Supplementary material

Ahmad et al., The effect of general adiposity and central body fat distribution on the circulating metabolome: a multi-cohort non-targeted metabolomics observational and Mendelian randomization study

Table S1 List of 97 BMI-associated genetic variants used for Mendelian randomization and their allele frequencies in TwinGene, PIVUS and ULSAM cohorts, respectively

Variant	Nearest Gene	Chr: Position	Effect/Other Allele	Published EAF	Published Beta	Published SE	TwinGene EAF	PIVUS EAF	ULSAM EAF
rs977747	*TAL1*	1:47,457,264	T/G	0.39	0.017	0.003	0.38	0.38	0.38
rs657452	*AGBL4*	1:49,362,434	A/G	0.39	0.023	0.003	0.38	0.38	0.40
rs11583200	*ELAVL4*	1:50,332,407	C/T	0.40	0.018	0.003	0.40	0.39	0.41
rs3101336	*NEGR1*	1:72,523,773	C/T	0.61	0.033	0.003	0.59	0.59	0.59
rs12566985	*FPGT-TNNI3K*	1:74,774,781	G/A	0.45	0.024	0.003	0.45	0.46	0.43
rs12401738	*FUBP1*	1:78,219,349	A/G	0.35	0.021	0.003	0.37	0.34	0.37
rs11165643	*PTBP2*	1:96,696,685	T/C	0.58	0.022	0.003	0.57	0.57	0.57
rs17024393	*GNAT2*	1:109,956,211	C/T	0.04	0.066	0.009	0.03	0.04	0.04
rs543874	*SEC16B*	1:176,156,103	G/A	0.19	0.048	0.004	0.22	0.23	0.21
rs2820292	*NAV1*	1:200,050,910	C/A	0.56	0.02	0.003	0.57	0.59	0.57
rs10182181	*ADCY3*	2:25,003,800	G/A	0.46	0.031	0.003	0.47	0.47	0.47
rs11126666	*KCNK3*	2:26,782,315	A/G	0.28	0.021	0.003	0.26	0.27	0.28
rs1016287	*LINC01122*	2:59,159,129	T/C	0.29	0.023	0.003	0.29	0.28	0.28
rs13021737	*TMEM18*	2:622348	G/A	0.83	0.06	0.004	0.83	0.83	0.84
rs11688816	*EHBP1*	2:62,906,552	G/A	0.52	0.017	0.003	0.55	0.54	0.51
rs2121279	*LRP1B*	2:142,759,755	T/C	0.15	0.025	0.004	0.15	0.14	0.15
rs1460676	*FIGN*	2:164,275,935	C/T	0.17	0.020	0.004	0.16	0.15	0.16
rs1528435	*UBE2E3*	2:181,259,207	T/C	0.63	0.018	0.003	0.64	0.65	0.63
rs17203016	*CREB1*	2:207,963,763	G/A	0.20	0.021	0.004	0.21	0.22	0.21
rs7599312	*ERBB4*	2:213,121,476	G/A	0.72	0.022	0.003	0.72	0.72	0.70
rs492400	*USP37*	2:219,057,996	C/T	0.42	0.016	0.003	0.42	0.41	0.43
rs2176040	*LOC646736*	2:226,801,046	A/G	0.37	0.014	0.003	0.37	0.36	0.40
rs6804842	*RARB*	3:25,081,441	G/A	0.57	0.019	0.003	0.59	0.59	0.58
rs2365389	*FHIT*	3:61,211,502	C/T	0.58	0.02	0.003	0.57	0.59	0.58
rs3849570	*GBE1*	3:81,874,802	A/C	0.36	0.019	0.003	0.37	0.38	0.35
rs13078960	*CADM2*	3:85,890,280	G/T	0.20	0.03	0.004	0.18	0.19	0.19

Variant	Nearest Gene	Chr: Position	Effect/Other Allele	Published EAF	Published Beta	Published SE	TwinGene EAF	PIVUS EAF	ULSAM EAF
rs16851483	*RASA2*	3:142,758,126	T/G	0.07	0.048	0.008	0.06	0.05	0.06
rs1516725	*ETV5*	3:187,306,698	C/T	0.87	0.045	0.005	0.89	0.89	0.88
rs10938397	*GNPDA2*	4:44,877,284	G/A	0.43	0.04	0.003	0.41	0.40	0.42
rs17001654	*SCARB2*	4:77,348,592	G/C	0.15	0.031	0.005	0.14	0.13	0.13
rs13107325	*SLC39A8*	4:103,407,732	T/C	0.07	0.048	0.007	0.04	0.04	0.04
rs11727676	*HHIP*	4:145,878,514	T/C	0.91	0.036	0.006	0.92	0.92	0.91
rs2112347	*POC5*	5:75,050,998	T/G	0.63	0.026	0.003	0.63	0.62	0.63
rs7715256	*GALNT10*	5:153,518,086	G/T	0.42	0.016	0.003	0.42	0.42	0.42
rs205262	*C6orf106*	6:34,671,142	G/A	0.27	0.022	0.004	0.27	0.26	0.26
rs2033529	*TDRG1*	6:40,456,631	G/A	0.29	0.019	0.003	0.29	0.29	0.30
rs2207139	*TFAP2B*	6:50,953,449	G/A	0.18	0.045	0.004	0.16	0.17	0.19
rs9400239	*FOXO3*	6:109,084,356	C/T	0.69	0.019	0.003	0.72	0.73	0.74
rs9374842	*LOC285762*	6:120,227,364	T/C	0.75	0.019	0.004	0.75	0.75	0.72
rs13201877	*IFNGR1*	6:137,717,234	G/A	0.14	0.023	0.005	0.15	0.15	0.17
rs13191362	*PARK2*	6:162,953,340	A/G	0.88	0.028	0.005	0.88	0.89	0.88
rs1167827	*HIP1*	7:75,001,105	G/A	0.55	0.02	0.003	0.57	0.54	0.58
rs2245368	*PMS2L11*	7:76,446,079	C/T	0.18	0.032	0.006	0.19	0.19	0.19
rs9641123	*CALCR*	7:93,035,668	C/G	0.43	0.019	0.004	0.40	0.40	0.41
rs6465468	*ASB4*	7:95,007,450	T/G	0.30	0.017	0.004	0.29	0.30	0.30
rs17405819	*HNF4G*	8:76,969,139	T/C	0.70	0.022	0.003	0.72	0.70	0.71
rs16907751	*ZBTB10*	8:81,538,012	C/T	0.91	0.035	0.007	0.90	0.90	0.89
rs2033732	*RALYL*	8:85,242,264	C/T	0.75	0.019	0.004	0.72	0.77	0.75
rs4740619	*C9orf93*	9:15,624,326	T/C	0.54	0.018	0.003	0.58	0.56	0.56
rs10968576	*LINGO2*	9:28,404,339	G/A	0.32	0.025	0.003	0.31	0.33	0.32
rs6477694	*EPB41L4B*	9:110,972,163	C/T	0.37	0.017	0.003	0.39	0.39	0.42
rs1928295	*TLR4*	9:119,418,304	T/C	0.55	0.019	0.003	0.53	0.53	0.52
rs10733682	*LMX1B*	9:128,500,735	A/G	0.48	0.017	0.003	0.48	0.50	0.51
rs7899106	*GRID1*	10:87,400,884	G/A	0.05	0.04	0.007	0.04	0.04	0.05
rs17094222	*HIF1AN*	10:102,385,430	C/T	0.21	0.025	0.004	0.19	0.20	0.21

