

Supplemental material

**Core and rare microbiota enhance bacterial community complexity
and stability along environmental gradients on the Loess Plateau**

Zhengchen Wang^{a,b}, Zijie Zhou^a, Ruoxi Hu^a, Siyu Lu^a, Qi Zhang^{a,b}, Fang Chen^{a,b},
Xiaojiao Wang^{a,b}, Yongzhong Feng^{a,b}, Chengjie Ren^{a,b}, Gaihe Yang^{a,b}, Fengxia Li^c,
Xing Wang^{a,b*}, Xinhui Han^{a,b*}

^a College of Agronomy, Northwest A&F University, Yangling 712100, Shaanxi, PR
China.

^b Shaanxi Engineering Research Center of Circular Agriculture, Yangling 712100,
Shaanxi, PR China.

^c Institute of Agricultural Resources and Environment, Academy of Agriculture and
Forestry Sciences, Ningxia, 750002, PR China.

*Corresponding author:

Xing Wang, wangxing1996@nwafu.edu.cn

Xinhui Han, hanxinhui@nwsuaf.edu.cn

Keyword:

Subcommunities; Community assembly; Biogeographic pattern; Microbial network;
Taxa; Loess plateau

23 **Table S1** Relationships between core taxa and diversity, stability, and complexity
24 across different ecosystems.

		Core MST			
		slope	F	R ²	P
AG	diversity	-0.05345	2.581145	0.012257	0.10966
	Δstability	-0.0534	24.20705	0.104248	< 0.001
	Δcomplexity	-0.00092	0.003399	1.63E-05	0.953562
AS	diversity	0.064897	1.789234	0.035229	0.187195
	Δstability	0.026582	10.63425	0.178325	0.002024
	Δcomplexity	-0.01571	0.158794	0.00323	0.692
AF	diversity	-0.13507	24.70984	0.106183	< 0.001
	Δstability	0.001369	0.082108	0.000395	0.774745
	Δcomplexity	-0.06093	4.374214	0.020597	0.037701
NG	diversity	-0.25379	31.31901	0.233169	< 0.001
	Δstability	-0.03227	12.11469	0.10524	< 0.001
	Δcomplexity	-0.01495	0.525128	0.005072	0.470305
NF	diversity	-0.04326	4.731328	0.007929	0.030013
	Δstability	0.011539	4.780345	0.00801	0.029177
	Δcomplexity	0.054417	6.083481	0.010172	0.013927

25 AG: artificial grassland, AS: artificial shrub, AF: artificial forest, NG: natural grassland, NF: natural
26 forest, *P* = p value. Significant results ($P \leq 0.05$) are shown in bold.

27 **Table S2** Relationships between Abundant taxa and diversity, stability, and complexity
 28 across different ecosystems.

		Abundant MST			
		slope	F	R ²	<i>P</i>
AG	diversity	-0.08094	3.766397	0.017786	0.053644
	Δstability	-0.05927	18.39775	0.081263	< 0.001
	Δcomplexity	-0.02591	1.738543	0.008289	0.188773
AS	diversity	-0.10432	3.170348	0.060769	0.081187
	Δstability	-0.00941	0.742849	0.014934	0.392948
	Δcomplexity	-0.01613	0.111661	0.002274	0.739687
AF	diversity	-0.27749	285.5216	0.578539	< 0.001
	Δstability	-0.01716	18.09545	0.080035	< 0.001
	Δcomplexity	-0.01139	0.193377	0.000929	0.660577
NG	diversity	-0.3777	162.4972	0.612049	< 0.001
	Δstability	-0.05004	44.14673	0.300018	< 0.001
	Δcomplexity	-0.06361	12.57959	0.108839	< 0.001
NF	diversity	-0.07612	39.38241	0.062375	< 0.001
	Δstability	-0.03404	127.4242	0.17712	< 0.001
	Δcomplexity	0.004207	0.091442	0.000154	0.762459

29 AG: artificial grassland, AS: artificial shrub, AF: artificial forest, NG: natural grassland, NF: natural
 30 forest, *P* = p value. Significant results ($P \leq 0.05$) are shown in bold.

