Lignin in the Organic Matter of the Soils of the Russian Plain as Biomarker of Palaeoenvironment

I V Kovalev, N O Kovaleva
Soil Science Faculty, Lomonosov Moscow State University, Russia, Moscow
E-mail: kovalevmsu@mail.ru

Abstract. It has been shown by the methods of biochemistry, nuclear magnetic resonance, and isotope geochemistry that the proportions of lignin phenols may be used as molecular traces of paleovegetation due to their biochemical and physiological specificity and high resistance to decomposition. Lignin structures have been detected in soils and in iron–manganese concretions. The comparison of the $^{13}$C NMR spectra of native lignin preparations isolated from different woody and herbaceous species with those of soil humic acids makes it possible to identify many characteristic shifts of lignin nature in humic acids at 56, 102, 115, 119, 131, 147, 151–152, 160, and 166 ppm. The information role of biomarker has been tested at the reconstruction of paleovegetation in the uplands of the Russian Plain. The representativeness of information has been increased using the isotope analysis ($\delta^{13}$C) and the radiocarbon dating; a new parameter—the composition of lignin phenols—has been introduced to the existing system of biomarkers.

1. Introduction
Biomarkers are organic molecules of the known structure and origin. Separate molecules of lignin phenols, amino acids, amino sugars, lipids can be reliably detected in living organisms, their residues, and waste, as well as in complex associates of humic substances in the soil [8]. Many biomarkers are resistant to degradation and mineralization in soils under specific conditions because of their specific structure and biochemical functions; therefore, they serve as molecular traces of paleobiota and land vegetation and as indicators of the rates of recent and past biochemical processes occurring in the biosphere. Individual organic compounds are well preserved not only in the normal profiles of postlithogenic soils, but also in their redeposition products, soilcolluvial and soilalluvial complexes, pedosediments, pedoliths, derivatives of separate horizons, bottom sediments of water bodies, etc [7].

The information role of biomarkers is important when the pool of molecular products of organic matter decomposition in soils is not identical to the sum of individual waste components because of the masking effect of the mineral matrix. The information role of biomarkers extracted from soils can be increased due to the determination of their structurally specific isotopes [8]. The diagnostics of the structural elements of individual nonspecific compounds in associates of humic substances, e.g., by NMR spectroscopy, in combination with the radiocarbon dating of soil humus increases the representativeness of the obtained information. Although the current database on the contents of similar individual compounds in soils is still poor, there is a significant body of data in the Russian soil science about the nonspecific groups of compounds like lipids, chlorophyll, amino acids, lignin, inositol phosphates, fungal melanins, etc. in soils [7]. A unique methodological basis for the interpretation of biochemical information from soil archives includes the studies of organic matter from buried soils and Fe-Mn nodules.
Lignin is the most prevalent natural phenolic compound of plant origin. The highest contents of lignin are in woody plants: lignin makes up 18–25% of wood biomass in deciduous species and 25–33% in coniferous species, while grasses contain about 4–9% lignin [2]. Chemically, lignin is an irregular threedimensional biopolymer of high molecular weight composed of branched phenylpropane units with colloidal properties [2]. The relative proportions of the component phenols are determined by the phylogenetic origin of plants and thus cause the appearance of many various low- and high-molecular-weight products of lignin decomposition in soils of different ecosystems. Different types of plants and plant communities produce a phylogenetically specific and strictly individual set of phenolic compounds.

The high stability of lignin phenols in buried soils is due to the fact that the phenylpropane moieties after burial enter into carboxylation and fragmentary condensation reactions, and their structural unites become less capable of entering into biochemical reactions, especially at the decreased biological activity [7]. Therefore, the aim of this work was to assess the information role of individual organic compounds of lignin origin and their applicability as molecular traces of paleovegetation, especially in soils of different ages and complex genesis.

2. Materials and methods

The objects of study included: light-gray forest soils (Greyic Phaeozems Albic; Eutric Podzoluvisols), light-gray forest gleic and gley soils (Greyic Phaeozems Albic; Gleyic Greyzems) and agrogray soils of the southern taiga (Moscov district, 38° 12’ E; 55° 02’ N); gray forest soils (Greyic Phaeozems Albic; Haplic Greyzems) under broad-leaved forest of Tul’skie Zaseki Reserve (Tula district, 37° 30’ E; 54° 00’N); dark gray forest soils on the microelevations and dark gray forest soils with the second humus horizon (Greyic Phaeozems Albic; Haplic Greyzems) on the microdepressions, as well as agrogray soils and gray forest gleic and gley soils (Greyic Phaeozems Albic; Gleyic Greyzems) at footslopes of the Bryansk district (33° 38’ E; 52° 34’ N).

