Hierarchical clustering analysis of blood plasma lipidomics profiles from mono- and dizygotic twin families

Harmen H.M. Draisma,[1,6,*] Theo H. Reijmers,[1] Jacqueline J. Meulman,[2] Jan van der Greef,[1] Thomas Hankemeier,[1,3,5] Dorret I. Boomsma[4,5]

Supplemental Material

[1]Division Analytical Biosciences, Leiden/Amsterdam Center for Drug Research, Einsteinweg 55, NL-2333CC Leiden, The Netherlands. Phone: +31-71-5274220. Fax: +31-71-5274277

[2]Leiden University, Mathematical Institute, Niels Bohrweg 1, room 224A, NL-2300RA Leiden, The Netherlands. Phone: +31-71-5273830

[3]Netherlands Metabolomics Centre, Einsteinweg 55, NL-2333CC Leiden, The Netherlands. Phone: +31-71-5274220. Fax: +31-71-5274277

[4]Department of Biological Psychology, VU University Amsterdam, Van der Boechorststraat 1, room 2B-41, NL-1081 BT Amsterdam, The Netherlands. Phone: +31-20-5988787. Fax: +31-20-5988832

[5]Both authors contributed equally to this study.

[6]Present address: Department of Biological Psychology and the EMGO Institute for Health and Care Research, Faculty of Psychology and Education, VU University Amsterdam, Van der Boechorststraat 1, room 2B-09, NL-1081 BT Amsterdam, The Netherlands. Phone: +31-20-5986871. Fax: +31-20-5988832

[*]Corresponding author. E-mail: hhm.draisma@psy.vu.nl

Table S1: Pearson correlations between cophenetic distances and original Euclidean distances among all study participants, before and after equating

Clustering algorithm	Before equating	After equating
Average linkage	0.75	0.60
Ward's method	0.64	0.52
Single linkage	0.60	0.12
Complete linkage	0.68	0.53
McQuitty's method	0.72	0.52
Median linkage	0.15	0.10
Centroid linkage	0.29	0.01

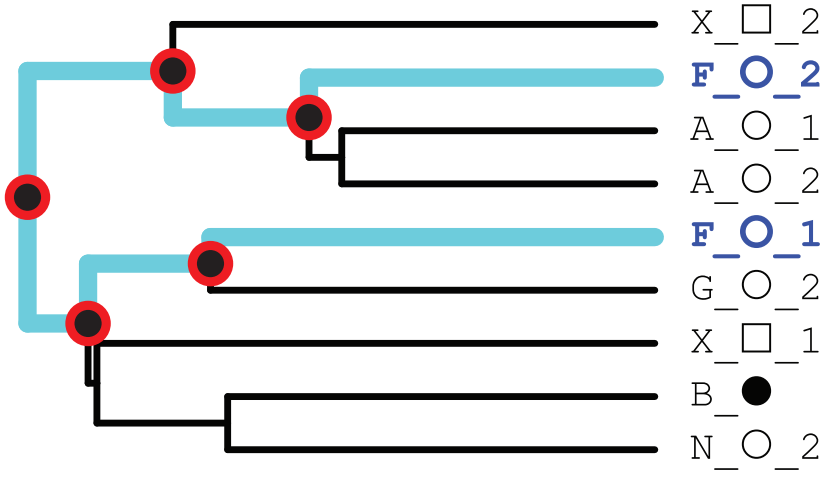

Figure S1: Example of 'node analysis': the distance between relatives in the hierarchical clustering dendrogram is assessed by counting the number of separating nodes. In this example, the highlighted nodes and edges illustrate that the pair of objects "F_○_1" and "F_○_2" is separated by five nodes.