Variant	Nearest Gene	Chr: Position	Effect/Other Allele	Published EAF	Published Beta	Published SE	TwinGene EAF	PIVUS EAF	ULSAM EAF
rs11191560	*NT5C2*	10:104,859,028	C/T	0.09	0.031	0.005	0.10	0.10	0.11
rs7903146	*TCF7L2*	10:114,748,339	C/T	0.71	0.023	0.003	0.75	0.76	0.72
rs11030104	*BDNF*	11:27,641,093	A/G	0.79	0.041	0.004	0.80	0.82	0.80
rs2176598	*HSD17B12*	11:43,820,854	T/C	0.25	0.02	0.004	0.25	0.24	0.23
rs3817334	*MTCH2*	11:47,607,569	T/C	0.41	0.026	0.003	0.41	0.40	0.40
rs4256980	*TRIM66*	11:8,630,515	G/C	0.65	0.021	0.003	0.65	0.64	0.66
rs12286929	*CADM1*	11:114,527,614	G/A	0.52	0.022	0.003	0.53	0.53	0.54
rs7138803	*BCDIN3D*	12:48,533,735	A/G	0.38	0.032	0.003	0.42	0.42	0.42
rs11057405	*CLIP1*	12:121,347,850	G/A	0.90	0.031	0.006	0.90	0.90	0.89
rs12016871	*MTIF3*	13:26,915,782	T/C	0.20	0.03	0.005	0.22	0.21	0.21
rs12429545	*OLFM4*	13:53,000,207	A/G	0.13	0.033	0.005	0.14	0.14	0.13
rs9540493	*MIR548X2*	13:65,103,705	A/G	0.46	0.017	0.004	0.46	0.43	0.45
rs1441264	*MIR548A2*	13:78,478,920	A/G	0.61	0.018	0.003	0.61	0.60	0.62
rs10132280	*STXBP6*	14:24,998,019	C/A	0.68	0.023	0.003	0.69	0.70	0.71
rs12885454	*PRKD1*	14:28,806,589	C/A	0.64	0.021	0.003	0.63	0.63	0.65
rs11847697	*PRKD1*	14:29,584,863	T/C	0.04	0.049	0.008	0.04	0.03	0.04
rs7141420	*NRXN3*	14:78,969,207	T/C	0.53	0.024	0.003	0.52	0.52	0.52
rs3736485	*DMXL2*	15:49,535,902	A/G	0.45	0.018	0.003	0.42	0.41	0.42
rs16951275	*MAP2K5*	15:65,864,222	T/C	0.78	0.031	0.004	0.70	0.78	0.79
rs7164727	*LOC100287559*	15:70,881,044	T/C	0.69	0.018	0.003	0.71	0.69	0.70
rs758747	*NLRC3*	16:3,567,359	T/C	0.24	0.022	0.004	0.24	0.24	0.22
rs12446632	*GPRC5B*	16:19,842,890	G/A	0.86	0.036	0.005	0.88	0.87	0.87
rs2650492	*SBK1*	16:28,240,912	A/G	0.30	0.021	0.004	0.28	0.32	0.33
rs3888190	*ATP2A1*	16:28,796,987	A/C	0.40	0.031	0.003	0.41	0.42	0.44
rs4787491	*INO80E*	16:29,922,838	G/A	0.51	0.016	0.003	0.51	0.48	0.51
rs9925964	*KAT8*	16:31,037,396	A/G	0.62	0.019	0.003	0.61	0.58	0.60
rs2080454	*CBLN1*	16:47,620,091	C/A	0.41	0.017	0.003	0.42	0.43	0.44
rs1558902	*FTO*	16:52,361,075	A/T	0.42	0.082	0.003	0.42	0.41	0.40
rs9914578	*SMG6*	17:1,951,886	G/C	0.21	0.02	0.004	0.22	0.23	0.22

Variant	Nearest Gene	Chr: Position	Effect/Other Allele	Published EAF	Published Beta	Published SE	TwinGene EAF	PIVUS EAF	ULSAM EAF
rs1000940	*RABEP1*	17:5,223,976	G/A	0.32	0.019	0.003	0.31	0.31	0.31
rs12940622	*RPTOR*	17:76,230,166	G/A	0.57	0.018	0.003	0.55	0.56	0.57
rs1808579	*C18orf8*	18:19,358,886	C/T	0.53	0.017	0.003	0.51	0.50	0.51
rs7239883	*LOC284260*	18:38,401,669	G/A	0.39	0.016	0.003	0.39	0.38	0.40
rs7243357	*GRP*	18:55,034,299	T/G	0.81	0.022	0.004	0.84	0.81	0.82
rs6567160	*MC4R*	18:55,980,115	C/T	0.24	0.048	0.004	0.25	0.24	0.25
rs17724992	*PGPEP1*	19:18,315,825	A/G	0.75	0.019	0.004	0.76	0.75	0.75
rs29941	*KCTD15*	19:39,001,372	G/A	0.67	0.018	0.003	0.67	0.66	0.68
rs2075650	*TOMM40*	19:50,087,459	A/G	0.85	0.026	0.005	0.85	0.84	0.83
rs2287019	*QPCTL*	19:50,894,012	C/T	0.80	0.036	0.004	0.78	0.79	0.78
rs3810291	*ZC3H4*	19:52,260,843	A/G	0.67	0.028	0.004	0.69	0.66	0.67
rs6091540	*ZFP64*	20:50,521,269	C/T	0.72	0.019	0.004	0.73	0.73	0.72
rs2836754	*ETS2*	21:39,213,610	C/T	0.61	0.016	0.003	0.64	0.64	0.64

In the TwinGene dataset, rs12016871 was merged into rs9581854 on January 27, 2015 (Build 36) https://www.ncbi.nlm.nih.gov/snp/rs12016871.

Chr: chromosome number; EAF: effect allele frequency; SE: standard error. Published Beta, Published SE and Published EAF were extracted from Locke *et al.* Nature 2015[1].

Table S2 List of 47 WHR-associated genetic variants used for Mendelian randomization and their allele frequencies in females from the TwinGene and PIVUS cohorts

Variant	Chr: Position	Nearest Gene	Effect Allele	Other Allele	Published EAF	Published Beta Females	P-values	TwinGene EAF	PIVUS EAF
rs905938	1:154991389	DCST2	T	C	0.75	0.034	4.9×10^{-10}	0.72	0.74
rs10919388	1:170372503	GORAB	C	A	0.72	0.033	4.8×10^{-10}	0.73	0.73
rs1569135	2:188115398	CALCRL	A	G	0.53	0.023	6.9×10^{-7}	0.52	0.49
rs10804591	3:129334233	PLXND1	A	C	0.79	0.04	6.1×10^{-13}	0.8	0.80
rs17451107	3:156797609	LEKR1	T	C	0.61	0.023	1.0×10^{-6}	0.61	0.62
rs3805389	4:56482750	NMU	A	G	0.28	0.027	4.6×10^{-8}	0.27	0.26
rs9991328	4:89713121	FAM13A	T	C	0.49	0.028	3.4×10^{-10}	0.47	0.48
rs303084	4:124066948	SPATA5-FGF2	A	G	0.80	0.029	3.4×10^{-7}	0.79	0.80
rs9687846	5:55861894	MAP3K1	A	G	0.19	0.041	3.8×10^{-12}	0.14	0.13
rs7759742	6:32381736	BTNL2	A	T	0.51	0.024	1.7×10^{-7}	0.49	0.48
rs1776897	6:34195011	HMGA1	G	T	0.08	0.052	6.8×10^{-9}	0.08	0.08
rs7801581	7:27223771	HOXA11	T	C	0.24	0.025	7.7×10^{-6}	0.24	0.25
rs7830933	8:23603324	NKX2-6	A	G	0.77	0.037	1.2×10^{-12}	0.78	0.78
rs12679556	8:72514228	MSC	G	T	0.25	0.033	2.1×10^{-10}	0.23	0.24
rs10991437	9:107735920	ABCA1	A	C	0.11	0.04	2.8×10^{-8}	0.12	0.12
rs7917772	10:104487443	SFXN2	A	G	0.62	0.027	5.5×10^{-9}	0.58	0.60
rs11231693	11:63862612	MACROD1-VEGFB	A	G	0.06	0.068	2.7×10^{-11}	0.07	0.07
rs4765219	12:124440110	CCDC92	C	A	0.67	0.037	1.0×10^{-14}	0.67	0.64
rs8042543	15:31708263	KLF13	C	T	0.79	0.023	6.7×10^{-5}	0.79	0.78
rs8030605	15:56504598	RFX7	A	G	0.15	0.031	1.0×10^{-5}	0.11	0.11
rs1440372	15:67033151	SMAD6	C	T	0.71	0.022	1.1×10^{-5}	0.7	0.71