31 **Table S3** Relationships between Rare taxa and diversity, stability, and complexity
32 across different ecosystems.

		Rare MST			
		slope	F	R ²	P
AG	diversity	-0.04119	2.592017	0.012308	0.10892
	Δstability	-0.01371	2.446143	0.011624	0.119334
	Δcomplexity	-0.0024	0.039634	0.000191	0.842392
AS	diversity	0.161875	19.28921	0.282464	< 0.001
	Δstability	-0.00329	0.173578	0.00353	0.67877
	Δcomplexity	0.036785	1.14408	0.022816	0.290033
AF	diversity	0.03121	3.132215	0.014835	0.078225
	Δstability	0.004837	2.716741	0.012893	0.10081
	Δcomplexity	-0.00779	0.183496	0.000881	0.668829
NG	diversity	-0.01847	0.197966	0.001918	0.657302
	Δstability	0.006131	0.611408	0.005901	0.436052
	Δcomplexity	-2.4E-05	2.12E-06	2.06E-08	0.998841
NF	diversity	0.069799	33.93716	0.054218	< 0.001
	Δstability	-0.01968	38.61337	0.061231	< 0.001
	Δcomplexity	0.047109	12.09541	0.020022	< 0.001

33 AG: artificial grassland, AS: artificial shrub, AF: artificial forest, NG: natural grassland, NF: natural
34 forest, P = p value. Significant results ($P \leq 0.05$) are shown in bold.

35

Table S4 Results of the principal component analysis (PCA) of climate factors, vegetation properties, and soil properties. Loadings and correlations for each factor with the first principal component are shown. MAP and MAT represent mean annual precipitation and mean annual temperature, respectively. SOC, TN, TP, C:N, MBC, MBN, and MBP represent soil organic carbon, total nitrogen, total phosphorus, carbon-to-nitrogen ratio, microbial biomass carbon, microbial biomass nitrogen, and microbial biomass phosphorus, respectively. Vegetation properties include Gleason's richness, Shannon-Wiener diversity index, and Pielou's evenness. Soil texture, pH, bulk density (BD), soil water content (SWC) and inorganic nitrogen forms (NH_4^+ and NO_3^-).

	Loading	Correlations			Loading	Correlations	
Climate factors				Vegetation properties			
MAP	0.81	0.90	***	Gleason	0.72	0.85	***
MAT	0.81	0.90	***	Shannon.Wie	0.99	0.99	***
				ner			
				Pielou	0.50	0.70	***
Soil properties							
SOC	0.87	0.93	***	Texture	0.40	-0.63	***
TN	0.81	0.90	***	pH	0.90	-0.95	***
TP	0.10	0.32	***	BD	0.42	-0.65	***
C/N	0.50	0.71	***	SWC	0.78	0.88	***
MBC	0.83	0.91	***	NH ₄ ⁺	0.60	0.78	***
MBN	0.80	0.90	***	NO ₃ ⁻	0.37	0.60	***
MBP	0.77	0.88	***				

