Supplementary Material

This file includes:

Supplementary text

Supplementary references

Supplementary tables S1 – S4

Supplementary figures S1 – S6

Supplementary Text

Compilation of data and user-defined source groups

In both case studies presented in this article, the source groups are user-defined, based on information from published plant *n*-alkane records. In case study 1, the three source groups (terrestrial plants, submerged aquatic macrophytes, and algae) were assigned to the published data records in Supplementary Material EA-2 based on either taxonomic identification or sample description from the original publications. Details on the taxonomic identification, sample description, lipid extraction methods, additional environmental information, and referenced publications are listed in Supplementary Material EA-2. In case study 2, the three source groups (rainforest C_3 plants, savanna C_3 plants, C_4 plants) were assigned to the published data records in in Supplementary Material EA-3 and EA-4 based on either taxonomic identification or description of sample location from the original publications. Details on the taxonomic identification, sample description, lipid extraction methods, and referenced publications are listed in Supplementary Material EA-3 and EA-4.

The potential influence of different lipid extraction methods

We listed the details regarding the lipid extraction methods used in each publication in the Supplementary Material, due to the potential influence of the different methods as well as the organic solvents on leaf wax lipid extraction yield and composition. In this study, the influence of extraction methods on the lipid yield and composition is assumed to be minimal. While extraction yield has been found to be different with different methods and solvents used (Ardenghi et al., 2017; Weber and Schwark, 2020), it most likely affects the chain-specific *n*-alkane concentrations of the samples, which are known to be associated with large uncertainties due to their log-normal distributions. The potential influence of extraction methods on sample specific *n*-alkane concentrations should have little effects on the prior distributions when a large number of samples are included in the calculation of prior parameters (Supplementary Material EA-2 to EA-4). Moreover, the results of sensitivity tests in section 3.3 illustrate that the model is more sensitive to the $\delta^{13}C$ values of the prior distribution and the likelihood evaluation than relative abundance. Since $\delta^{13}C$ values are independent to the *n*-alkane concentrations, extraction

methods should have little influence on the central tendency of the posterior distribution of the mixing ratios or chain-specific mixing ratios reported in sections 3.1 and 3.2. More systematic investigations will help to evaluate the influence of extraction methods on chain-specific *n*-alkane concentrations, and subsequently, to constrain model uncertainty.

Comparing prior and posterior distributions

Because the model explores all possible combinations of parameter values in the multivariate parameter space, the posterior distributions of some parameters may differ from their priors, depending on the data values. Such deviations can be informative on how constraint the parameters are, as well as any potential biases in the prior parameter estimations. Here, we report the comparisons between prior and posterior distributions using examples from both case studies.

The first comparison is based on the model output of the QHS-5S sample. For per sample *n*-alkane concentrations, the posterior distributions are quite similar to the prior distributions, with some small deviations observed in the terrestrial source in the *n*-C_{29} and the *n*-C_{31} chains (Figure S3). For per sample *n*-alkane $\delta^{13}C$, the posterior distributions are identical to the prior distributions (Figure S4).

The second comparison is based on the model output of the Asso sample. For per sample *n*-alkane concentrations, the posterior distributions are quite similar to the prior distributions, with some small deviations observed in the C_4 plants in all three chains, in the savanna C_3 plants in the *n*-C_{33} chain, as well as in the rainforest C_3 plants in the *n*-C_{31} and the *n*-C_{33} chains (Figure S5). For per sample *n*-alkane $\delta^{13}C$, the posterior distributions are identical to the prior distributions (Figure S6).

Because the proxy system model used here is simple, the posterior distributions of the two case studies are generally quite similar to the priors (Figures S3 to S6). When deviations from the prior distribution do occur (Figures S3 and S5), they reflect the fact that the *n*-alkane chain length distribution of the sedimentary samples differ substantially from the central tendency of any source. For instance, the QHS-5S sample displays a *n*-C_{31} dominance (Figure 5), which is only possible if the terrestrial source has a greater *n*-C_{31} dominance than reflected in the prior (Figures S3). Similarly, the strong *n*-C_{29} dominance in the Asso sample (Figure 8) is best explained by somewhat higher than expected *n*-C_{29} abundance in the rainforest C_3 source (Figures S5). Such information can be used to refine our prior assumptions associated with the model, and to provide information on potential biases or alternative interpretations.

