

Supporting Information

Quantitative structure-activity relationship (QSAR) study of carcinogenicity of polycyclic aromatic hydrocarbons (PAHs) in atmospheric particulate matter by random forest (RF)

Nan Li,^a Juan Qi,^a Ping Wang,^a Xin Zhang,^a Tianlong Zhang,^{*a} Hua Li,^{*a,b}

^a Key Laboratory of Synthetic and Natural Functional Molecular Chemistry of Ministry of Education, College of Chemistry and Material Science, Northwest University, Xi'an, 710127, China. E-mail: tlzhang@nwu.edu.cn

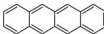
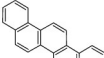
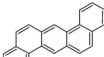
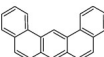
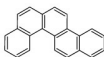
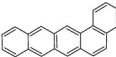
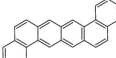
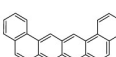
^b College of Chemistry and Chemical Engineering, Xi'an Shiyou University, Xi'an 710065, China. E-mail: huali@nwu.edu.cn

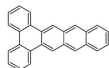
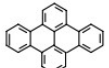
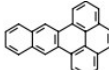
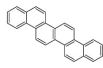
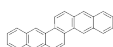
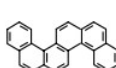
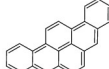
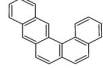
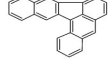
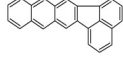
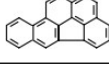
Contents

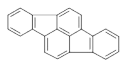
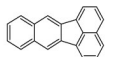
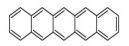
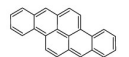
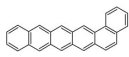
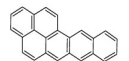
Table S1 The information of the 91 PAHs

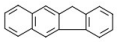
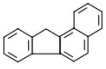
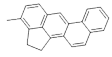
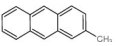
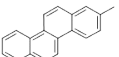
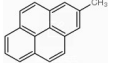
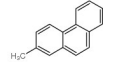
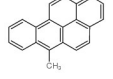
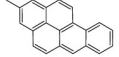
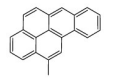
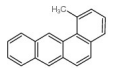
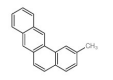
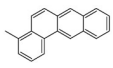
Table S2 The molecular descriptors of Anthracene (1#)

Table S1 The information of the 91 PAHs

No.	Name	CAS#	Molecular formula	Structure	Carcinogenicity
1	Anthracene	120-12-7	C ₁₄ H ₁₀		1
2	Naphthacene	92-24-0	C ₁₈ H ₁₂		1
3	Benz[a]anthracene	56-55-3	C ₁₈ H ₁₂		1
4	Chrysene	218-01-9	C ₁₈ H ₁₂		1
5	Pyrene	129-00-0	C ₁₆ H ₁₀		1
6	Benzo[c]phenanthrene	195-19-7	C ₁₈ H ₁₂		1
7	Perylene	198-55-0	C ₂₀ H ₁₂		1
8	Benzo[a]pyrene	50-32-8	C ₂₀ H ₁₂		1
9	Benzo[e]pyrene	192-97-2	C ₂₀ H ₁₂		1
10	Picene	213-46-7	C ₂₂ H ₁₄		1
11	Pentaphene	222-93-5	C ₂₂ H ₁₄		1
12	Benzo[b]chrysene	214-17-5	C ₂₂ H ₁₄		1
13	Dibenz[a,h]anthracene	53-70-3	C ₂₂ H ₁₄		1
14	Dibenz[a,j]anthracene	224-41-9	C ₂₂ H ₁₄		1
15	Benzo[b]triphenylene	215-58-7	C ₂₂ H ₁₄		1
16	Benzo[c]chrysene	194-69-4	C ₂₂ H ₁₄		1
17	Dibenzo[c,g]phenanthrene	188-52-3	C ₂₂ H ₁₄		1
18	Benzo[a]naphthacene	226-88-0	C ₂₂ H ₁₄		1
19	Dibenzo[a,j]naphthacene	227-04-3	C ₂₆ H ₁₆		1
20	Dibenzo[a,l]naphthacene	226-86-8	C ₂₆ H ₁₆		1