Lignin phenols were isolated using Amelung’s version [1] of the Ertel–Hedges procedure [2]. The determination of lignin in the soils included the alkaline oxidation of the sample with copper oxide at 170°C under pressure in a nitrogen atmosphere, the precipitation of HAs, and the preconcentration of phenolic products on compact disposable C18 columns under pressure. Lignin preparations were isolated by the evaporation of ethyl acetate on a rotary evaporator. The phenol components of lignin were derivatized into trimethylsilyl ethers and separated on a gasliquid chromatograph. They were determined on a gas chromatograph–mass spectrometer (Hewlett-Packard, Palo Alto, CA, USA) with a flame ionization detector and a capillary column. Individual reaction products (vanillin, syringic aldehyde (al), syringic acid (ac), p-coumaric acid, and ferulic acid) were identified by comparing the retention times and peaks with those of the known components and amounts used as external standards. The reproducibility of results was 95%. It should be emphasized that the oxidation products of lignin include only methylated lignin structures without any changes in their ring fragments. The alkaline oxidation of vascular pant tissues and their remains by copper oxide in the soil yields 11 phenols [2], which can be grouped in accordance with their chemical nature into three structural families: vanillyl (guaiacyl) (V), syringyl (S), and cinnamyl (C) ones. The two first phenol types are found in mixed oxidation products of plant tissues in the form of aldehydes, ketones, and acids, while the last have only acid forms: ferulic (F) and p-coumaric (K) acids. Thus, the sum of oxidation products reflects the total content of lignin in the sample. 13C NMR spectra were recorded for HA preparations. The preparations were isolated from the studied soils, including iron–manganese nodules, by triple extraction with a 0.1 M NaOH + 0.4 M NaF mixture from the samples decalcified with 0.05 N H2SO4. [6]. The dialyzed and frozen-out HA preparations (50 mg) were dissolved in 0.6 mL of 0.3 M NaOD/D2O. Spectra were recorded on a Bruker Avance DRX 500 NMR spectrometer at 25.18 MHz and 290 K.

The isotope composition of carbon in soil organic matter was determined on a Thermo V Plus isotope ratio mass spectrometer and a Thermo Flash 1112 elemental analyzer [8]. Radiocarbon analyses of extractable humic acids were carried out at the Kyiv Radiocarbon laboratory.
3. Results and discussion

Although insufficient data are available about the biochemical compositions of different organisms participating in humification, it is considered established [2] that different types of plant tissues (gymnospermous and angiospermous, woody and nonwoody, cereal and herbaceous) have contrast lignin parameters. Our earlier studies [4, 5, 7] confirmed the revealed regularities and found three known lignin types in different samples of plant tissues. The first type is the lignin of coniferous plants (soft wood lignin), which contains vanillyl (guaiacyl) phenols as the main structural units. The content of cinnamyl (coumaryl and ferulyl) phenols in tissues of gymnospermous plants is low; syringic acids and aldehydes are almost absent, and the S/V ratio is 0. The second type is the lignin of deciduous plants (hard wood lignin) mainly consisting of similar amounts of vanillyl (guaiacyl) and syringyl structures. The lignin in the tissues of deciduous species (beech, chestnut) contains similar amounts of vanillyl and syringyl phenols. The C/V ratios are higher than 0 but lower than 1. The content of cinnamyl (p-coumaryl) phenols is close to 0. In the tissues of deciduous plants from the temperate zone (Moscow region), the V: S: C composition ratio is 3:4:1 for smallleaved species, 6:5:1 for broadleaved species (Tula region) [4]. The third type is the lignin of herbaceous plants containing the largest amounts of cinnamyl structural units (coumaric and ferulic acids), which exceeds their content in woody tissues by 4–6 times. The content of syringyl phenols in herbs is similar to their content in the wood of angiospermous plants but exceeds that in leaves by 5–6 times. Ferulic acids are associated with hemicelluloses in the cell walls of grasses; similar amounts of coumaryl and ferulyl phenols are found in herbaceous tissues [5].