Supplementary References

Ardenghi, N., Mulch, A., Pross, J., and Maria Niedermeyer, E.: Leaf wax n-alkane extraction: An optimised procedure, Org Geochem, 113, 283-292, https://doi.org/10.1016/j.orggeochem.2017.08.012, 2017.

Weber, J. and Schwark, L.: Epicuticular wax lipid composition of endemic European Betula species in a simulated ontogenetic/diagenetic continuum and its application to chemotaxonomy and paleobotany, Sci Total Environ, 730, 138324, https://doi.org/10.1016/j.scitotenv.2020.138324, 2020.

Supplementary Tables

Table S1: Means and variance-covariance matrices of *n*-alkane $\delta^{13}C$ of terrestrial, aquatic macrophyte and algae n-alkane sources around Lake Qinghai.

$\delta^{13}C$ Means (‰, VPDB)	*n*-C_{27}	*n*-C_{29}	*n*-C_{31}
Terrestrial	-32.8	-33.5	-33.5
Aquatic macrophyte	-22.6	-23.3	-23.8
Algae	-31.2	-31.0	-32.1
Terrestrial $\delta^{13}C$	*n*-C_{27}	*n*-C_{29}	*n*-C_{31}
V-covariance matrix			
n-C_{27}	1.9684	-0.3273	0.8149
n-C_{29}	-	1.4068	-0.0304
n-C_{31}	-	-	2.0529
Aquatic macrophyte $\delta^{13}C$	*n*-C_{27}	*n*-C_{29}	*n*-C_{31}
V-covariance matrix			
n-C_{27}	18.1873	16.5000	16.1755
n-C_{29}	-	15.6888	15.4095
n-C_{31}	-	-	18.9184
Algae $\delta^{13}C$	*n*-C_{27}	*n*-C_{29}	*n*-C_{31}
V-covariance matrix			
n-C_{27}	6.3562	-4.2521	-1.9812
n-C_{29}	-	9.9662	4.9155
n-C_{31}	-	-	2.8562

Table S2: Means and variance-covariance matrices of ln-transformed n-alkane concentration of terrestrial, aquatic macrophyte and algae n-alkane sources around Lake Qinghai.

ln(concentration) ln(μg/g)	*n*-C_{27}	*n*-C_{29}	*n*-C_{31}
Terrestrial	3.6781	4.3077	3.5228
Aquatic macrophyte	2.7477	2.2110	-0.0476
Algae	-0.8770	-1.0454	-1.6011
Terrestrial ln(concentration) V-covariance matrix	*n*-C_{27}	*n*-C_{29}	*n*-C_{31}
n-C_{27}	1.1286	0.9208	0.5662
n-C_{29}	-	1.8794	1.3238
n-C_{31}	-	-	2.7186
Aquatic macrophyte ln(concentration) V-covariance matrix	*n*-C_{27}	*n*-C_{29}	*n*-C_{31}
n-C_{27}	1.1519	0.9627	0.4091
n-C_{29}	-	0.9453	0.5098
n-C_{31}	-	-	1.1875
Algae ln(concentration) V-covariance matrix	*n*-C_{27}	*n*-C_{29}	*n*-C_{31}
n-C_{27}	0.7614	0.6063	0.4345
n-C_{29}		0.7170	0.5284
n-C_{31}			0.6045

Table S3: Means and variance-covariance matrices of n-alkane $\delta^{13}C$ of C_4 grasses, savanna C_3 plants and rainforest C_3 plants of sub-Saharan Africa.

$\delta^{13}C$ Means (‰, VPDB)	*n*-C_{29}	*n*-C_{31}	*n*-C_{33}
C_4 grasses	-21.7	-21.6	-22.0
Savanna C_3	-34.0	-34.5	-34.9
Rainforest C_3	-38.1	-38.4	-38.9
C_4 grasses $\delta^{13}C$	*n*-C_{29}	*n*-C_{31}	*n*-C_{33}
V-covariance matrix			
n-C_{29}	3.2710	2.4707	2.2354
n-C_{31}	-	3.1060	2.6191
n-C_{33}	-	-	3.5775
Savanna C_3 $\delta^{13}C$	*n*-C_{29}	*n*-C_{31}	*n*-C_{33}
V-covariance matrix			
n-C_{29}	6.6406	5.5126	3.8458
n-C_{31}	-	6.7642	5.4598
n-C_{33}	-	-	6.8613
Rainforest C_3 $\delta^{13}C$	*n*-C_{29}	*n*-C_{31}	*n*-C_{33}
V-covariance matrix			
n-C_{29}	6.8823	7.4840	7.4830
n-C_{31}	-	9.0163	9.4427
n-C_{33}	-	-	11.1913