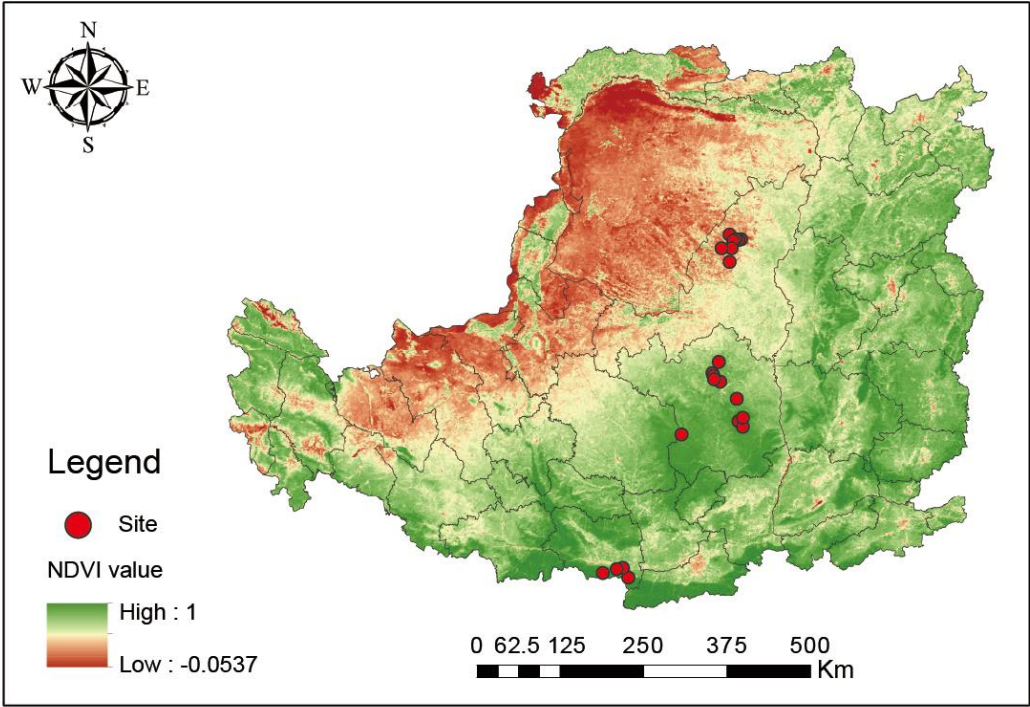
Asterisks denote levels of significance (*** $p < 0.001$).

Table S5 Mantel, Partial mantel, and Moran's I test for evaluating spatial effects in bacterial subcommunities and network properties.

	Mantel test		Partial mantel test	
	Geo	Env	Geo Env	Env Geo
Whole	0.806 ***	0.735 ***	0.595 ***	0.388 ***
Core	0.734 ***	0.644 ***	0.513 ***	0.255 ***
Abundant	0.784 ***	0.705 ***	0.567 ***	0.338 ***
Rare	0.623 ***	0.576 ***	0.370 ***	0.242 ***
Moran's I				
	Basic I	Basic P	Sensitivity I	Sensitivity P
Core composition	0.269	0.004	-0.075	0.548
Abundant composition	0.273	0.004	-0.139	0.185
Rare composition	0.419	< 0.001	-0.16	0.122
Core richness	0.138	0.077	-0.22	0.022
Abundant richness	0.357	0.001	-0.214	0.034
Rare richness	0.022	0.643	-0.177	0.084
Complexity	0.162	0.047	-0.096	0.422
Stability	-0.05	0.77	-0.153	0.142