Thus, as shown in our earlier series of publications [4, 5, 7], the nature of plant tissues in soils is clearly identified, and the oxidation of lignin biopolymers and their chromatographic separation into simple phenols provide information about the types of plant tissues in soils: the S/V ratio may be used for the separation of gymnospermous and angiospermous plant tissues, the C/V ratio for the separation of organic materials of woody and nonwoody origins, and the K/F ratio for the identification of herbaceous and cereal plant residues.

The $^{13}$C NMR spectroscopic study of HA preparations showed that the lignin of higher plants is involved in the formation of specific humus compounds. In general, HAs are similar to soil samples in the content of lignin oxidation products and lignin parameters [7]; they inherit characteristic properties of plant tissues but show a better ordering of structural lignin fragments. Their proportions of vanillyl, syringyl, and cinnamyl phenols strongly vary among the tissue types but are similar within the homotypic ecosystems (1:1:1 for agrocenoses and 3:4:1 for woody cenoses). These facts are embodied in the NMR spectra (figure 1). It is known that the peaks at 147 ppm (region of aromatic bonds) and 56 ppm (region of aliphatic bonds) in the $^{13}$C NMR spectra of HAs are due to compounds of lignin origin. The coefficient of correlation between the content of lignin (VSC) in the humus soil horizons of humid landscapes and the $^{13}$C NMR peak area of the chemical bonds of lignin origin is 0.94 for the aliphatic region of the HA spectrum (56 ppm) and 0.93 for the aromatic region of the HA spectrum (147 ppm).

![Figure 1](image_url)  
**Figure 1.** Comparison of the $^{13}$C NMR spectra of native lignin preparations and soil humic acids from (a) woody plants and Light-gray forest soil and (b) cereal crops and Dark agrogray soil.
The comparison of the $^{13}$C NMR spectra of native lignin preparations from different woody and herbaceous plant species obtained by Karmanov and Kocheva [3], with the spectra of our HA preparations [3], revealed that the signals inherited by HAs from plant tissues are also clearly identified at 102, 115, 119, 131, 152, and 160 ppm (table 1). The sets of peaks vary in the plant species; hence, they should also vary in the HAs from different soils; i.e., the chemical structures of lignin from coniferous and deciduous or cereal and herbaceous species differ not only in the content of the main types of intermonomer bonds $\beta$–$\beta$–$\beta$, $\beta$–$\gamma$–$\beta$, and $\gamma$–$\gamma$–$\gamma$, but also in the composition of macromolecules, i.e., the proportions of the three types of monomeric units: vanillyl (guaiacyl), syringyl, and pcoumaryl ones (V : S : C). For example, the characteristic signals of the pcoumaryl units from herbs include characteristic peaks at 131.4–131.5 ppm (C-2 and C-6 in phloxyphenylpropene units) and signals at 160.1 and 166.7 ppm related to the C-4 and C-6 atoms in esters of $p$-coumaric structures. The signals at 53.5 and 53.8 ppm indicate the presence of coumaran and pinoresinol structures (in the macromolecules, monomeric units are linked by the $\beta$–$\gamma$ (phenylcoumaran) and $\beta$–$\beta$ (pinoresinol) bonds) [6].

| Table 1. Signals of lignin units in the $^{13}$C NMR spectra of soil humic acids. |
|-----------------------------------------------|
| Soil, horizon, cm  | 53.5–53.8 | 55.7 | 55.9 | 102–104 | 115.5 | 119.1 | 131.4–131.5 | 151.0–152.0 | 160.1–166.7 |
|                  | ppm coumarans | ppm pinoresinol | ppm syringyl | ppm vanilly | ppm C-2 | ppm C-5 | ppm p-coumaryl | ppm C-3, C-5-OCH$_3$ |
|                  | Gleys light gray soil, A 0–10 | + | + | - | - | + | - | - | - |
|                  | Moscow region |  |  |  |  |  |  |  |  |
|                  | Gray soil, A 0–30 | - | - | + | + | + | + | + | + |
|                  | Bryansk region |  |  |  |  |  |  |  |  |
|                  | Dark gray soil, A 0–30 | + | + | + | + | + | + | + | + |
|                  | G 60–70 | + | + | + | + | + | + | + | + |