Table S4: Means and variance-covariance matrices of ln-transformed n-alkane concentration of C_4 grasses, savanna C_3 plants and rainforest C_3 plants of sub-Saharan Africa.

ln(concentration) ln(µg/g)	*n*-C_{29}	*n*-C_{31}	*n*-C_{33}
C_4 grasses	3.4433	4.4287	3.7499
Savanna C_3	3.4862	3.6974	2.4946
Rainforest C_3	3.9780	3.8278	2.0584
C_4 grasses ln(concentration)	*n*-C_{29}	*n*-C_{31}	*n*-C_{33}
V-covariance matrix			
n-C_{29}	1.3710	0.6492	0.4509
n-C_{31}	-	1.5669	0.6833
n-C_{33}	-	-	1.6065
Savanna C_3 ln(concentration)	*n*-C_{29}	*n*-C_{31}	*n*-C_{33}
V-covariance matrix			
n-C_{29}	1.9701	1.3619	1.3795
n-C_{31}	-	2.5040	2.3458
n-C_{33}	-	-	4.0725
Rainforest C_3 ln(concentration)	*n*-C_{29}	*n*-C_{31}	*n*-C_{33}
V-covariance matrix			
n-C_{29}	2.7555	1.1830	0.6894
n-C_{31}	-	1.9080	1.4103
n-C_{33}	-	-	2.2133