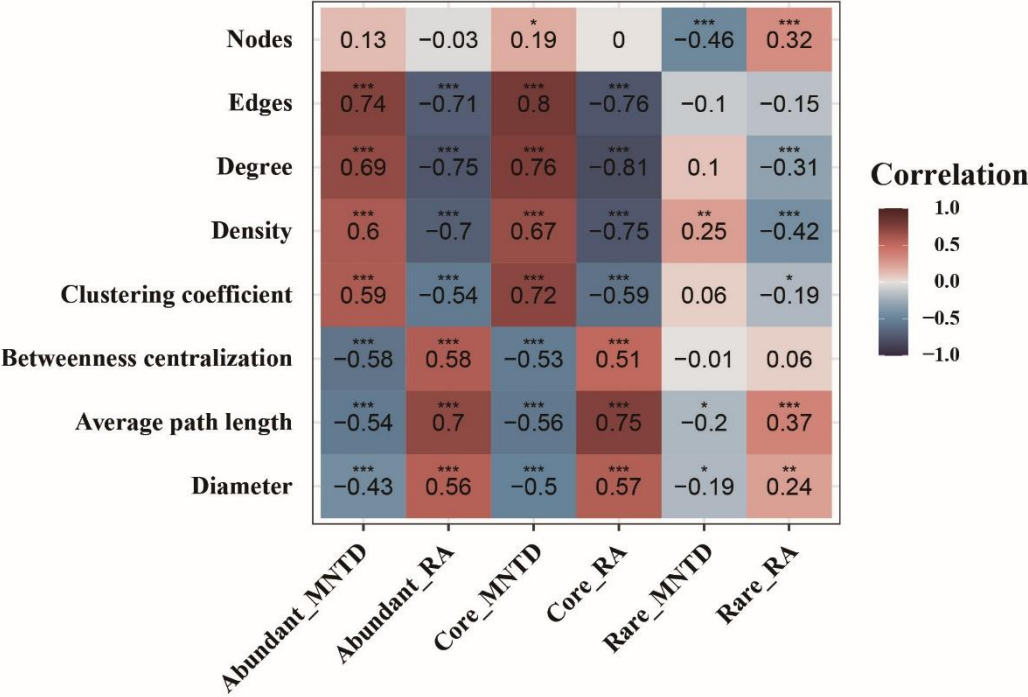
Mantel and partial Mantel tests were based on Bray-Curtis dissimilarities. Geo, geographic distance; Env, environmental distance. Geo | Env indicates the partial Mantel correlation between community dissimilarity and geographic distance after controlling for environmental distance, whereas Env | Geo indicates the partial Mantel correlation between community dissimilarity and environmental distance after controlling for geographic distance. *** $P < 0.001$. Moran's I was used to evaluate residual spatial autocorrelation after accounting for environmental gradients and vegetation background. "Basic" models included Climate_PC1, Soil_PC1, and Vegetation_PC1; "Sensitivity" models additionally included ecosystem type and vegetation type. Composition refers to NMDS1. Bold P values indicate $P < 0.05$.

58 **Fig. S1** Geographic distribution of sampling sites across the Loess Plateau. The NDVI
59 background was derived from the MODIS/Terra MOD13Q1 Version 6.1 product. The
60 Loess Plateau boundary and provincial boundaries were obtained from the National
61 Earth System Science Data Center of China, provided by the National Earth System
62 Science Data Center of China. The map was created using ArcGIS (Esri, Redlands,
63 CA, USA).