Variant	Chr: Position	Nearest Gene	Effect Allele	Other Allele	Published EAF	Published Beta Females	P-values	TwinGene EAF	PIVUS EAF
rs2925979	16:81534790	*CMIP*	T	C	0.31	0.032	3.4×10^{-11}	0.31	0.32
rs4646404	17:17420199	*PEMT*	G	A	0.66	0.034	5.3×10^{-11}	0.63	0.64
rs8066985	17:68453345	*KCNJ2*	A	G	0.51	0.026	4.0×10^{-9}	0.55	0.56
rs12454712	18:60845884	*BCL2*	T	C	0.61	0.035	1.1×10^{-9}	0.57	0.56
rs4081724	19:33824946	*CEBPA*	G	A	0.85	0.033	9.2×10^{-7}	0.87	0.87
rs979012	20:6623374	*BMP2*	T	C	0.35	0.026	1.0×10^{-7}	0.35	0.35
rs6090583	20:45558831	*EYA2*	A	G	0.48	0.029	2.8×10^{-10}	0.46	0.45
rs2645294	1:119574587	*TBX15-WARS2*	T	C	0.58	0.035	1.5×10^{-14}	0.6	0.64
rs714515	1:172352990	*DNM3-PIGC*	G	A	0.43	0.029	1.8×10^{-10}	0.45	0.46
rs2820443	1:219753509	*LYPLAL1*	T	C	0.72	0.062	5.7×10^{-35}	0.71	0.70
rs10195252	2:165513091	*GRB14-COBLL1*	T	C	0.59	0.052	4.7×10^{-30}	0.56	0.57
rs17819328	3:12489342	*PPARG*	G	T	0.43	0.035	4.6×10^{-14}	0.44	0.45
rs2276824	3:52637486	*PBRM1*	C	G	0.43	0.028	3.7×10^{-9}	0.44	0.44
rs2371767	3:64718258	*ADAMTS9*	G	C	0.72	0.056	1.2×10^{-26}	0.74	0.74
rs1045241	5:118729286	*TNFAIP8-HSD17B4*	C	T	0.71	0.035	6.6×10^{-12}	0.72	0.72
rs7705502	5:173320815	*CPEB4*	A	G	0.32	0.027	1.9×10^{-8}	0.35	0.34
rs1294410	6:6738752	*LY86*	C	T	0.63	0.037	1.6×10^{-15}	0.65	0.64
rs1358980	6:43764551	*VEGFA*	T	C	0.47	0.06	3.7×10^{-34}	0.45	0.42
rs1936805	6:127452116	*RSPO3*	T	C	0.51	0.052	3.7×10^{-30}	0.53	0.53
rs1534696	7:26397239	*SNX10*	C	A	0.44	0.027	2.1×10^{-8}	0.4	0.40
rs10245353	7:25858614	*NFE2L3*	A	C	0.20	0.041	7.9×10^{-13}	0.17	0.18
rs10842707	12:26471364	*ITPR2-SSPN*	T	C	0.23	0.041	6.1×10^{-15}	0.25	0.25
rs1443512	12:54342684	*HOXC13*	A	C	0.24	0.04	1.1×10^{-14}	0.23	0.24
rs2294239	22:29449477	*ZNRF3*	A	G	0.59	0.028	6.9×10^{-10}	0.58	0.59

Chr: chromosome number; EAF: effect allele frequency. Published Beta and Published EAF estimates were extracted from Shungin et al. Nature 2015[2].

Table S3 List of 22 WHR-associated genetic variants used for Mendelian randomization and their allele frequencies in males from the TwinGene, PIVUS and ULSAM cohorts

Variant	Chr: Position	Nearest Gene	Effect Allele	Other Allele	Published EAF	Published Beta Males	P-values	TwinGene EAF	PIVUS EAF	ULSAM EAF
rs1385167	2:66200648	MEISI	G	A	0.14	0.036	2.3×10^{-7}	0.14	0.16	0.15
rs1569135	2:188115398	CALCRL	A	G	0.53	0.019	1.5×10^{-4}	0.52	0.49	0.52
rs17451107	3:156797609	LEKR1	T	C	0.62	0.03	1.4×10^{-8}	0.61	0.62	0.63
rs6556301	5:176527577	FGFR4	T	G	0.36	0.029	1.0×10^{-6}	0.36	0.35	0.36
rs7759742	6:32381736	BTNL2	A	T	0.5	0.023	5.5×10^{-6}	0.49	0.48	0.49
rs7801581	7:27223771	HOXA11	T	C	0.24	0.029	2.4×10^{-6}	0.24	0.25	0.26
rs8042543	15:31708263	KLF13	C	T	0.79	0.03	1.0×10^{-6}	0.79	0.78	0.77
rs8030605	15:56504598	RFX7	A	G	0.15	0.031	5.9×10^{-5}	0.11	0.11	0.12
rs1440372	15:67033151	SMAD6	C	T	0.70	0.027	1.4×10^{-6}	0.7	0.71	0.71
rs12608504	19:18389135	JUND	A	G	0.35	0.028	1.1×10^{-7}	0.33	0.32	0.35
rs4081724	19:33824946	CEBPA	G	A	0.86	0.039	1.4×10^{-7}	0.87	0.87	0.87
rs979012	20:6623374	BMP2	T	C	0.34	0.028	6.6×10^{-8}	0.35	0.35	0.34
rs224333	20:34023962	GDF5	G	A	0.63	0.036	9.0×10^{-12}	0.64	0.61	0.61
rs2645294	1:119574587	TBX15-WARS2	T	C	0.58	0.027	1.5×10^{-7}	0.6	0.64	0.61
rs714515	1:172352990	DNM3-PIGC	G	A	0.43	0.025	8.5×10^{-7}	0.45	0.46	0.44
rs2276824	3:52637486	PBRM1	C	G	0.43	0.02	1.4×10^{-4}	0.44	0.44	0.45
rs7705502	5:173320815	CPEB4	A	G	0.33	0.027	2.3×10^{-7}	0.35	0.34	0.34
rs1294410	6:6738752	LY86	C	T	0.63	0.025	1.4×10^{-6}	0.65	0.64	0.65
rs1936805	6:127452116	RSPO3	T	C	0.51	0.031	3.1×10^{-10}	0.53	0.53	0.54
rs10245353	7:25858614	NFE2L3	A	C	0.20	0.027	1.4×10^{-5}	0.17	0.18	0.18
rs10842707	12:26471364	ITPR2-SSPN	T	C	0.23	0.022	1.4×10^{-4}	0.25	0.25	0.25
rs2294239	22:29449477	ZNRF3	A	G	0.59	0.024	2.3×10^{-6}	0.58	0.59	0.61

Chr: chromosome number; EAF: effect allele frequency. Published Beta and Published EAF estimates were extracted from Shungin et al. Nature 2015[2].

Table S4. BMI-associated metabolites in meta-analysis of the ULSAM (N=1,035) and PIVUS (N=970) cohorts.

Metabolite	Beta (95% CI)	P-value
2-Methylbutyroylcarnitine	0.04 (0.03, 0.05)	1.9E-15
L-Acetylcarnitine	0.02 (0.01, 0.03)	7.4E-05
L-Carnitine	0.02 (0.01, 0.03)	2.3E-03
Piperine	0.06 (0.05, 0.07)	3.0E-26
1, 7 Dimethyluric acid	0.02 (0.01, 0.04)	1.6E-03
Creatine	0.02 (0.01, 0.03)	5.6E-03
DL-2-Aminooctanoic acid	-0.03 (-0.05, -0.02)	7.09E-06
Hippuric acid	-0.02 (-0.03, -0.01)	5.8E-03
L-Leucine	0.04 (0.02, 0.05)	5.6E-08
L-Tyrosine	0.03 (0.01, 0.04)	7.0E-04
Ornithine	0.02 (0.01, 0.04)	6.0E-04
Chenodeoxycholic acid	0.03 (0.02, 0.05)	6.5E-10
Chenodeoxycholic acid glycine conjugate	0.01 (0.00, 0.03)	3.8E-02
Deoxycholic acid	0.04 (0.02, 0.05)	2.0E-10
Deoxycholic acid glycine conjugate	0.04 (0.03, 0.05)	9.1E-10
Glycocholic acid	0.01 (0.00, 0.03)	0.0096
Hyodeoxycholic acid	0.03 (0.02, 0.04)	1.7E-06
4E,15Z-Bilirubin IX[a]	-0.02 (-0.03, -0.01)	1.6E-03
Bilirubin	-0.02 (-0.03, -0.01)	1.2E-04
Biliverdin a	-0.02 (-0.03, -0.01)	5.4E-04
I-Urobilin	0.03 (0.01, 0.04)	5.1E-04
$C_{12}H_{14}O_5$	-0.02 (-0.04, -0.01)	1.3E-05
Flavone	0.02 (0.01, 0.03)	1.6E-03
1-Linoleoyl-2-stearoyl-sn-glycerol	0.02 (0.00, 0.04)	0.012
Diacylglycerol(34:1)	0.03 (0.01, 0.06)	8.7E-03
Monoacylglycerol(14:0)	0.03 (0.02, 0.04)	1.5E-07
Monoacylglycerol(16:0)	0.03 (0.02, 0.04)	1.2E-07
Monoacylglycerol(16:1)	0.05 (0.04, 0.06)	1.8E-17
Monoacylglycerol(18:0)	0.01 (0.00, 0.03)	8.8E-03
Monoacylglycerol(18:1)	0.04 (0.03, 0.05)	2.8E-11
Monoacylglycerol(18:2)	0.05 (0.04, 0.06)	4.7E-17
Lysophosphatidylethanolamine(18:1)c	-0.04 (-0.05, -0.03)	1.0E-11
Lysophosphatidylethanolamine(18:1-P)	-0.04 (-0.06, -0.02)	1.1E-03
Lysophosphatidylethanolamine(18:2)	-0.05 (-0.06, -0.04)	6.4E-18
Lysophosphatidylethanolamine(20:4)	-0.02 (-0.03, -0.01)	7.4E-05
Phosphoethanolamine(38:2)	-0.04 (-0.06, -0.01)	6.6E-03
Lyso-PAF C-18	-0.05 (-0.08, -0.02)	0.0027
Lysophosphatidylcholine(0:0/18:2)	-0.05 (-0.08, -0.03)	1.0E-05
Lysophosphatidylcholine(17:0)	-0.07 (-0.08, -0.05)	6.0E-22
Lysophosphatidylcholine(18:1)[a]	-0.05 (-0.08, -0.03)	1.99E-05
Lysophosphatidylcholine(18:1)[b]	-0.05 (-0.07, -0.03)	3.7E-06
Lysophosphatidylcholine(18:2/0:0)	-0.05 (-0.08, -0.03)	3.4E-07
Lysophosphatidylcholine(18:3)	-0.04 (-0.06, -0.02)	8.1E-04