Supplementary Figures

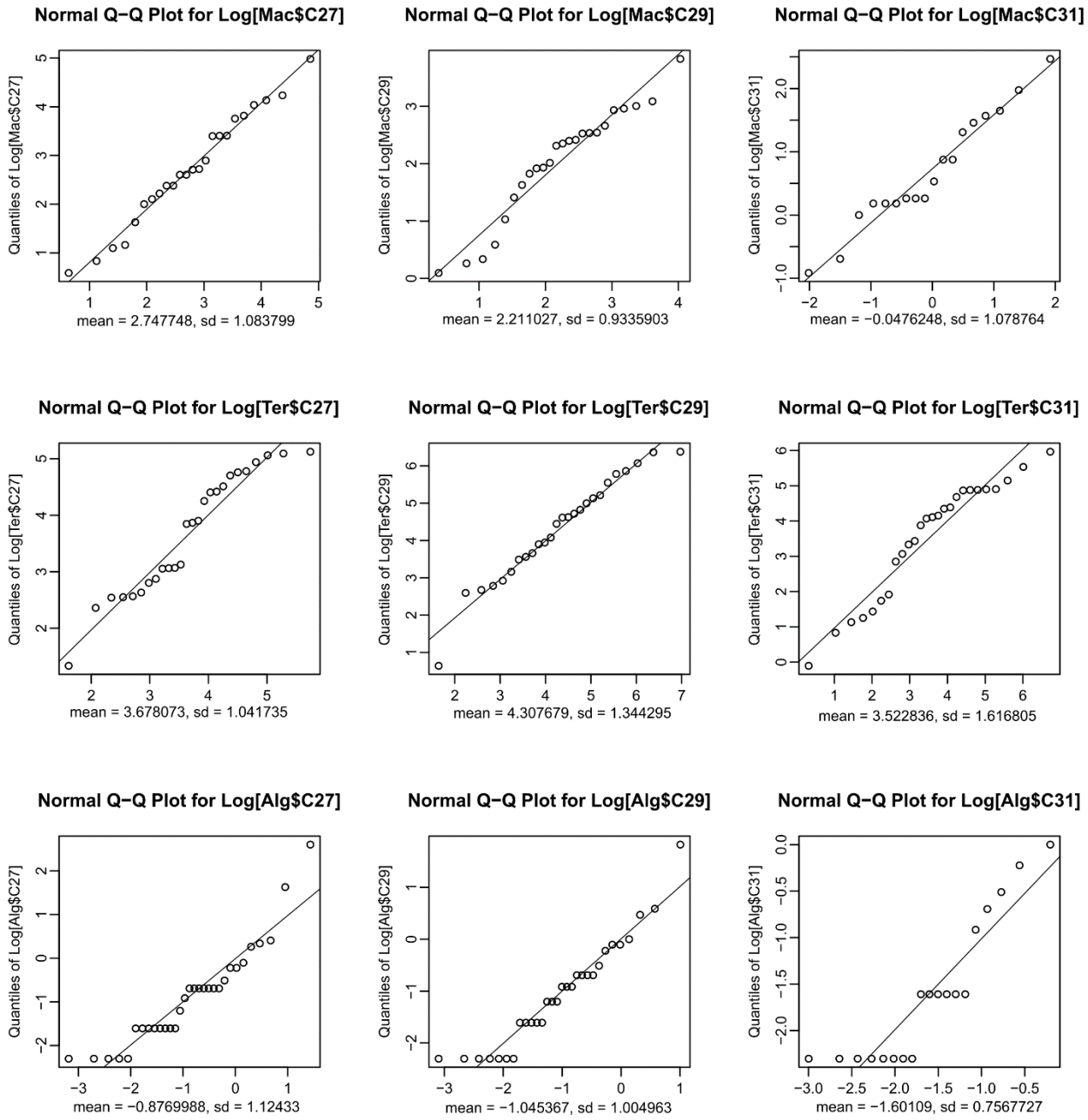


Figure S1: Quantile-quantile plots between parameterized log-normal distribution and empirical data used in case study 1, demonstrating goodness of fit between the data and the model; columns are the three *n*-alkane chains used in the case study, rows are the end members used in the case study; diagonal lines indicate the 1:1 relationship.