The signals in the region of 100–160 ppm are due to the presence of aromatic structural units in HAs from soils under forest. This region can be subdivided into four intervals: 100–117 ppm, signals of tertiary aromatic carbon atoms located in the ortho position to the C atoms with an oxygen function (C-2 and C-5 in the noncondensed vanillyl (guaiacyl) units or C-2 and C-6 in the syringyl units); 117–125 ppm, signals of tertiary aromatic carbon atoms without C atoms with an oxygen function in the ortho position (C-2 and C-6 in the phloxyphenylpropene units and C-6 in the vanillyl (guaiacyl) units); 125–142 ppm, signals of quaternary aromatic carbon atoms, mainly C-1 and C-5; and 142–160 ppm, signals due to etherified aromatic carbon atoms. In the $^{13}$C NMR spectra of HAs from soils of coniferous forest, characteristic signals are observed at 152.1–152.4 ppm, which are due to the C-3 and C-5 atoms bound to the OCH$_3$ groups. The unsubstituted C-2 and C-6 atoms in the syringyl units also give signals at 102–104 ppm, and the signals at 119 ppm indicate the presence of the vanillyl (guaiacyl) units (C-6). In the spectrum of coniferous lignin, there are nopeaks at 152.2 and 152.5 ppm, which are due to the C-3 and C-5 atoms bound to the OCH$_3$ groups. However, the uncondensed units of coniferous lignin contain a significant number of the C-5 atoms, which give a characteristic shift at 115.6 ppm. The general view of HA spectra (figure 1) shows that lignin from plants of the southern taiga zone becomes a source of more developed, elongated HA molecules with a well-developed
aliphatic moiety (high peaks in the aliphatic region of the spectrum); feruryl and coumaryl phenols from steppe plants form compact structures of chernozem HAs with a well-defined aromatic moiety. Thus, different types of plants and plant communities produce phylogenetically specific and strictly individual set of phenolic acids and aldehydes, which determine the structural features of HAs and may be used as biomarkers in the study of buried soils and the paleoecological reconstructions of past regional environments. We shall present several examples.

It follows from the data presented in table 2 and figure 2 that tissues of angiospermous (woody and herbaceous) plants participate in the formation of humus in the surface and second humus horizons of gray forest soils. Their lignin consists of similar amounts of syringyl and vanillyl structures, which is typical for the tissues of woody species and is embodied in the value of the S/V ratio (about 0.8) (table 2).

**Table 2.** Lignin parameters in Gray soils of the Tul’skie Zaseki Reserve, mg/g C<sub>org</sub>.

| Soil             | Horizon, depth, cm | VSC, mg g<sup>-1</sup>C<sub>org</sub> | (ac/al)<sub>v</sub> | (ac/al)<sub>s</sub> | S/V  | δ<sup>13</sup>C, ‰ |
|------------------|--------------------|---------------------------------------|---------------------|---------------------|------|-------------------|
| Gray forest soil | A 0–10             | 14.60                                 | 0.27                | 0.40                | 0.80 | -26.91            |
|                  | Bt1 50–60          | 2.30                                  | 0.25                | 0.24                | 0.55 | -                 |
| Gray forest soil | A 0–10             | 12.20                                 | 0.24                | 0.37                | 0.79 | -26.73            |
|                  | Bt2 50–60          | 1.20                                  | 0.36                | 0.48                | 0.99 | -                 |

This ratio is also observed in the soil surface horizons of the Tul’skie Zaseki Reserve (0.7–0.8) (table 2), as well as in the second humus horizons of soils of the Bryansk region (0.7–0.9). The radiocarbon age of the latter soils is 2180 ± 60 years. The vanillyls:syringyls:coumaryls ratios are about 6:5:1 and 5:3:1, respectively, which corresponds to the residues of nonwoody tissues of such broad-leaved species as maple and alder. The cinnamyl (styryl and ferulyl) phenols are mainly extracted from herbaceous tissues; therefore, their complete disappearance from the second humus horizon can also indicate the contribution of woody species to humus formation in the late Holocene. The herbaceous genesis of cultivated cereal residues was revealed in the plow horizons of soils of the Bryansk region (V: S: C = 1:2.1; C/F = 0.79–0.95) and in the humus horizon of agrogray soil on the microdepression. The predominance of herbaceous elements was established from the values C/V = 1.5 and C/F = 1.7 with the domination of coumaryl units in the gley soils of hydromorphic landscape positions. Peak identifiers of cereals at 53.5 and 113.1 ppm are also present in the NMR spectra of HAs from the plow horizons of agrogray soil (Table 1). However, the second humus horizon of soil in the microdepression has no peaks of coumaryl structures (at 113.1 and 166.7 ppm) typical for cereals. On the contrary, the NMR spectra of HAs contain signals of syringyl and vanillyl structures, which confirm the participation of soft wood lignin in the formation of their humus. The isotope mass spectrometry data also reveal contrast isotope profiles for agrogray soils of microdepressions and microelevations (Table 3, figure 2).