Metabolite	Beta (95% CI)	P-value
Lysophosphatidylcholine(20:0)	-0.07 (-0.08, -0.06)	5.7E-38
Lysophosphatidylcholine(20:1)	-0.05 (-0.08, -0.01)	5.3E-03
Lysophosphatidylcholine(20:2)	-0.03 (-0.05, -0.01)	9.81E-05
Lysophosphatidylcholine(20:3)[a]	0.02 (0.01, 0.03)	2.5E-04
Lysophosphatidylcholine(20:3)[b]	0.02 (0.01, 0.03)	1.3E-03
Lysophosphatidylcholine(22:5)[a]	-0.02 (-0.03, -0.01)	0.0059
Lysophosphatidylcholine(22:5)[b]	-0.02 (-0.03, -0.01)	5.0E-03
Lysophosphatidylcholine(P-16:0)	-0.05 (-0.07, -0.04)	3.2E-09
Phosphatidylcholine(16:2)	-0.04 (-0.06, -0.03)	4.9E-09
Phosphatidylcholine(18:1)	-0.04 (-0.06, -0.02)	5.8E-04
Phosphatidylcholine(32:0)	-0.02 (-0.03, -0.01)	7.9E-05
Phosphatidylcholine(32:1)	0.02 (0.01, 0.03)	3.0E-05
Phosphatidylcholine(36:2)	-0.02 (-0.03, -0.00)	9.4E-03
Phosphatidylcholine(36:5)	0.02 (0.01, 0.03)	3.6E-03
Phosphatidylcholine(38:3)	0.03 (0.01, 0.04)	2.8E-04
Phosphatidylcholine(38:4)	0.02 (0.00, 0.03)	4.3E-02
Phosphatidylcholine(38:6)	0.01 (0.00, 0.02)	4.7E-02
Phosphatidylcholine(40:5)	0.01 (0.00, 0.02)	3.5E-02
Phosphatidylcholine(42:7)	-0.02 (-0.04, -0.00)	2.1E-02
Caffeine	0.04 (0.03, 0.05)	1.2E-10
Paraxanthine; Theophylline	0.02 (0.01, 0.03)	1.9E-04
Uric acid	0.04 (0.02, 0.07)	6.0E-04
3-Indolepropionic acid	-0.02 (-0.03, -0.00)	1.4E-02
Indolelactic acid	0.01 (0.00, 0.03)	3.6E-02
Creatinine	0.01 (0.00, 0.02)	6.4E-03
Sum-Hexose	0.05 (0.04, 0.06)	1.6E-14
Propranolol	0.02 (0.01, 0.03)	1.4E-03
1-N-(tetradecanoyl)-1-b-lactosyl-sphing-4-enine	-0.03 (-0.05, -0.02)	8.2E-05
Dipeptide	0.02 (0.01, 0.04)	5.4E-03
Gamma-Glutamyl-leucine	0.04 (0.03, 0.05)	4.4E-13
Phosphoethanolamine(P-34:1)	-0.04 (-0.06, -0.02)	1.3E-04
Phosphoethanolamine(40:0)	-0.03 (-0.04, -0.01)	1.3E-04
Lactosyl ceramide(d18:1/16:0)	-0.04 (-0.06, -0.02)	3.9E-06
Sphingomyelin(32:1-OH)	-0.03 (-0.04, -0.02)	8.9E-08
Prostaglandin J2	-0.02 (-0.03, -0.01)	2.0E-03
L-proline-betaine	-0.02 (-0.03, -0.00)	4.8E-03
Dodecanedioic acid	-0.04 (-0.05, -0.02)	4.1E-10
Dodecanoic acid	-0.02 (-0.04, -0.01)	4.7E-04
Heptadecanoic acid	-0.02 (-0.03, -0.01)	7.1E-04
Palmitic acid	0.02 (0.00, 0.03)	5.0E-03
Sphingomyelin(28:1)	-0.03 (-0.05, -0.02)	1.3E-10
Sphingomyelin(34:1)	-0.03 (-0.05, -0.02)	3.6E-10
Sphingosine	0.04 (0.03, 0.05)	3.7E-08
Sphingomyelin(32:1)	-0.02 (-0.03, -0.01)	5.4E-04
Sphingomyelin(41:2)	-0.02 (-0.03, -0.01)	3.7E-04

Metabolite	Beta (95% CI)	P-value
Sphingomyelin(42:3)	-0.01 (-0.02, -0.00)	2.2E-02
Cholesterol	-0.01 (-0.02, -0.00)	2.4E-02
Corticosterone	0.02 (0.01, 0.04)	3.9E-03
Cortisol	-0.03 (-0.05, -0.02)	5.8E-10
Dehydroepiandrosterone sulfate (sodium salt)	-0.02 (-0.03, -0.01)	7.2E-05
10-Nitro-9E-octadecenoic acid	0.02 (0.00, 0.03)	6.1E-03
Alpha-Linolenic acid	0.02 (0.00, 0.03)	4.1E-03
Arachidonic acid	0.02 (0.01, 0.03)	3.8E-03
Arachidonic acid ethyl ester	0.03 (0.02, 0.04)	1.9E-08
Docosahexaenoic Acid	0.02 (0.01, 0.03)	1.4E-04
Docosapentaenoic acid	0.02 (0.01, 0.03)	1.8E-04
Eicosadienoic acid	-0.01 (-0.02, -0.00)	3.6E-02
Eicosatrienoic Acid	0.03 (0.02, 0.04)	1.1E-06
Eicosatrienoic Acid methyl ester	0.02 (0.00, 0.03)	2.7E-02
Linolenyl aldehyde	-0.03 (-0.05, -0.01)	9.5E-03
Palmitoleic acid	0.02 (0.00, 0.04)	3.6E-02
Vaccenic acid	-0.02 (-0.03, -0.00)	4.1E-02
Pantothenic acid	0.02 (0.00, 0.03)	8.8E-03
Vitamin D3 derivative I	0.03 (0.02, 0.04)	5.2E-09
Vitamin D3 derivative II	-0.01 (-0.03, -0.00)	5.0E-02
Alpha-Tocopherol	-0.02 (-0.03, -0.01)	1.1E-03

Analysis were adjusted for both age and sex in the PIVUS cohort but age only in the ULSAM.

Table S5. 77 metabolites associated with BMI in meta-analysis of the ULSAM (N=1,135) and PIVUS (N=970), and replicated in TwinGene (N= 2,059).

Metabolite	ULSAM and PIVUS		TwinGene	
	Beta (95% CI)	P-value	Beta (95% CI)	P-value
Acyl Carnitines				
L-Acetylcarnitine	0.02 (0.01, 0.03)	7.4E-05	0.01 (0.00, 0.02)	0.031
2-Methylbutyroylcarnitine	0.04 (0.03, 0.05)	1.9E-15	0.03 (0.02, 0.04)	2.7E-07
L-Carnitine	0.02 (0.01, 0.03)	0.0023	0.03 (0.02, 0.05)	6.0E-08
Alkaloids and Derivatives				
Piperine	0.06 (0.05, 0.07)	3.0E-26	0.04 (0.03, 0.05)	2.4E-08
1, 7 Dimethyluric acid	0.02 (0.01, 0.04)	0.0016	0.02 (0.00, 0.03)	0.014
Amino Acids and Derivatives				
L-Leucine	0.04 (0.02, 0.05)	5.6E-08	0.05 (0.04, 0.07)	3.1E-14
Ornithine	0.02 (0.01, 0.04)	6.0E-04	0.04 (0.02, 0.05)	1.6E-08
Creatine	0.02 (0.01, 0.03)	0.0056	0.02 (0.01, 0.03)	0.0059
DL-2-Aminooctanoic acid	-0.03 (-0.05, -0.02)	7.1E-06	-0.03 (-0.04, -0.01)	8.0E-05
L-Tyrosine	0.03 (0.01, 0.04)	7.0E-04	0.04 (0.02, 0.05)	9.0E-09
Hippuric acid	-0.02 (-0.03, -0.01)	0.0058	-0.02 (-0.03, -0.01)	0.0027
Bile Acids				
Deoxycholic acid glycine conjugate	0.04 (0.03, 0.05)	9.1E-10	0.03 (0.02, 0.05)	1.4E-08
Deoxycholic acid	0.04 (0.02, 0.05)	2.0E-10	0.04 (0.03, 0.06)	8.9E-14
Hyodeoxycholic acid	0.03 (0.02, 0.04)	1.7E-06	0.04 (0.03, 0.05)	2.6E-09
Chenodeoxycholic acid	0.03 (0.02, 0.05)	6.5E-10	0.05 (0.03, 0.06)	1.5E-13
Glycocholic acid	0.01 (0.00, 0.03)	0.0096	0.02 (0.01, 0.03)	0.0011
Bilirubins				
4E,15Z-Bilirubin IX[a]	-0.02 (-0.03, -0.01)	0.0016	-0.02 (-0.04, -0.01)	0.0043
I-Urobilin	0.03 (0.01, 0.04)	5.1E-04	0.03 (0.02, 0.04)	3.1E-06
Biliverdin a	-0.02 (-0.03, -0.01)	5.4E-04	-0.03 (-0.04, -0.01)	1.5E-04
Bilirubin	-0.02 (-0.03, -0.01)	1.2E-04	-0.02 (-0.04, -0.01)	0.0045