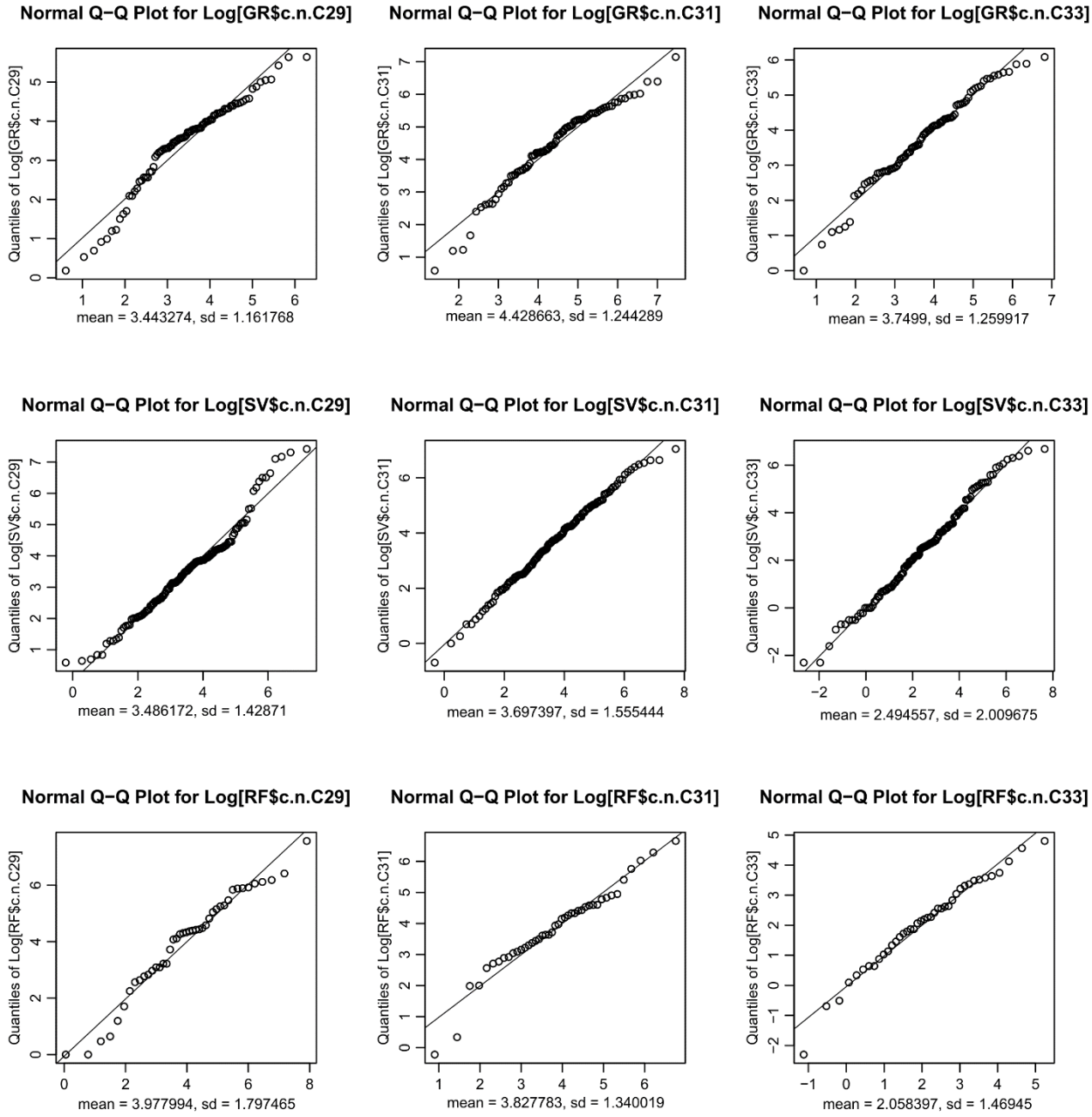


Figure S2: Quantile-quantile plots between parameterized log-normal distribution and empirical data used in case study 2, demonstrating goodness of fit between the data and the model; columns are the three *n*-alkane chains used in the case study, rows are the end members used in the case study; diagonal lines indicate the 1:1 relationship.