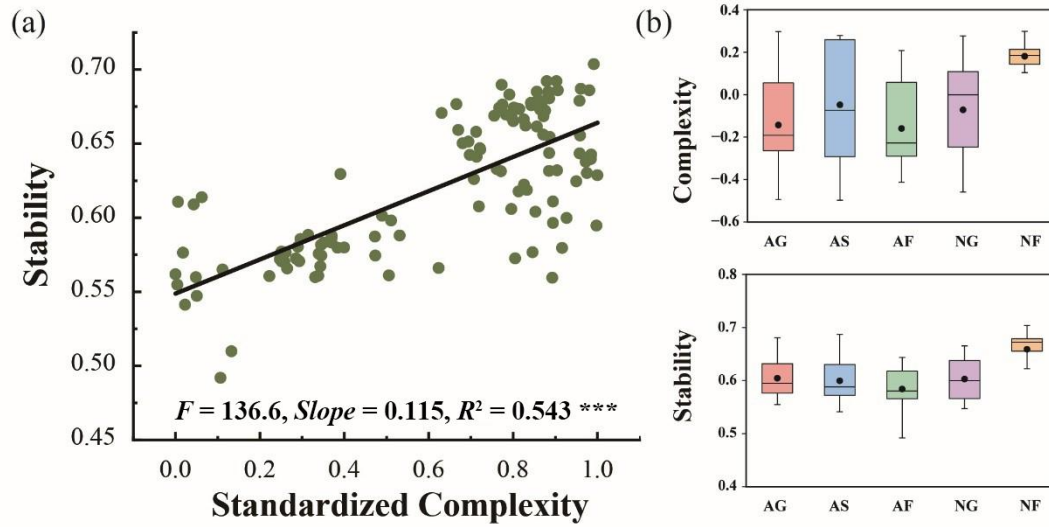


64
65

Fig. S2 Heatmap of correlations between bacterial taxa of different abundances and community network topological properties. The heatmap illustrates the correlations of core, abundant, and rare bacterial taxa under mean nearest taxon distance (MNTD) and relative abundance (RA) conditions with network properties, including node count, edge count, node degree, density, clustering coefficient, betweenness centrality, average path length, and network diameter. Positive correlations are represented in red, and negative correlations in blue, with the intensity of the color indicating the strength of the correlation. Asterisks indicate significance levels (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$).



76 **Fig. S3 (a)** Relationship between community complexity and stability. The scatter plot
 77 illustrates the positive correlation between community complexity and stability. The x-
 78 axis represents the Min-Max scaled standardized community complexity (Standardized
 79 Complexity), and the y-axis represents stability (Stability). Each dot represents a
 80 sample, and the black line indicates the fitted linear regression. Asterisks denote the
 81 significance level ($***P < 0.001$). **(b)** Boxplots of complexity and stability for different
 82 land use types: AG (artificial grassland), AS (artificial shrubland), AF (artificial forest),
 83 NG (natural grassland), and NF (natural forest). The central line in each box represents
 84 the median, the black dot indicates the mean.



86 **Fig. S4** Changes in community complexity and stability along environmental and
 87 geographical distances. Phylogenetic distance-decay curves showing community
 88 complexity and stability against environmental **(a)** and geographical **(b)** distance.
 89 Community complexity is represented in dark green, and stability in light green. The
 90 correlation coefficients of the linear regression models are provided for each panel, with
 91 asterisks indicating significance levels (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$).

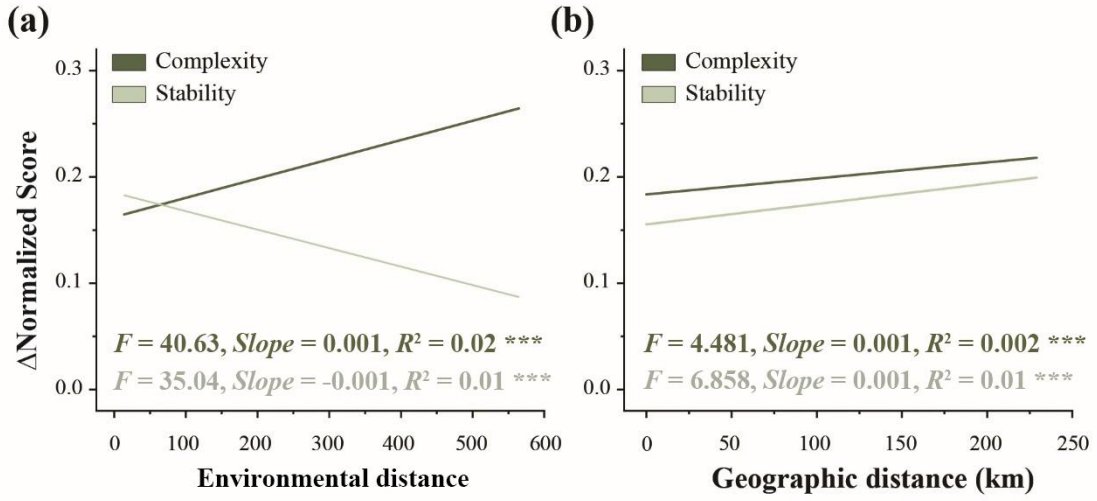


Fig. S5 Principal component analysis (PCA) of environmental factors. PCA of environmental factors using Type II scaling. Arrows indicate the direction and magnitude of the projections of environmental factors, with arrow length reflecting the contribution of each factor to the principal components. Environmental factors include: Gleason index (Gleason), Shannon-Wiener index (Shannon-Wiener), Pielou index (Pielou); soil texture (Texture), pH (Ph), bulk density (BD), soil water content (SWC); ammonium nitrogen (NH_4), nitrate nitrogen (NO_3); soil organic carbon (SOC), total nitrogen (TN), total phosphorus (TP), carbon-to-nitrogen ratio (C:N); microbial biomass carbon (MBC), microbial biomass nitrogen (MBN), microbial biomass phosphorus (MBP); mean annual precipitation (MAP), and mean annual temperature (MAT).