Metabolite	ULSAM and PIVUS		TwinGene	
	Beta (95% CI)	P-value	Beta (95% CI)	P-value
Cinnamic Acid and Derivatives				
$C_{12}H_{14}O_5$	-0.02 (-0.04, -0.01)	1.3E-05	-0.03 (-0.04, -0.01)	0.00014
Glycerolipids				
1-Linoleoyl-2-stearoyl-sn-glycerol	0.02 (0.00, 0.04)	0.012	0.03 (0.02, 0.04)	4.5E-09
Monoacylglycerol(18:0)	0.01 (0.00, 0.03)	0.0088	0.06 (0.05, 0.07)	1.4E-21
Monoacylglycerol(16:0)	0.03 (0.02, 0.04)	1.2E-07	0.07 (0.06, 0.08)	2.1E-27
Monoacylglycerol(18:1)	0.04 (0.03, 0.05)	2.8E-11	0.06 (0.05, 0.07)	1.8E-19
Diacylglycerol(34:1)	0.03 (0.01, 0.06)	0.0087	0.05 (0.04, 0.06)	3.1E-15
Monoacylglycerol(14:0)	0.03 (0.02, 0.04)	1.5E-07	0.06 (0.05, 0.07)	1.8E-23
Monoacylglycerol(18:2)	0.05 (0.04, 0.06)	4.7E-17	0.05 (0.04, 0.06)	1.5E-15
Monoacylglycerol(16:1)	0.05 (0.04, 0.06)	1.8E-17	0.06 (0.05, 0.07)	3.1E-20
Glcerophosphoethanolamines				
Lysophosphatidylethanolamine(18:2)	-0.05 (-0.06, -0.04)	6.4E-18	-0.02 (-0.04, -0.01)	1.3E-04
Phosphoethanolamine(38:2)	-0.04 (-0.06, -0.01)	0.0066	-0.03 (-0.04, -0.02)	4.3E-07
Lysophosphatidylethanolamine(18:1-P)	-0.04 (-0.06, -0.02)	0.0011	-0.03 (-0.04, -0.01)	1.9E-05
Imidazopyrimidines				
Uric acid	0.04 (0.02, 0.07)	0.0006	0.06 (0.05, 0.08)	4.0E-17
Indoles				
3-Indolepropionic acid	-0.02 (-0.03, -0.00)	0.014	-0.03 (-0.05, -0.02)	5.0E-07
Monosaccharides				
Sum-Hexose	0.05 (0.04, 0.06)	1.6E-14	0.05 (0.04, 0.07)	1.1E-13
Peptides				
Gamma-Glutamyl-leucine	0.04 (0.03, 0.05)	4.4E-13	0.05 (0.04, 0.06)	1.4E-12
Phosphoethanolamine				
Phosphoethanolamine(P-34:1)	-0.04 (-0.06, -0.02)	1.3E-04	-0.03 (-0.04, -0.02)	1.8E-06
Phosphoethanolamine(40:0)	-0.03 (-0.04, -0.01)	1.3E-04	-0.02 (-0.04, -0.01)	0.0031
Phosphosphingolipids				
Sphingomyelin(32:1-OH)	-0.03 (-0.04, -0.02)	8.9E-08	-0.02 (-0.03, -0.01)	4.3E-04

Metabolite	ULSAM and PIVUS		TwinGene	
	Beta (95% CI)	P-value	Beta (95% CI)	P-value
Prostaglandins				
Prostaglandin J2	-0.02 (-0.03, -0.01)	0.002	-0.02 (-0.04, -0.01)	7.1E-04
Pyrrolidines				
L-Proline-betaine	-0.02 (-0.03, -0.00)	0.0048	-0.02 (-0.03, -0.00)	0.015
Saturated Fatty Acids				
Dodecanedioic acid	-0.04 (-0.05, -0.02)	4.1E-10	-0.02 (-0.04, -0.01)	4.3E-06
Palmitic acid	0.02 (0.00, 0.03)	0.005	0.03 (0.01, 0.04)	4.6E-05
Spingolipids				
Sphingosine	0.04 (0.03, 0.05)	3.7E-08	0.05 (0.04, 0.07)	1.5E-15
Spingomyelins				
Sphingomyelin(34:1)	-0.03 (-0.05, -0.02)	3.6E-10	-0.02 (-0.03, -0.01)	0.0027
Sphingomyelin(28:1)	-0.03 (-0.05, -0.02)	1.3E-10	-0.02 (-0.03, -0.01)	9.2E-04
Steroid and Steroid Derivatives				
Dehydroepiandrosterone sulfate (sodium salt)	-0.02 (-0.03, -0.01)	7.2E-05	-0.04 (-0.05, -0.02)	4.0E-09
Cortisol	-0.03 (-0.05, -0.02)	5.8E-10	-0.02 (-0.04, -0.01)	0.0079
Corticosterone	0.02 (0.01, 0.04)	0.0039	0.03 (0.02, 0.04)	8.3E-06
Unsaturated Fatty Acids				
Eicosatrienoic Acid	0.03 (0.02, 0.04)	1.1E-06	0.05 (0.04, 0.06)	7.6E-15
10-Nitro-9E-octadecenoic acid	0.02 (0.00, 0.03)	0.006	0.01 (0.00, 0.02)	0.035
Docosapentaenoic acid	0.02 (0.01, 0.03)	1.8E-04	0.03 (0.01, 0.04)	2.0E-04
Arachidonic acid ethyl ester	0.03 (0.02, 0.04)	1.9E-08	0.04 (0.03, 0.05)	1.3E-09
Arachidonic acid	0.02 (0.01, 0.03)	0.0038	0.02 (0.01, 0.04)	1.9E-04
Alpha-Linolenic acid	0.02 (0.00, 0.03)	0.004	0.03 (0.02, 0.04)	2.6E-06
Linolenyl aldehyde	-0.03 (-0.05, -0.01)	0.01	-0.03 (-0.04, -0.02)	5.7E-06
Vitamin B				
Pantothenic acid	0.02 (0.00, 0.03)	0.0088	0.02 (0.00, 0.03)	6.8E-03
Vitamin D3 Derivatives				
Vitamin D3 derivative I	0.03 (0.02, 0.04)	5.2E-09	0.04 (0.02, 0.05)	1.1E-07