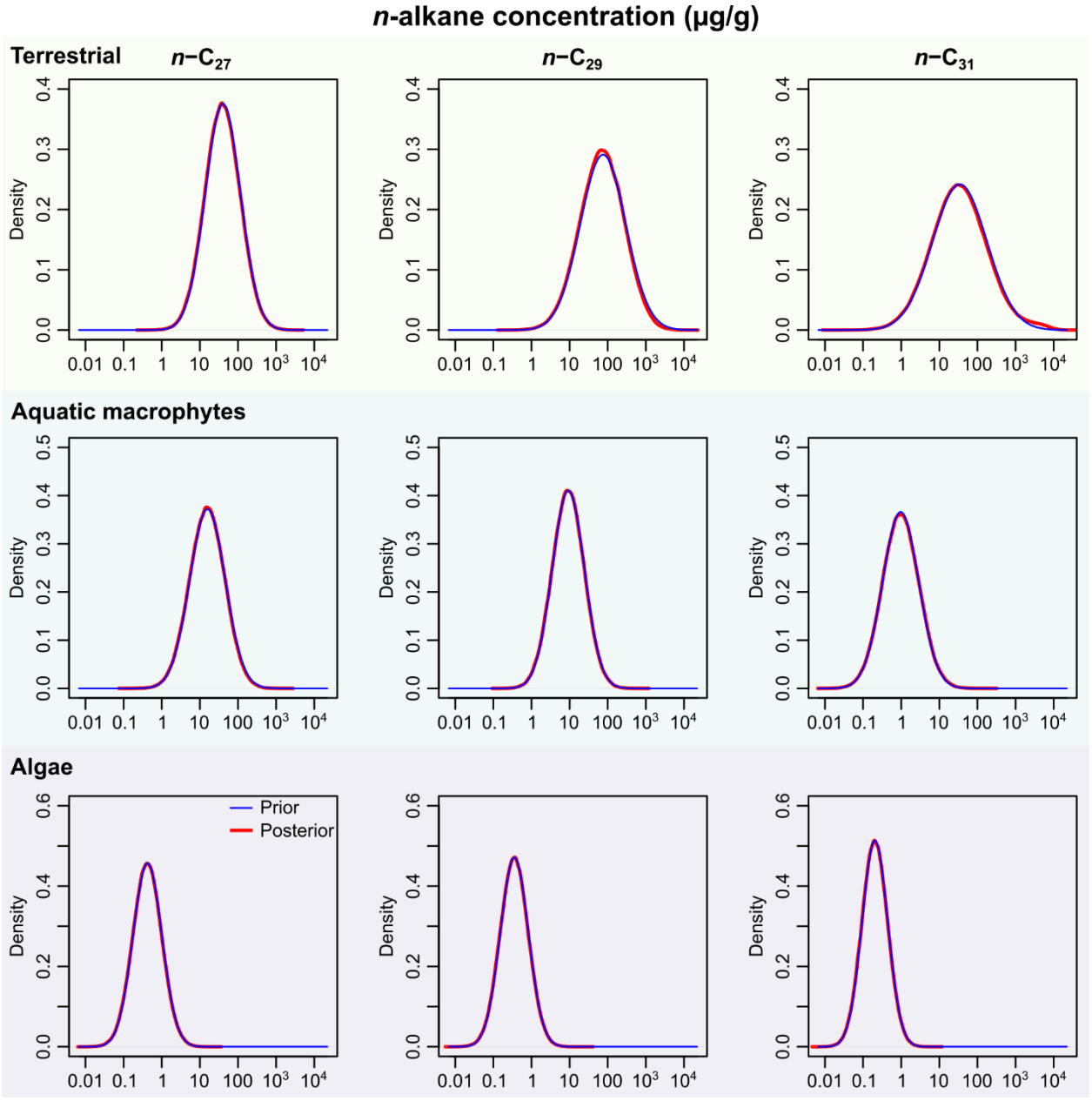


Figure S3. Comparison between the *n*-alkane concentration prior distribution of case study 1 and the posterior distribution based on the QHS-5S sample.