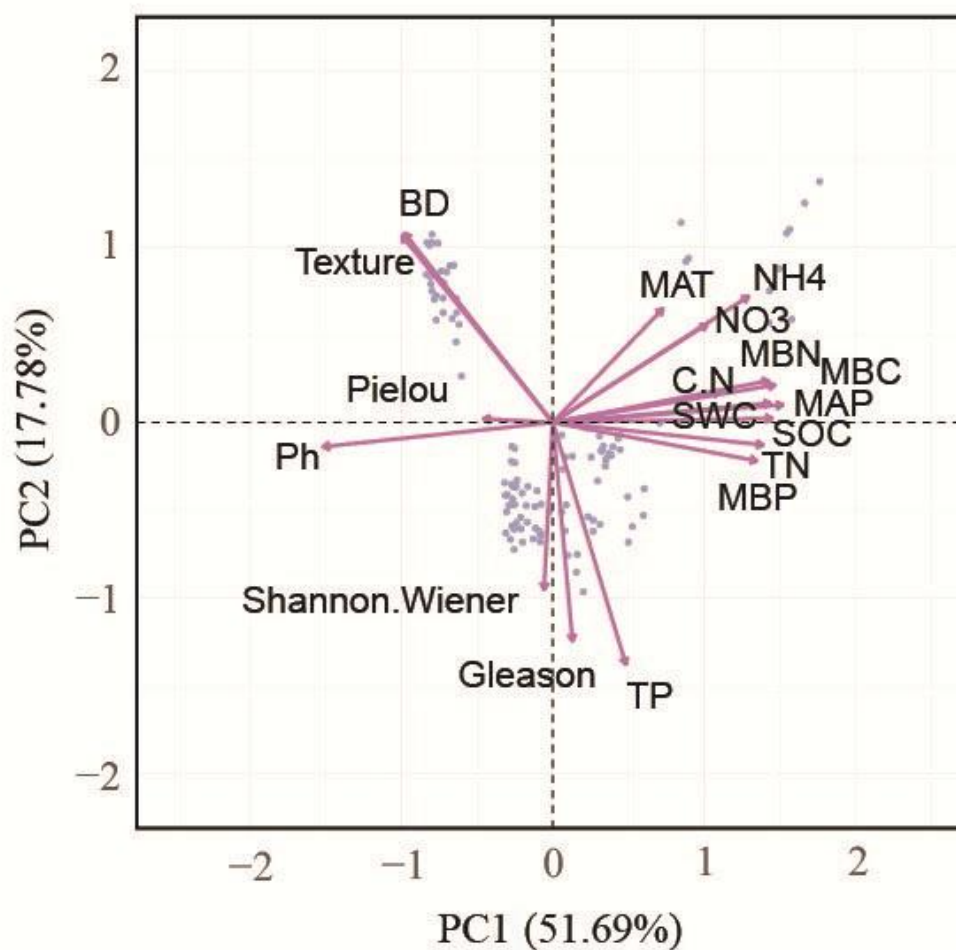


Fig. S6 The prior model illustrates the role of bacterial taxa with different abundances in shaping community complexity and stability along environmental gradients. This model integrates the effects of climate factors, soil characteristics, and vegetation traits on bacterial taxa (core, abundant, and rare taxa) and their subsequent impacts on microbial network complexity and stability. Arrows represent potential causal relationships between variables.