21	Dibenzo[a.c]naphthacene	216-00-2	C ₂₆ H ₁₆		1
22	Dibenzo[el]pyrene	192-51-8	C ₂₄ H ₁₄		1
23	Dibenzo[de.qr]tetracene	193-09-9	C ₂₆ H ₁₄		1
24	Dibenzo[g.p]chrysene	191-68-4	C ₂₆ H ₁₆		1
25	Benzo[c]picene	217-37-8	C ₂₆ H ₁₆		1
26	Benzo[ghi]perylene	191-24-2	C ₂₂ H ₁₂		1
27	Dibenzo[b.k]chrysene	217-54-9	C ₂₆ H ₁₆		1
28	Dibenzo[cl]chrysene	—	C ₂₆ H ₁₆		1
29	Benzo[b]perylene	197-70-6	C ₂₄ H ₁₄		1
30	Naphtho[2.3-g]chrysene	196-64-5	C ₂₆ H ₁₆		1
31	Benzo[h]pentaphene	214-91-5	C ₂₆ H ₁₆		1
32	Coronene	191-07-1	C ₂₄ H ₁₂		1
33	Naphtho[1.2.3.4-def]chrysene	192-65-4	C ₂₄ H ₁₄		1
34	Dibenzo[def.p]chrysene	191-30-0	C ₂₄ H ₁₄		1
35	Benzo[rst]pentaphene	189-55-9	C ₂₄ H ₁₄		1
36	Benzo[g]chrysene	196-78-1	C ₂₂ H ₁₄		1
37	Dibenzo[b.g]phenanthrene	195-06-2	C ₂₂ H ₁₄		1
38	Dibenz[a.e]aceanthrylene	5385-75-1	C ₂₄ H ₁₄		1
39	Dibenzo[a.k]fluoranthene	84030-79-5	C ₂₄ H ₁₄		1
40	Naphtho[2.3-k]fluoranthene	207-18-1	C ₂₄ H ₁₄		1
41	Fluoranthene	206-44-0	C ₁₆ H ₁₀		-1
42	Dibenzo[k.mno]fluoranthene	203-25-8	C ₂₂ H ₁₂		1

43	Benzo[c]fluorene	205-12-9	C ₁₇ H ₁₂		1
44	Benzo[ghi]fluoranthene	203-12-3	C ₁₈ H ₁₀		1
45	Benzo[a]aceanthrylene	203-33-8	C ₂₀ H ₁₂		1
46	Indeno[1.2.3.cd]pyrene	193-39-5	C ₂₂ H ₁₂		1
47	Indeno[1.2.3-cd]fluoranthene	193-43-1	C ₂₂ H ₁₂		1
48	Cyclopenta[cd]pyrene	27208-37-3	C ₁₈ H ₁₀		1
49	Benzo[j]fluoranthene	205-82-3	C ₂₀ H ₁₂		1
50	Benzo[k]fluoranthene	207-08-9	C ₂₀ H ₁₂		1
51	Dibenz[e.k]acephenanthrylene	206-06-4	C ₂₄ H ₁₄		1
52	Benzo[b]fluoranthene	205-99-2	C ₂₀ H ₁₂		1
53	Naphthalene	91-20-3	C ₁₀ H ₈		-1
54	Phenanthrene	85-01-8	C ₁₄ H ₁₀		-1
55	Triphenylene	217-59-4	C ₁₈ H ₁₂		-1
56	Pentacene	135-48-8	C ₂₂ H ₁₄		-1
57	Dibenzo[b.def]chrysene	189-64-0	C ₂₄ H ₁₄		-1
58	Dibenzo[def.mno]chrysene	191-26-4	C ₂₂ H ₁₂		-1
59	Benzo[a]perylene	191-85-5	C ₂₄ H ₁₄		-1
60	Dibenzo[de.mn]naphthacene	214-63-1	C ₂₆ H ₁₄		-1
61	Benzo[a]pentacene	239-98-5	C ₂₆ H ₁₆		-1
62	Naphtho[2.1.8-qla]naphthacen	196-42-9	C ₂₄ H ₁₄		-1
63	Acenaphthylene	208-96-8	C ₁₂ H ₈		-1
64	1,2-Dihydroacenaphthylene	83-32-9	C ₁₂ H ₁₀		-1