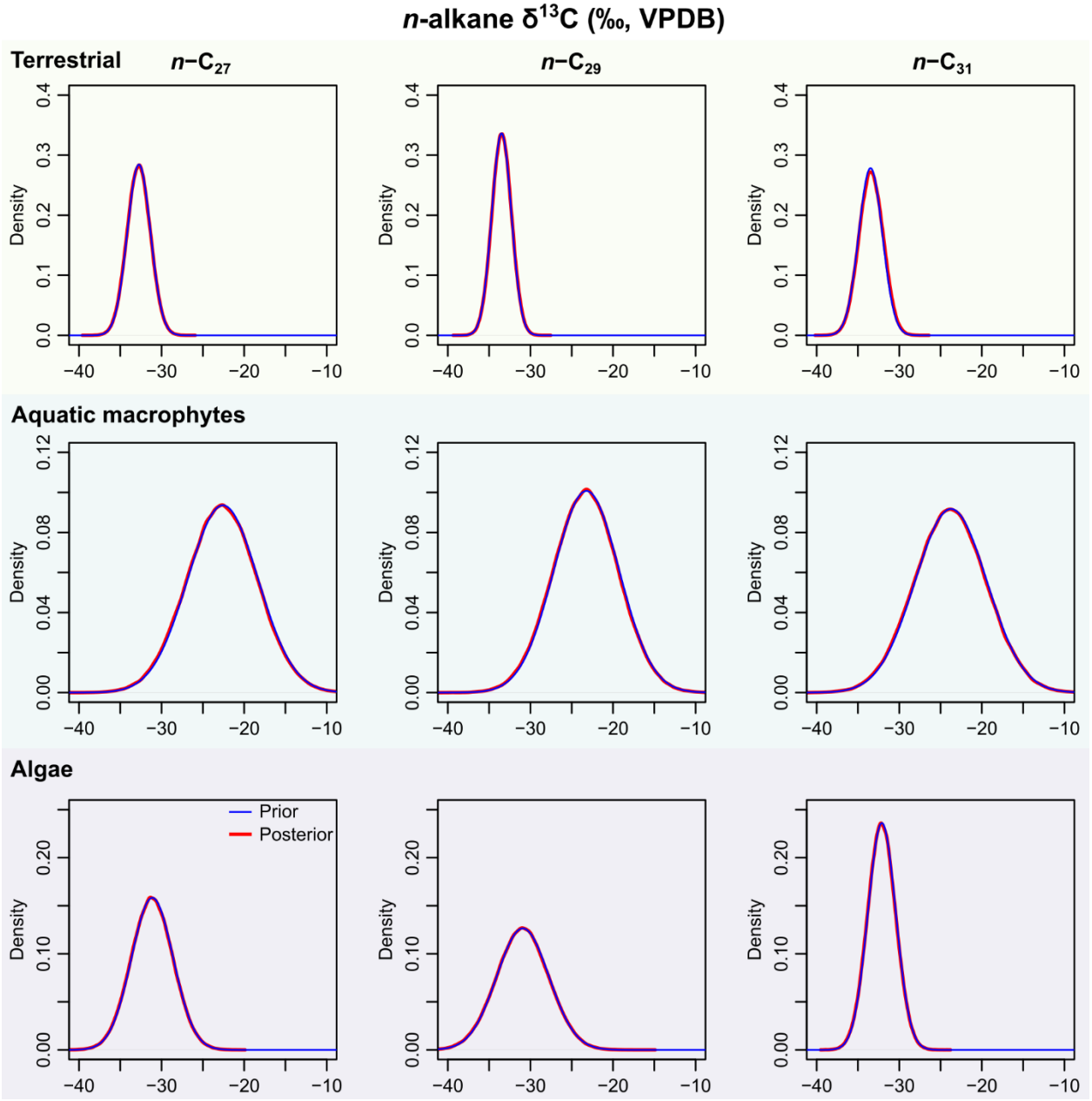


Figure S4. Comparison between the *n*-alkane $\delta^{13}C$ prior distribution of case study 1 and the posterior distribution based on the QHS-5S sample.