The lightening of humus isotope composition in the second humus horizon of the soil in the microdepression to −28.4‰ confirms the conclusion about the participation of deciduous wood tissues in humus formation 2180 ± 60 years ago (Ki-17415) (table 3). A similar isotope composition of humus (−27.6‰) is found in the surface horizons of gray forest soils 330 ± 50 years old (Ki-18777) under the deciduous forest of Bryansk Forest Reserve. At the same time, cinnamyl phenols make heavier the isotope composition of humus carbon to −26.1–26.4‰ in the recent plow horizons and the horizons of the Atlantic age (6690 ± 110, Ki-18776) and to −24.7‰ in the lower part of the humus horizon of agrogray soil on the microdepression. The herbaceous meadow composition of plants is proved by the δ<sup>13</sup>C values of −29.49 to −32.69‰ in the second humus horizons of soils in the hydromorphic positions of the landscapes.
Figure 2. Products of lignin oxidation (% of total phenols) in recent (Ap, 0–33 cm) and second humus ([A], 33–63 cm) horizons of gray forest soils of the Bryansk region.

Table 3. Lignin parameters in Agrogray soil of Bryansk district, mg/g C$_{org}$.

| Horizon, depth, cm | VSC, mg $\delta^{13}$C$_{org}$ | Coumaric acids | Age, $^{14}$C | S/V | C/V | C/F | V:S:C | $\delta^{13}$C, % |
|--------------------|---------------------------------|-----------------|--------------|-----|-----|-----|-------|-----------------|
| Ap 0–33            | Agrogray soils with a second humus horizon (microdepression) | 9.09            | 0.81         | -   | 2.80| 0.85| 0.95  | 1:3:1           | -26.41          |
| A[hh] 33-63        | 0.72                           | 0.00            | 2180 ± 60    | 0.91| 0.00| 0.00| 6:5:1 | -28.41          |
| Ap 0-20            | Agrogray soils with a second humus horizon (microdepression) | 6.77            | 0.98         | -   | 2.13| 1.41| 0.88  | 1:2:1           | -26.10          |
| A[hh] 50-80        | 0.79                           | 0.00            | 6690 ± 110   | 0.65| 0.00| 0.00| 5:3:1 | -24.10          |
|                    | Gleyic dark gray soil (ravine bottom) |                 |              |     |     |     |       |                 |
| A 0-32             | 12.49                          | 1.54            | -            | 1.50| 0.94| 0.83| 1:2:1 | -29.49          |
| G 32-54            | 4.99                           | 0.94            | -            | 2.46| 1.49| 1.68| 1:2:1 | -32.69          |
|                    | Agrogray soil (microelevation) |                 |              |     |     |     |       |                 |
| Ap 0-30            | 5.80                           | 0.59            | -            | 2.84| 1.16| 0.79| 1:3:1 | -26.11          |
| EB 30-55           | 1.85                           | 0.74            | -            | 3.09| 5.43| 2.39| 1:2:5 | -25.39          |