Metabolite	ULSAM and PIVUS		TwinGene	
	Beta (95% CI)	P-value	Beta (95% CI)	P-value
Glycerophospholipids				
Lyso-PAF C-18	-0.05 (-0.08, -0.02)	0.0027	-0.04 (-0.05, -0.03)	8.0E-10
Lysophosphatidylcholine(20:1)	-0.05 (-0.08, -0.01)	0.0053	-0.04 (-0.05, -0.02)	3.0E-06
Lysophosphatidylcholine(20:2)	-0.03 (-0.05, -0.01)	9.8E-05	-0.02 (-0.03, -0.01)	7.5E-04
Lysophosphatidylcholine(22:5)[a]	-0.02 (-0.03, -0.01)	0.0059	-0.02 (-0.03, -0.01)	0.0022
Lysophosphatidylcholine(P-16:0)	-0.05 (-0.07, -0.04)	3.2E-09	-0.04 (-0.05, -0.03)	3.6E-11
Lysophosphatidylcholine(18:2/0:0)	-0.05 (-0.08, -0.03)	3.4E-07	-0.06 (-0.07, -0.05)	2.6E-17
Phosphatidylcholine(18:1)	-0.04 (-0.06, -0.02)	5.8E-04	-0.02 (-0.03, -0.00)	9.5E-03
Lysophosphatidylcholine(20:3)[b]	0.02 (0.01, 0.03)	0.0013	0.04 (0.03, 0.05)	1.2E-09
Lysophosphatidylcholine(18:1)[b]	-0.05 (-0.07, -0.03)	3.7E-06	-0.05 (-0.06, -0.03)	1.2E-11
Lysophosphatidylcholine(18:3)	-0.04 (-0.06, -0.02)	8.1E-04	-0.03 (-0.05, -0.02)	2.9E-08
Lysophosphatidylcholine(20:0)	-0.07 (-0.08, -0.06)	5.7E-38	-0.06 (-0.08, -0.05)	7.9E-17
Lysophosphatidylcholine(18:1)[a]	-0.05 (-0.08, -0.03)	2.0E-05	-0.04 (-0.06, -0.03)	1.2E-10
Lysophosphatidylcholine(17:0)	-0.07 (-0.08, -0.05)	6.0E-22	-0.03 (-0.04, -0.02)	2.5E-05
Lysophosphatidylcholine(20:3)[a]	0.02 (0.01, 0.03)	2.5E-04	0.04 (0.03, 0.05)	1.0E-09
Lysophosphatidylcholine(22:5)[b]	-0.02 (-0.03, -0.01)	0.005	-0.02 (-0.03, -0.01)	0.0012
Phosphatidylcholine(32:1)	0.02 (0.01, 0.03)	3.0E-05	0.02 (0.01, 0.03)	0.0056
Phosphatidylcholine(36:2)	-0.02 (-0.03, -0.00)	0.009	-0.01 (-0.02, -0.00)	0.038
Phosphatidylcholine(38:3)	0.03 (0.01, 0.04)	2.8E-04	0.03 (0.02, 0.04)	3.6E-06
Lysophosphatidylcholine(0:0/18:2)	-0.05 (-0.08, -0.03)	1.0E-05	-0.06 (-0.07, -0.05)	1.0E-16

Regression models were adjusted for age and sex. In the ULSAM cohort, the regression models were only adjusted for age as the cohort contain only male participants. [a] or [b] indicates that two distinct peaks of the same metabolite (isomers) were detected.

Table S6. Meta-analysis of WHRadjBMI-metabolites estimates across the ULSAM (N=1,112) and PIVUS (N=478) cohorts among males.

Metabolite	Beta (95% CI)	P-value
L-Acetylcarnitine	1.52 (0.44, 2.60)	5.7E-03
L-Carnitine	1.34 (0.24, 2.44)	1.7E-02
Piperine	1.96 (0.91, 3.02)	2.6E-04
Creatine	3.06 (1.44, 4.67)	2.0E-04
L-Aspartyl-L-phenylalanine	2.03 (0.29, 3.78)	2.3E-02
Chenodeoxycholic acid glycine conjugate	1.62 (0.14, 3.10)	3.2E-02
Cholic acid	-1.39 (-2.73, -0.05)	4.2E-02
Deoxycholic acid glycine conjugate	1.59 (0.53, 2.64)	3.1E-03
Glycocholic acid	1.90 (0.42, 3.38)	1.2E-02
$C_{12}H_{14}O_5$	-2.09 (-3.17, -1.01)	1.5E-04
Monoacylglycerol(16:1)	1.37 (0.28, 2.46)	1.3E-02
Lysophosphatidylcholine(0:0/16:1)	1.15 (0.07, 2.24)	3.8E-02
Lysophosphatidylcholine(17:0)	-1.98 (-3.73, -0.23)	2.7E-02
Lysophosphatidylcholine(18:1)a	-1.20 (-2.36, -0.04)	4.3E-02
Lysophosphatidylcholine(18:2/0:0)	-2.27 (-4.33, -0.21)	3.0E-02
Lysophosphatidylcholine(20:0)	-1.12 (-2.18, -0.07)	3.7E-02
Lysophosphatidylcholine(20:2)	-2.01 (-3.09, -0.93)	2.7E-04
Phosphatidylcholine(32:1)	1.53 (0.20, 2.87)	2.5E-02
Theobromine	2.84 (0.94, 4.73)	3.3E-03
Propranolol	1.23 (0.14, 2.31)	2.6E-02
Ceramide phosphoethanolamine(38:2)	-1.11 (-2.15, -0.07)	3.7E-02
Lactosyl ceramide(d18:1/16:0)	-1.53 (-2.85, -0.20)	2.4E-02
Sphingosine	1.94 (0.13, 3.76)	3.6E-02
Sphingomyelin(32:2)	-1.03 (-2.06, -0.00)	4.9E-02
Sphingomyelin(36:3)	-1.41 (-2.45, -0.36)	8.3E-03
Sphingomyelin(40:2)	-1.23 (-2.30, -0.15)	2.6E-02
3a,6b,7b-Trihydroxy-5b-cholanoic acid	-1.69 (-2.80, -0.58)	2.8E-03
Corticosterone	1.20 (0.15, 2.25)	2.6E-02
Dehydroepiandrosterone sulfate (sodium salt)	-1.92 (-3.00, -0.85)	4.3E-04
Palmitoleic acid	1.39 (0.32, 2.46)	1.1E-02
Pantothenic acid	2.17 (0.39, 3.95)	1.7E-02

Waist-to-hip ratio (WHR) analysis were adjusted for age and BMI.

Table S7. 11 WHRadjBMI-associated metabolites in females in the PIVUS (N=487) and TwinGene (N=879) cohorts, and 4 WHRadjBMI-associated metabolites in males in the PIVUS (N=483) and ULSAM (N=1,112) cohorts, and replicated in TwinGene (N=1,167).

	WHRadjBMI			
Metabolite	**Beta (95% CI)**	**P-value**	**Beta (95% CI)**	**P-value**
Female	**PIVUS**		**TwinGene**	
Amino Acids and Derivatives				
L-Proline	2.55 (0.89, 4.22)	0.003	1.20 (0.29, 2.11)	0.010
Bile Acids				
Hyodeoxycholic acid	3.69 (2.03, 5.34)	1.5E-05	0.72 (0.04, 1.39)	0.037
Glycerolipids				
Monoacylglycerol(20:5)	2.67 (0.95, 4.39)	0.0025	0.81 (0.04, 1.58)	0.038
Monoacylglycerol(18:2)	2.64 (0.84, 4.44)	0.004	1.53 (0.60, 2.46)	0.001
Monoacylglycerol(14:0)	3.18 (1.41, 4.94)	0.0005	1.64 (0.79, 2.50)	0.0002
Diacylglycerol(34:1)	3.57 (1.88, 5.26)	4.2E-05	0.91 (0.02, 1.81)	0.044
Monoacylglycerol(18:1)	3.84 (2.08, 5.59)	2.3E-05	1.75 (0.83, 2.68)	0.0003
Monoacylglycerol(16:1)	3.01 (1.24, 4.77)	0.0009	1.46 (0.52, 2.40)	0.002
Monoacylglycerol(16:0)	3.74 (1.96, 5.51)	4.4E-05	1.92 (1.01, 2.83)	4.0E-05
Peptides				
Gamma-Glutamyl-leucine	2.73 (1.25, 4.21)	0.0003	1.40 (0.67, 2.13)	0.0002
Sphingomyelins				
Sphingomyelin(32:2)	-2.46 (-3.78, -1.14)	0.0003	-0.71 (-1.30, -0.12)	0.019
Male	**PIVUS and ULSAM**		**TwinGene**	
Amino Acids and Derivatives				
Creatine	3.06 (1.44, 4.67)	0.0002	1.40 (0.60, 2.20)	0.0006
Cinnamic Acid and Derivates				
$C_{12}H_{14}O_5$	-2.09 (-3.17, -1.01)	0.0001	-0.89 (-1.73, -0.06)	0.036
Steroids and Steroid Derivatives				
Dehydroepiandrosterone sulfate (sodium salt)	-1.92 (-3.00, -0.85)	0.0004	-1.22 (-2.03, -0.40)	0.004

Waist-to-hip ratio (WHR) analysis were adjusted for age and BMI.

Table S9. Mendelian randomization analysis using the combined sample (N=3,610) from the ULSAM, PIVUS and TwinGene cohort, employing weighted median method and MR-Egger method.

Metabolite	Weighted Median Method		MR-Egger Method	
	Beta (95% CI)	P-value	Beta (95% CI)	P-value
BMI				
Dodecanedioic acid	-0.36 (-0.72, 0.01)	0.05	0.04 (-0.56, 0.65)	0.89
Lysophosphatidylcholine(P-16:0)	-0.26 (-0.64, 0.11)	0.17	0.02 (-0.60, 0.64)	0.95
Arachidonic acid	-0.49 (-0.83, -0.15)	0.005	-0.29 (-0.86, 0.27)	0.31
Creatine	0.095 (-0.32, 0.51)	0.65	0.25 (-0.39, 0.89)	0.44
WHRadjBMI (Females)				
Sphingomyelin(32:2)	-0.31 (-0.87, 0.25)	0.29	-0.32 (-1.63, 0.98)	0.62

Table S10. Mendelian randomization analysis using the weighted median method and MR-Egger method in the KORA/TwinsUK cohort (N=7,373), the CHARGE consortium (N= 8,631), the DIRECT consortium (N=3,029) and the FHS cohort (N=2,076).