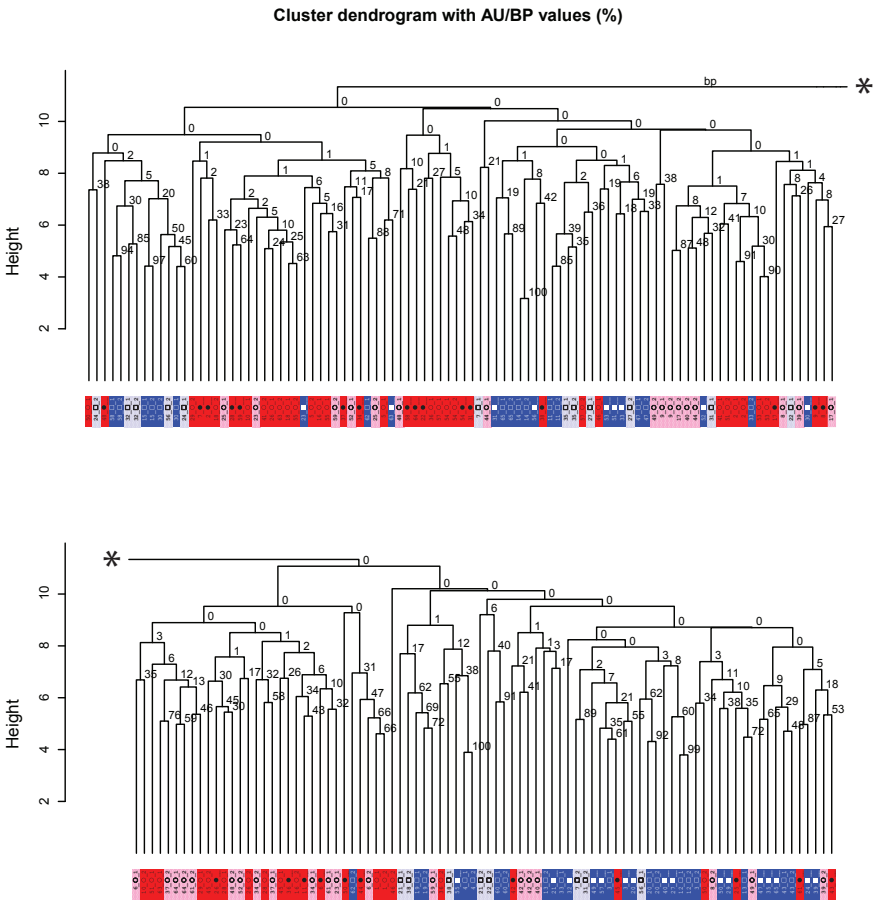

Figure S2: Clustering dendrogram on the basis of combined equated B1–B2 data sets, with associated probability values based on nonparametric bootstrap procedure. Numbers near the branching points in the dendrogram indicate bootstrap probability (bp) values; high values indicate high stability of the corresponding node during bootstrapping. The dendrogram structure in this Figure is equal to that of the dendrogram displayed at the top of the heatmap in Figure 1B in the main document. The dendrogram was split at the highest level (as indicated by asterisks) to enhance the legibility of the object labels. Participants are denoted as follows: the family identifier (1–65) is followed by a square (□, for males) or a circle (○, for females) to indicate the sex of the participant, and, in case of twins, a "1" or a "2" to indicate the first and second members of the twin pair, respectively. Nontwin siblings are indicated by filled squares (■) or filled circles (●) for males and females, respectively. For the participants from B1, see Table S5 for a comparison between the labeling as used in Draisma *et al.* (2008)[1] and the labeling used in the current manuscript.

Table S2: Numbers of monozygotic co-twins (A), dizygotic co-twins (B), and sex-matched nontwin siblings (C) separated by particular numbers of nodes, with respect to chance observation[1]

A

I	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
II	15	1	1	0	0	1	1	1	0	1	4	3	2	1	1	0	2	2	0	1	0	0	0	0	0
III	0.0*	15.0	20.8	100.0	100.0	46.8	57.6	69.3	100.0	81.4	17.7	43.2	74.8	95.2	97.6	100	87.4	82.0	100.0	85.0	100.0	100.0	100.0	100.0	100.0
IV	0.00	0.30	0.42	0.00	0.00	0.51	0.50	0.49	0.00	0.39	0.34	0.48	0.42	0.20	0.14	0.00	0.34	0.38	0.00	0.40	0.00	0.00	0.00	0.00	0.00

B

I	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
II	4	0	0	0	2	3	1	2	3	0	0	1	1	1	1	0	1	3	3	0	1	1	0	0	0
III	0.0*	100.0	100.0	100.0	4.4*	1.1*	47.7	21.6	4.3*	100.0	100.0	84.8	87.3	89.9	93.9	100.0	93.6	40.5	26.6	100.0	60.1	44.7	100.0	100.0	100.0
IV	0.00	0.00	0.00	0.00	0.22	0.11	0.48	0.41	0.20	0.00	0.00	0.36	0.28	0.27	0.23	0.00	0.26	0.53	0.40	0.00	0.53	0.50	0.00	0.00	0.00

