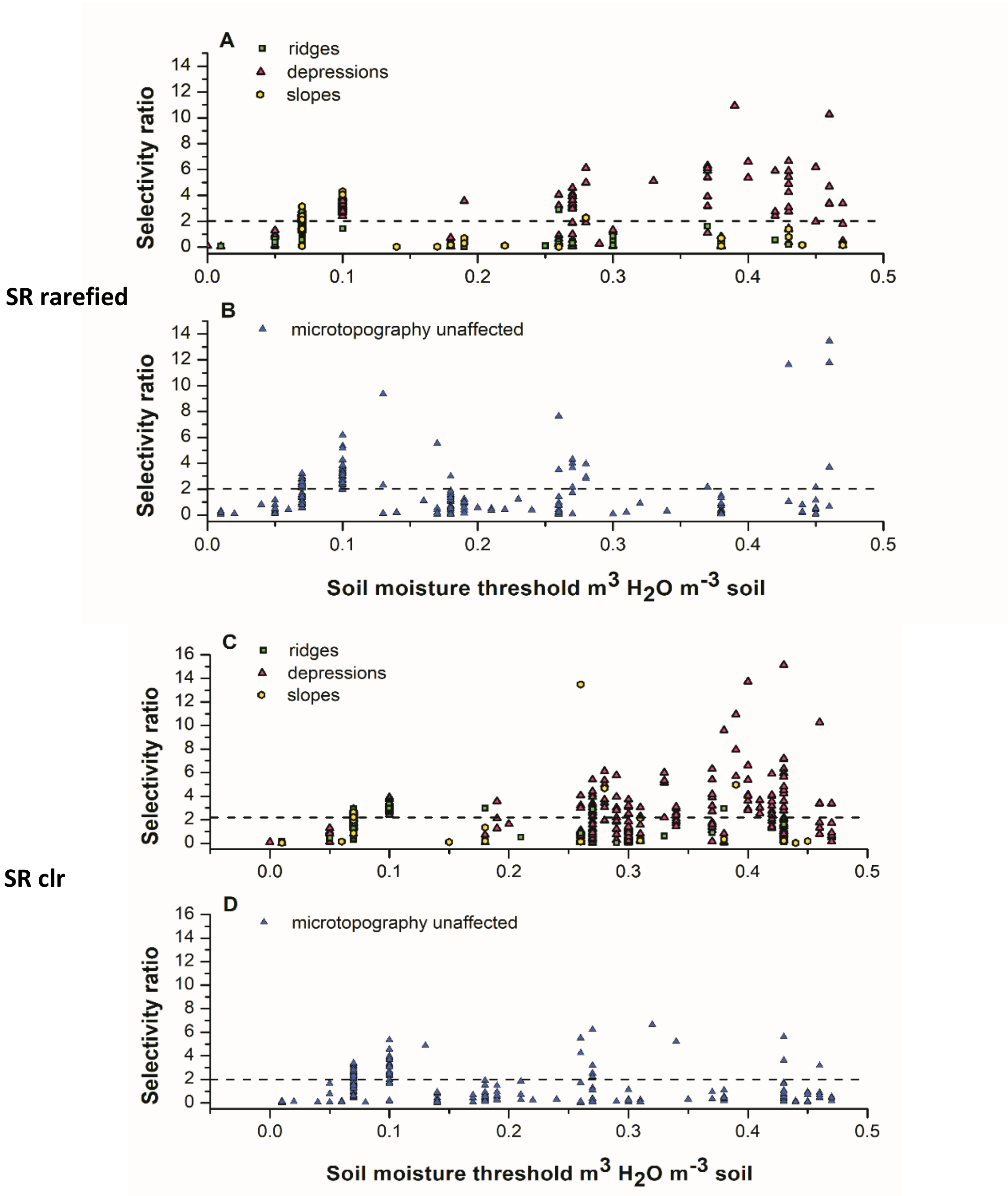

Figure S4: Plots displaying the maximum selectivity ratio (SR) in dependence on the 5-year soil moisture threshold values of each taxonomic group (TG) evaluated by using a rarefied TG dataset (A + B) or a centered log-ratio transformed dataset (C + D). Different colors indicate TGs that are affected by different micro-topographic expositions (A + C) or unaffected by micro-topography (C + D) according to STAMP. The dotted line indicates the SR cut-off at the value of 2, below which variables were considered to have no explanatory effect. Plots A and B are identical to those in figure 7 and are included here to allow direct comparison. The comparison of SR values derived from a rarefied dataset are largely in agreement with those derived from a clr-transformed dataset.