KORA/TwinsUK Cohort

Metabolite	Weighted Median Method		MR-Egger Method	
	Beta (95% CI)	P-value	Beta (95% CI)	P-value
Dodecanedioic acid	-0.01 (-0.06, 0.04)	0.64	-0.01 (-0.09, 0.07)	0.79
Arachidonic acid	-0.02 (-0.05, 0.01)	0.19	-0.04 (-0.09, 0.02)	0.16
Creatine	0.04 (-0.01, 0.10)	0.11	0.07 (-0.01, 0.15)	0.10
CHARGE Consortium				
Arachidonic acid	-0.58 (-0.98, -0.18)	0.004	-0.73 (-1.40, -0.05)	0.034
DIRECT Consortium				
Arachidonic acid	0.04 (-0.15, 0.23)	0.70	0.07 (-0.24, 0.39)	0.66
Creatine	0.08 (-0.12, 0.27)	0.43	-0.17 (-0.49, 0.16)	0.31
FHS Cohort				
Creatine	-0.05 (-0.51, 0.41)	0.84	0.34 (-0.42, 1.11)	0.38

Table S11. Pearson correlation between BMI and metabolites with evidence of causal association in Mendelian randomization analysis in the Swedish cohorts.

	PIVUS (N=970)		ULSAM (N=1,138)		TwinGene (N=2,059)	
	r	P-value	r	P-value	r	P-value
Dodecanedioic acid	-0.13	<0.0001	-0.14	<0.0001	-0.12	<0.0001
Lysophosphatidylcholine(P-16:0)	-0.27	<0.0001	-0.15	<0.0001	-0.19	<0.0001
Arachidonic acid	0.08	0.011	0.05	0.11	0.11	<0.0001
*Creatine	0.08	0.013			0.07	0.001

*Creatine was not available in the ULSAM cohort.

r represents Pearson correlation coefficient.

Table S12. Metabolite class enrichment analysis for the BMI-associated metabolites, and WHRadjBMI among females and males, respectively.

	BMI	WHRadjBMI females	WHRadjBMI males
Class	p_{adj}*	p_{adj}*	p_{adj}*
AcylCarnitines	0.06	0.75	0.68
Amino Acids and Derivatives	0.11	0.83	0.68
Glycerophosphoethanolamines	0.64	0.33	0.48
Glycerophospholipids	0.30	0.33	0.68
Unsaturated Fatty Acids	0.61	0.86	0.48

p_{adj} represents the Benjamini-Hochberg-corrected P-value.

Table S13. Mean correlation coefficients across the randomized duplicate injections for the metabolites discovered through Mendelian randomization analysis in the ULSAM, PIVUS and TwinGene.

BMI-associated Metabolites

	PIVUS	TwinGene	ULSAM
Dodecanedioic acid	0.65	n.d.	n.d.
Lysophosphatidylcholine(P-16:0)	0.61	0.74	0.69
Arachidonic acid	0.74	0.54	0.63
Creatine	0.74	0.54	0.63

WHRadjBMI-associated Metabolites (Females)

	PIVUS	TwinGene	ULSAM
Sphingomyelin(32:2)	0.86	0.63	0.65

n.d. denotes no duplicates are available.

Supplemental text 1

Metabolite annotation: In total 7,522 metabolomic features were detected from PIVUS, 10,162 from ULSAM, and 9,755 from the TwinGene cohorts were identified. Common metabolomic features between ULSAM, PIVUS, and TwinGene cohorts were identified through matching retention time, mass-to-charge (m/z) ratio, and fragmentation patterns. For each metabolomic feature, retention time, m/z, and fragmentation pattern were compared to in-house standards, as well as with public database reference libraries, and were matched according to Metabolomics Standard Initiative guidelines[3]. All metabolomic features with a retention time <35s were excluded from the current analysis. We combined strongly correlated features with shared retention times for the construction of representative fragmentation spectra for annotation according to the 4-level metabolites accuracy classification suggested by Metabolomics Standards Initiative[3]. In this 4-level accuracy approach, level 1 indicates identification based upon matching by retention time, mass, and fragmentation patterns to in-house standards while level 4 represents unknown metabolites which are not matched in the public databases in relation to retention time, mass-to-charge ratio, and fragmentation patterns with named metabolites (level 2) or based upon chemical classes (level 3). Details about the annotations for the metabolites discovered through Mendelian randomization (MR) analysis for general as well as well as for central body fat distribution have been presented: dodecanedioic acid (Supplemental Figure 3), lysophosphatidylcholine P-16:0 (Supplemental Figure 4), arachidonic acid (Supplemental Figure 5), creatine (Supplemental Figure 6) and sphingomyelin (32:2) (Supplemental Figure 7).

The metabolite data transformation procedure is described in detail elsewhere[4]. In summary, the metabolites features were $\log_2$ transformed to approximate normal transformation. Potential sample outliers were identified and removed through plotting the total sample intensity of each sample, as, the samples show extreme intensities might be due to degradation or technical

errors. ANOVA-type normalization was used for taking into accounts the factors of unwanted consideration. This normalization procedure outperforms the other commonly normalization approaches. Normalization procedure was performed through regression each metabolite intensity feature against several factors of unwanted variability. Residuals from the regression were used as intensity features. In each of the Swedish cohorts (TwinGene, ULSAM and PIVUS), the intensity features were identified through using following technical variables,

TwinGene: retention time correction, analysis date, storage time, unknown cluster effect;

ULSAM: retention time correction, analysis date, sample collection, plate effect;

PIVUS: retention time correction, analysis date, storage time, season effect.

The intensity features were averaged between technical duplicates to rule the potential effect of inherent instrumental variability and features with poor correlation (if p-values threshold was >0.05) between duplicates were removed[4]. Finally, intensities were SD-transformed before association analysis.

For quality control (QC), prior to each batch of two 96-well plates of samples, instrument maintenance (cone cleaning, mass calibration, and detector gain calibration) was performed, and an external QC standard mix was injected containing $2\,\mu g$ mL-1 each of caffeine, terfenadine, sulfadimethoxime, and reserpine. The QC standards were evaluated for retention time ($+/-$ 0.05 min), signal intensity ($<25\%$ relative standard deviation), and mass accuracy (<3 ppm). All samples were randomized prior to instrumental analysis. Since internal standards were not available at the time of analysis, randomized duplicate injections were performed to mitigate potential within-sample variation originating from the instrumental analysis. Spearman correlation between features across duplicate injections was assessed and a correlation between technical duplicates was considered significant with Bonferroni P value <0.05. Average peak

areas of the duplicate injections were then used for the relative quantitation, features with poor correlation between the duplicate injections were excluded. The mean Spearman correlation between duplicate injections across samples was 0.43 in PIVUS, 0.38 in TwinGene, and 0.46 in ULSAM. The mean feature correlation for the top BMI-associated metabolic features are provided in Supplemental Information Table S13 and ranged from 0.48-0.87. Moreover, all three studies were comparable in terms of mean coefficient of variation across features; 2.9 % in PIVUS, 3.7% in TwinGene, and 5.2 % in ULSAM as has been previously described in Ganna *et al* 2014[5] and Fall *et al* 2016[6].

Supplemental text 2

We applied a random effect meta-analysis as it assumes that all the studies in a meta-analysis are estimating different yet unrelated true underlying effects, each effect representing a random sample from a particular distribution of effect sizes. In a random effect meta-analysis approach, the study weights are more similar compared to the fixed effect meta-analysis approach (small studies that are part of meta-analysis gain influence while the larger sample size studies lose influence) and summary effect confidence intervals are usually larger[7].

In the below system of equations, T_i represents the observed effect, θ_i represents the true effect and ε_i represents the deviation from the true effect due to sampling error. θ represents the average effect across studies, and u_i the individual study deviation from the average effect. Both ε_i and u_i are assumed to follow a normal distribution.

$$T_i = \theta_i + \varepsilon_i$$
$$\theta_i = \theta + u_i$$

Supplemental text 3

Details and equations about the Mendelian randomization approaches that we used in the current analysis have previously been described in detail[8-10].