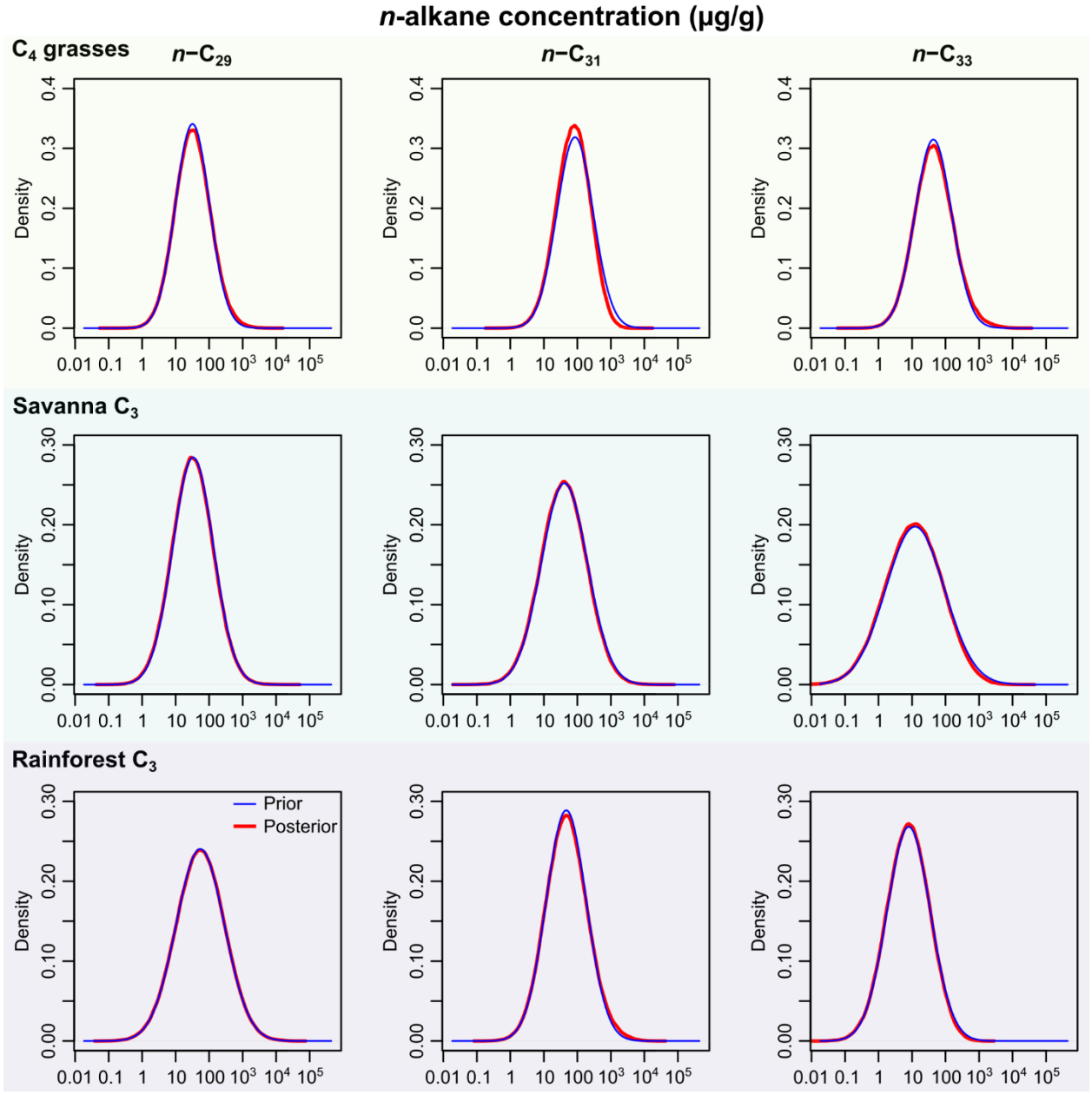
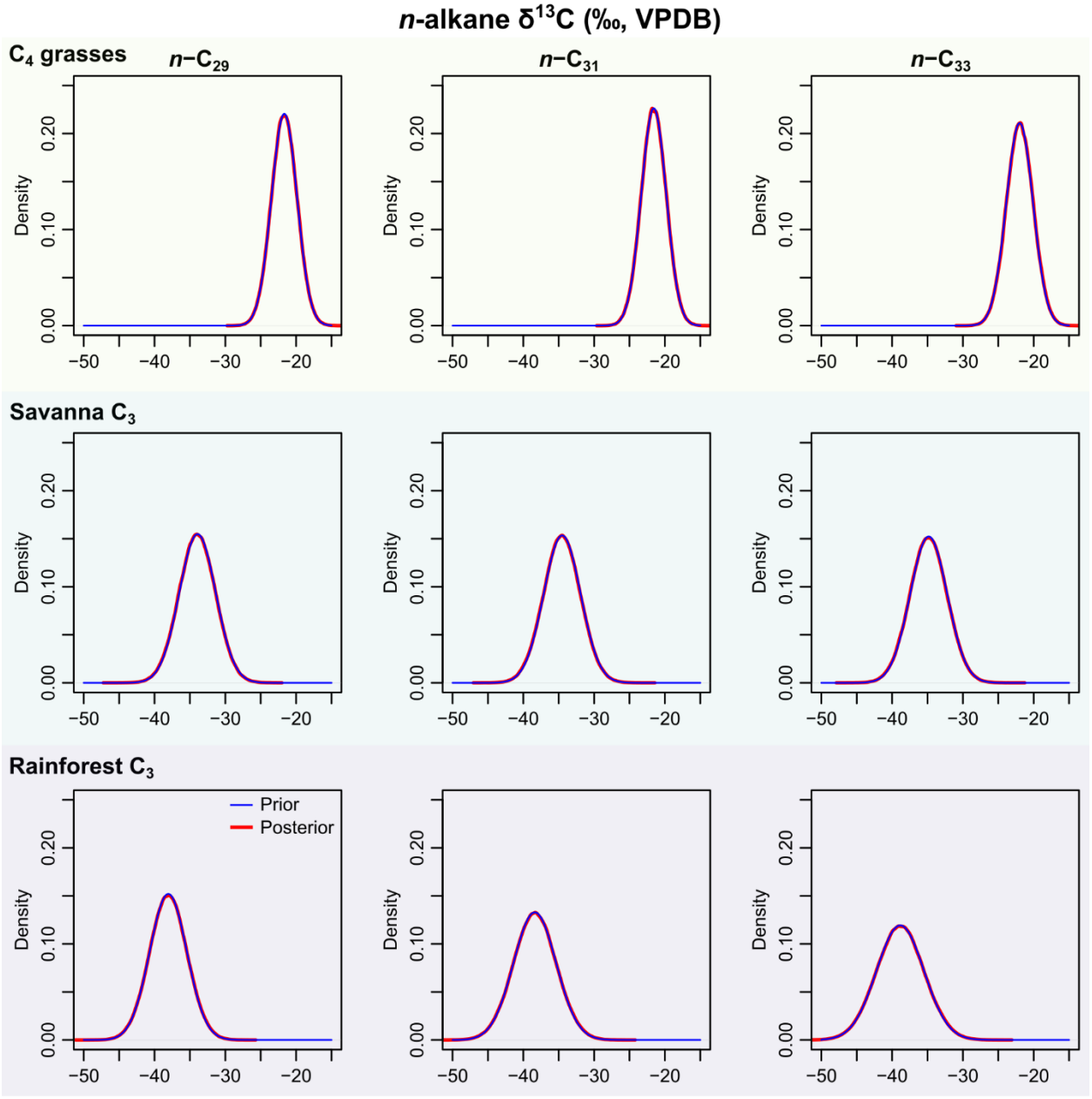


Figure S5. Comparison between the *n*-alkane concentration prior distribution of case study 2 and the posterior distribution based on the Asso sample.



95 **Figure S6. Comparison between the *n*-alkane $\delta^{13}C$ prior distribution of case study 2 and the posterior distribution based on the Asso sample.**