C

I	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
II	0	2	2	2	3	0	3	6	1	3	2	2	2	5	7	2	5	4	7	4	4	0	0	0	0
III	100.0	2.4*	4.8*	9.0	2.8*	100.0	12.7	0.7*	80.0	43.7	84.7	87.8	90.5	43.0	25.7	97.2	59.0	66.9	6.0*	28.9	10.0	100.0	100.0	100.0	100.0
IV	0.00	0.14	0.22	0.29	0.16	0.00	0.37	0.09	0.43	0.41	0.35	0.31	0.30	0.49	0.45	0.18	0.54	0.46	0.23	0.52	0.36	0.00	0.00	0.00	0.00

[1] The rows of each panel represent: number of nodes separating co-twins (row I); observed number of occasions where siblings are separated by the number of nodes as given in row I (row II); average p-value ($\times 100\%$) over 100 permutation tests (10 000 iterations per permutation test; direct comparison of the observed frequencies as in row II with the chance distribution generated by each permutation test) (row III); and standard deviation of the p-value ($\times 100\%$) as in row III, over the 100 permutation tests (row IV). Asterisks indicate average p-values <0.05

Table S3: Description of monozygotic twin pairs separated by only one node in the dendrograms of Figure 1B and Figure S2[1]

Twin pair	*Description*
1_○	Both co-twins had eaten rolls with jam and had drunk soft drink for breakfast at the day of sampling; furthermore, in the sample of 1_○_1 some hemolysis had occurred. Both co-twins had reported recent flu-like symptoms more than one week prior to sampling. Also, the menstrual cycles of both co-twins were not completely synchronous.
54_○	54_○_2 smoked 4 cigarettes per day at the time of sampling while 54_○_1 did not smoke. 54_○_1 and 54_○_2 had had a cold more than one week and more than one month prior to sampling, respectively.
58_□	Both co-twins used antihistamine as medication for chronic hay fever; 58_□_1 had suffered from hay fever in the week prior to sampling.
47_□	Both co-twins had had a cold more than one month prior to sampling.
11_□	Both co-twins had had a cold more than one month prior to sampling.
43_□	43_□_1 and 43_□_2 had had a cold more than one month and less than one week prior to sampling, respectively. Also, both co-twins had left their parents' home approximately half a year prior to sampling.
20_□	Both co-twins had had a cold more than one month prior to sampling.
2_○	2_○_1 and 2_○_2 had had a cold more than one month and more than one week prior to sampling, respectively. Furthermore, 2_○_2 suffered from allergy.
12_□	12_□_2 had eaten something during the fasting period. Both co-twins smoked at the time of sampling; 12_□_1 had been smoking 15 cigarettes/day for 3.5 years, whereas 12_□_2 had been smoking 8 cigarettes/day for 5 years. Furthermore, 12_□_1 had suffered from fatigue and headache more than one week prior to sampling, whereas 12_□_2 had suffered from flu accompanied by fever more than one month prior to sampling.
65_□	65_□_1 and 65_□_2 smoked 30 and 20 cigarettes/day at the time of sampling, respectively. Both co-twins had smoked less than one hour prior to sampling, and had had a cold less than one week prior to sampling.
15_□	Both co-twins had suffered from flu accompanied by fever more than one month prior to sampling.

[1]For an explanation of the labeling of families and participants, see the legend to Figure S2.

Table S3: Description of monozygotic twin pairs separated by one node (continued)

Twin pair	*Description*
60_□	Both co-twins had had muesli with diary products for breakfast at the day of sampling. 60_□_1 suffered from chronic back pain and had suffered from stomach flu accompanied by fever more than one month prior to sampling; 60_□_2 had had a cold more than one month prior to sampling.
53_○	53_○_1 and 53_○_2 had suffered from a cold and from stomach ache more than one month prior to sampling, respectively.
14_□	14_□_2 had eaten a roll for breakfast at the day of sampling whereas 14_□_1 had not. 14_□_1 had had a cold more than one week prior to sampling; 14_□_2 had suffered from flu accompanied by fever more than one month prior to sampling.
4_□	4_□_1 and 4_□_2 had had a cold less than one week and more than one month prior to sampling, respectively; furthermore, 4_□_1 suffered from allergy.