In the agrogray soils of the southern taiga (plowland in the place of a former forest), the S/V value is slightly higher: about 1 (table 4). This corresponds to the values typical for fresh birch tissues, which are characterized by similar proportions of syringyl and vanillyl units [2]. The development time of these former forest soils was apparently insufficient for the complete transformation of decomposition-resistant lignin compounds in woody tissues. Moreover, the composition proportions of lignin phenols typical for the birch ecosystems of southern taiga remained unchanged not only in the soils, but also in nodules, where the age of humus is from 1600 to 1990 years. The isotope composition of organic carbon in nodules (−26.7‰) agrees with the paleoecological scenario of their formation under variable moisture conditions 1990 ± 90 years ago, which correspond to the current ecosystem of small-leaved (birch or aspen) forest. However, the isotope compositions of some nodule fractions is lightened because of the presence of coumaric structures from meadow plants to −27.41‰, which indicates more hydromorphic conditions of their formation 1600 ± 80 years ago that replaced the drier period (−26.7‰) of about 2000 years ago. The contents of lignin compounds (VSC) are regularly accumulated in the gley soils and in the second humus horizons and reflect the degree of soil hydromorphism.
Thus, analysis of the presented data on the composition of lignin phenols in the recent and buried horizons of different polygenetic soils, reveals that the composition of lignin phenols in soils serves as a molecular trace of land vegetation. In spite of the significant number of lignin peaks in the NMR spectra of HAs, most of them allow only the qualitative identification of biomarker and the fixation of changes in the natural environment during the time intervals determined by radiocarbon dating.

Acknowledgements
We are grateful to Russian Scientific Fund for the financial support of this investigation (grant 17-14-01120) and professor W. Zech (Bayreuth University of FRG)

References
[1] Amelung W 1999 Zum Klimaeinfluß auf die Organische Substanz Nordamerikanischer Praärieböden. Bayreuth. p 131
[2] Ertel J R and Hedges J I 1984 The lignin component of humic substances: distribution among the soil and sedimentary humic, fulvic and baseinsoluble fractions Geochim. Cosmochim. Acta 48 pp 2065–2074
[3] Karmanov A P, Kocheva I S and Borisenkov M F 2013 Antioxidant properties of lignins Vestnik of Biology Institute, Komi SC UrDRAS 10-11 pp 31–38. (in Russian)
[4] Kovalev I V and Kovaleva N O 2008 Biochemistry of lignin in soils of periodic excessive moistening (from the example of agrogray soils) Eurasian Soil Science vol. 41, 10 pp 1066–1076
[5] Kovaleva N O and Kovalev I V 2002 Aromatic Struktures of Lignin in the Organic Matter of Gray Forest Soils Moscow University Bulletin 17. Soil Science 2 pp 23–29
[6] Kovalev N O and Kovalev I V 2003 Peculiarities of organic matter of Fe-Mn nodules in gray forest soils determined by C NMR-spectroscopy Moscow University Bulletin 17. Soil Science, 2 pp 23–29
Kovaleva N O and Kovalev I V 2009 Transformation of lignin in surface and buried soils of mountainous landscapes Eurasian Soil Science vol. 42, 11 pp 1270–1281
[7] Schatz A., Zech M., Buggle B., Gulyas S., Hambach U., Markovic S., Sumegi P., and Scholten T. 2011 The Quaternary loess record of Tokaj, Hungary – reconstructing paalaeoenvironment, vegetation and climate using stable C and N isotopes and biomarkers Quaternary International 240 (1-2) pp 52–61

Table 4. Lignin parameters in Agrogray soil of Moscow district, mg/g Corg.

| Soil                          | VSC, mg g⁻¹ Corg | Age, ¹³C | V:S:C | S/V   | δ¹³C, %o |
|-------------------------------|------------------|----------|-------|-------|----------|
| Agrogray soil (micro elevation) |                  |          |       |       |          |
| Fe-Mn nodules, (1-2 mm)       | 1.29 ± 0.51      | -        | 2:2:1 | 0.85  | -26.70   |
| Gleyic agrogray in micro depression | 11.97 ± 0.80 | -        | 2:2:1 | 1.02  | -27.08   |
| Fe-Mn nodules, (1–2 mm)       | 1.34 ± 0.02      | 1990 ± 90 (Ki–17759) | 1.6:1:5:1 | 0.93 | -26.73 ± 0.13 |
| Fe-Mn nodules, (2–3 mm)       | 1.41 ± 0.30      | 1600 ± 80 (Ki–17411) | 1.6:2:1 | 1.16 | -26.76 ± 0.36 |
| Fe-Mn nodules, (3–5 mm)       | 0.81 ± 0.02      | 1690 ± 110 (Ki–17412) | 1.8:3:1 | 1.51 | -26.89 ± 0.36 |