65	9H-Fluorene	86-73-7	C ₁₃ H ₁₀		-1
66	Benzo[b]fluorene	243-17-4	C ₁₇ H ₁₂		-1
67	Benzo[a]fluorene	238-84-6	C ₁₇ H ₁₂		-1
68	3-Methylcholanthren	56-49-5	C ₂₁ H ₁₆		1
69	1-Methylphenanthrene	832-69-9	C ₁₅ H ₁₂		1, -1
70	9-Methylanthracene	779-02-2	C ₁₅ H ₁₂		-1
71	2-Methylanthracene	613-12-7	C ₁₅ H ₁₂		-1
72	2-Methylchrysene	3351-32-4	C ₁₉ H ₁₄		-1
73	5-Methylchrysene	3697-24-3	C ₁₉ H ₁₄		1
74	6-Methylchrysene	1705-85-7	C ₁₉ H ₁₄		1
75	12-Methylbenzo[a]anthracene	2422-79-9	C ₁₉ H ₁₄		1
76	2-Methylpyrene	3442-78-2	C ₁₇ H ₁₂		-1
77	1-Methylpyrene	2381-21-7	C ₁₇ H ₁₂		1
78	2-Methylphenanthrene	2531-84-2	C ₁₅ H ₁₂		-1
79	6-Methylbenzo[a]pyrene	2381-39-7	C ₂₁ H ₁₄		1
80	2-Methylbenzo[a]pyrene	16757-82-7	C ₂₁ H ₁₄		1
81	12-Methylbenzo[a]pyrene	4514-19-6	C ₂₁ H ₁₄		1
82	1-Methylbenzo[a]anthracene	2498-77-3	C ₁₉ H ₁₄		-1
83	2-Methylbenzo[a]anthracene	2498-76-2	C ₁₉ H ₁₄		1
84	4-Methylbenzo[a]anthracene	316-49-4	C ₁₉ H ₁₄		-1
85	6-Methylbenzo[a]anthracene	316-14-3	C ₁₉ H ₁₄		1
86	7-Methylbenzo[a]anthracene	2541-69-7	C ₁₉ H ₁₄		1

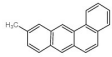
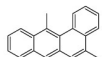
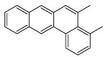
87	10-Methylbenzo[a]anthracene	2381-15-9	C ₁₉ H ₁₄		1
88	5,12-Dimethylbenzo[a]anthracene	21297-22-3	C ₂₀ H ₁₆		-1
89	6,8-Dimethylbenzo[a]anthracene	317-64-6	C ₂₀ H ₁₆		1
90	4,5-Dimethylbenzo[a]anthracene	18429-70-4	C ₂₀ H ₁₆		1
91	1-Methyltriphenylene	2871-91-2	C ₁₉ H ₁₄		-1

Table S2 The molecular descriptors of Anthracene (1#)

Categories	Descriptors	Values
Structure descriptors	Sum of atomic polarizabilities	17.81
	Mean atomic polarizability	0.74
	Mean electrotopological state	1.9
	Number of atoms	24
	Number of non-H atoms	14
	Number of bonds	26
	Number of Hydrogen atoms	10
	Number of non-H bonds	16
	Number of multiple bonds	16
	Aromatic ratio	1
	Number of rings	3
	Number of circuits	6
	Number of aromatic bonds	16
	Number of Carbon atoms	14
	...	...
Topology descriptors	First Zagreb index M1	76
	Second Zagreb index M2	90
	Quadratic index	13
	Pogliani index	28
	Ramification index	4
	Polarity number	21
	Xu index	13.815
	Superpendentic index	0
	Wiener W index	279
	Mean Wiener index	3.066
	Harary H index	25.316
	Detour index	1015
	Hyper-detour index	6272
	Kier symmetry index	27.303
	Eccentricity	80
	Average eccentricity	5.714
	Eccentric	0.898
	Unipolarity	33
	Centralization	96
	Variation	16
	Balaban centric index	0
	Lopping centric index	0
	...	...

Molecular connectivity index	Connectivity index chi-0	9.38
	Connectivity index chi-1	6.933
	Connectivity index chi-2	6.081
	Connectivity index chi-3	5.344
	Connectivity index chi-4	4.588
	Connectivity index chi-5	4.052
	Solvation connectivity index chi-0	9.38
	Solvation connectivity index chi-1	6.933
	Solvation connectivity index chi-2	6.081
	Solvation connectivity index chi-3	5.344
	Solvation connectivity index chi-4	4.588
	Solvation connectivity index chi-5	4.052
Geometric descriptors	...	...
	3D-Wiener index	1167.951
	3D-Balaban index	1.857
	3D-Harary index	85.634
	D/D index	301.115
	Span R	4.778
	Average span R	0.446
	Molecular eccentricity	1
	Spherosity	1
	Asphericity	0.574
	Folding degree index	1
	3D Petitjean shape index	0.798
	Jug RC index	1.411
	Aromaticity index	0.987
	HOMA total	15.061
	...	...