Table S4: Description of monozygotic twin pairs separated by more than one node in the dendrograms of Figure 1B and Figure S2[1]

Twin pair	*Description*
46_○	46_○_1 had reported sickness and headache more than 1 week prior to blood sampling. Both twins had synchronous menstrual cycles, although 46_○_2 appeared to suffer from oligomenorrhea.
3_□	3_□_1 had self-reportedly been ill without having a fever less than 1 week prior to blood sampling.
5_○	5_○_1 had smoked in the past (2 cigarettes/day) for half a year 1.5 years prior to blood sampling. Furthermore, 5_○_2 had had a cold less than one week prior to sampling. Also, the co-twins did not have completely synchronous menstrual cycles.
10_○	Both twins had self-reportedly suffered from a cold less than 1 week prior to blood sampling.
13_□	13_□_1 had had a cold less than 1 week prior to blood sampling.
62_□	62_□_2 had suffered from infectious mononucleosis more than 1 month prior to sampling. Moreover, during sample handling, in the sample of this twin hemolysis had occurred.
16_○	16_○_2 had been smoking five cigarettes per day for 6 years and had smoked 2 h before blood sampling; 16_○_1 had quit smoking a half year prior to sampling, after having smoked 10 cigarettes per day for 5 years. Furthermore, 16_○_2 had had a half cup of sugared tea for breakfast on the day of blood sampling. Both twins did not have synchronous menstrual cycles.
18_○	18_○_1 had self-reportedly suffered from flu-like symptoms less than 1 week prior to blood sampling. Both twins did not have synchronous menstrual cycles.
28_○	Twin 28_○_2 had been using the drug Fluoxetine for depression. Both twins did not have synchronous menstrual cycles.
30_□	30_□_2 had had a sip of cola during the fasting period prior to sampling. Both co-twins smoked at the time of sampling. 30_□_2 suffered from hay fever.
41_○	Both twins had self-reportedly been ill less than 1 week prior to blood sampling: 41_○_1 had suffered from a cold, whereas 41_○_2 had had flu-like symptoms accompanied by fever. 41_○_2 used oral contraceptives while 41_○_1 did not; furthermore, their menstrual cycles were not synchronous.
45_○	More than one week prior to sampling 45_○_1 had had a cold. 45_○_2 had suffered from stomach flu more than one week prior to sampling.

[1]For an explanation of the labeling of families and participants, see the legend to Figure S2.

Table S4: Description of monozygotic twin pairs separated by more than one node (continued)

Twin pair	*Description*
50_○	In the week prior to sampling, 50_○_2 had suffered from nausea and fatigue whereas 50_○_1 had not. 50_○_1 used terbinafine hydrochloride while 50_○_2 did not.
51_□	More than one month prior to sampling, 51_□_1 had had a cold and 51_□_2 had suffered from flu with fever, respectively.
55_○	Both co-twins did not have synchronous menstrual cycles. Furthermore, both co-twins had had a cold in the week prior to sampling.
57_○	57_○_1 suffered from chronic hay fever; 57_○_2 suffered from chronic asthma, for which she used budesonide/formoterol as medication.
63_□	63_□_1 had suffered from flu with fever and laryngitis in the week prior to sampling, for which she used feneticilline. Also, in the blood sample from 63_□_1 some hemolysis had occurred. 63_□_2 suffered from irritable bowel syndrome. Furthermore, 63_□_2 had smoked in the past (15 cigarettes/day), and had quit smoking two years prior to blood sampling after having smoked for two years.
26_○	26_○_1 had had a cold more than one week prior to sampling; 26_○_2 suffered from severe eczema for which she used a corticosteroid cream as a medication, from lymphedema in a leg, and from chronic respiratory disease.
29_○	In the blood sample of 29_○_2, hemolysis had occurred; furthermore, 29_○_2 had left her parents home about 4 months prior to sampling, while 29_○_1 had not. Both co-twins had had a cold in the week prior to sampling, and their menstrual cycles were not completely synchronous.
33_□	Both co-twins used fluticasone propionate as medication for slight asthma.
36_○	No tentative explanation for non-clustering on basis of available information
19_□	19_□_1 had been ill and 19_□_2 had had a cold more than one month prior to sampling, respectively

Table S5: Conversion table between labeling in this chapter and labeling in Draisma *et al.* (2008)[1] for families from B1[1]

Family label in this paper	Family label in Draisma *et al.* (2008)[1]
1_○	A_○
2_○	B_○
3_□	C_□
4_□	D_□
5_○	E_○
6_○	F_○
10_○	G_○
11_□	H_□
12_□	I_□
13_□	J_□
14_□	K_□
15_□	L_□
16_○	M_○
18_○	N_○
19_□	P_□
20_□	Q_□
21_□	R_□
28_○	S_○
30_□	T_□
41_○	U_○
46_○	V_○
60_□	W_□
62_□	X_□

[1] For an explanation of the labeling of families and participants, see the legend to Figure S2.

References

[1] Draisma HHM, Reijmers TH, Bobeldijk-Pastorova I *et al*: Similarities and differences in lipidomics profiles among healthy monozygotic twin pairs. *OMICS* 2008; **12**: 17–31.