Inverse-Variance Weighted (IVW) Method

The causal effect of an exposure on the outcome, which is the ratio of the gene variants outcome association to the genetic variants exposure association, can be estimated using *n* number of variants conditional on the variants being uncorrelated. Then the ratio can be estimated using the below formula which is called inverse variance weighted (IVW) estimator.

$$\hat{\beta}_{IVW} = \frac{\sum_j \hat{\gamma}_j^2 \sigma_{Yj}^{-2} \hat{\beta}_j}{\sum_j \hat{\gamma}_j^2 \sigma_{Yj}^{-2}}$$

$\hat{\gamma}_j$ denotes the estimated coefficient for the j_{th} variant for the exposure-instrument regression. $\hat{\beta}_j$ is the ratio between the coefficient for the exposure-instrument regression and the coefficient for the outcome-instrument regression for the j_{th} variant. σ_{Yj} is the standard error of the outcome-instrument association for the j_{th} variant. The IVW method is asymptotically equal to the two-stage least squares estimator which is commonly used for individual level data. If all the studied genetic variants satisfy the instrumental variable (IV) assumptions, then the IVW estimate is a consistent estimate of the causal effect, as it is a weighted mean of the individual ratio estimates.

Weighted Median Method (WMM)

The IVW method is an efficient approach to assess the causal estimate between an exposure to an outcome, however, it is biased when even one genetic variant is invalid e.g. due to pleiotropy. However, an estimator can provide consistent causal estimates even if 50% of the genetic variants are valid using the weighted median method (WMM).

A regular median estimate can be obtained through calculating the causal ratio estimates from each genetic variants theta$_j$ = betaY$_j$/betaX$_j$, ordering them, and finding the median. This estimator is inefficient, and the weighted median estimator improves by instead calculating percentiles P$_j$ according to the formula

$$P_j = 100(s_j + w_j/2)$$

Where w$_j$ is a weight assigned to estimate *j* and s$_j$ is the sum of all weights from the lowest estimate up to estimate *j*. Weights w$_j$ are calculated according to the formula:

$$w'_j = \frac{\hat{\gamma}_j^2}{\sigma_{Yj}^2}$$

The WMM estimate will then be a weighted average of the largest estimate smaller than the 50th percentile and the smallest estimate larger than the 50th percentile. The weights are derived from the delta method for the variance of the ratio of two random variables, and represent the reciprocal of the variance of the ratio estimates.

MR-Egger Method

MR-Egger is another method for Mendelian randomization that is less sensitive to bias from pleiotropy. MR-Egger performs a weighted linear regression of the genetic variants-outcome coefficients on the genetic variants-exposure ($\hat{\gamma}$j) coefficients:

$$\hat{\Gamma}_j = \beta_{0E} + \beta_E \hat{\gamma}_j$$

The value of the intercept term "β$_{0E}$" can be interpreted as the average pleiotropic effect across the genetic variants. Pleiotropic effect can be defined as the effect of the genetic variant on the outcome that is not mediated by the exposure.

MR-Egger provides a causal estimate even when all the genetic variants are invalid but under a weaker assumption known as the InSIDE (instrument strength independent of direct effect) assumption. For example, if the association of the jth variant with the outcome Ѓj is

$$Ѓj = βγj + αj$$

Where αj is the pleiotropic effect of the variant, then the InSIDE assumption postulates that the pleiotropic effects **α** must be independent of the parameters **γ** denoting the strength of the instrument.
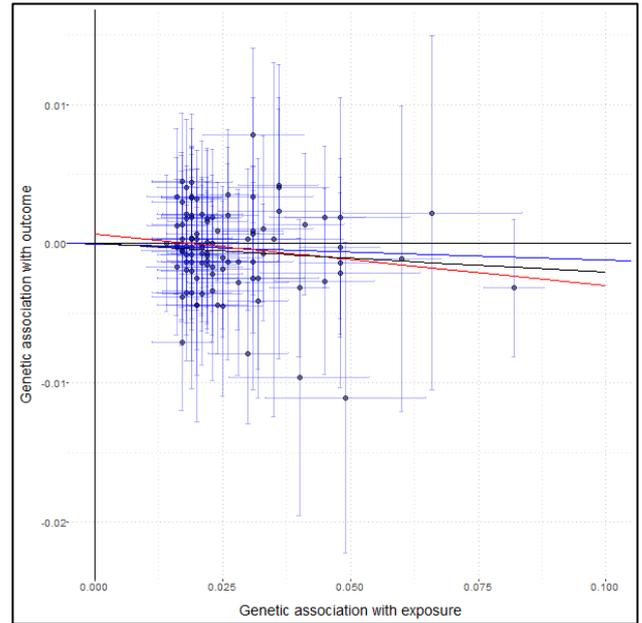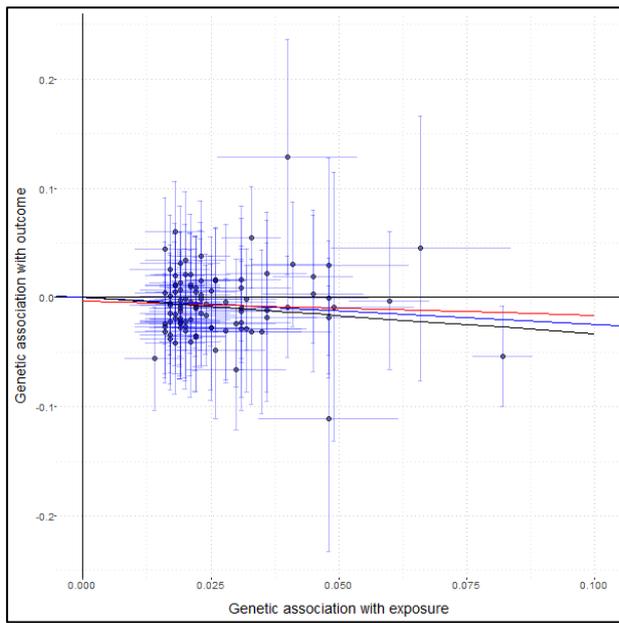
Supplemental Figure 1A Scatter plots of genetic association between exposure (BMI) and outcome (arachidonic acid) across the combined sample of three Swedish cohorts (PIVUS, ULSAM and TwinGene) (left) and KORA/TwinsUK (right). Regression lines indicate instrumental variable estimates for Mendelian randomization analyses using the inverse variance weighted method (blue), weighted median method (black), and MR-Egger method (red).
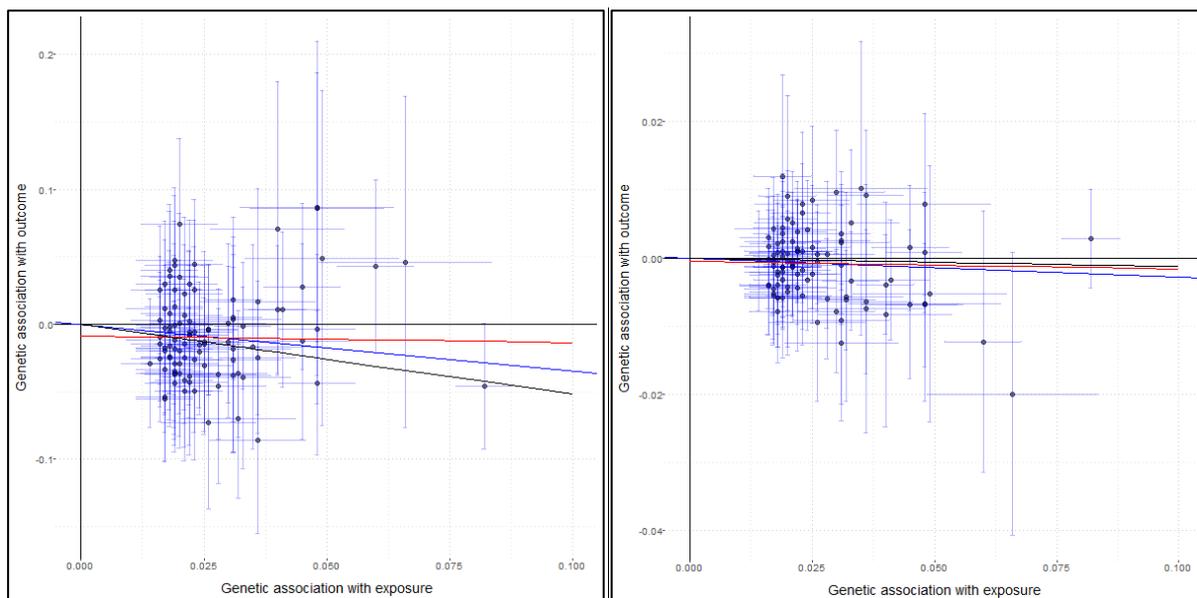
Supplemental Figure 1B Scatter plots of genetic association between exposure (BMI) and outcome (dodecanedioic acid) across the combined sample of three Swedish cohorts (PIVUS, ULSAM and TwinGene) (left) and KORA/TwinsUK (right). Regression lines indicate instrumental variable estimates for Mendelian randomization analyses using the inverse variance weighted method (blue), weighted median method (black), and MR-Egger method (red).
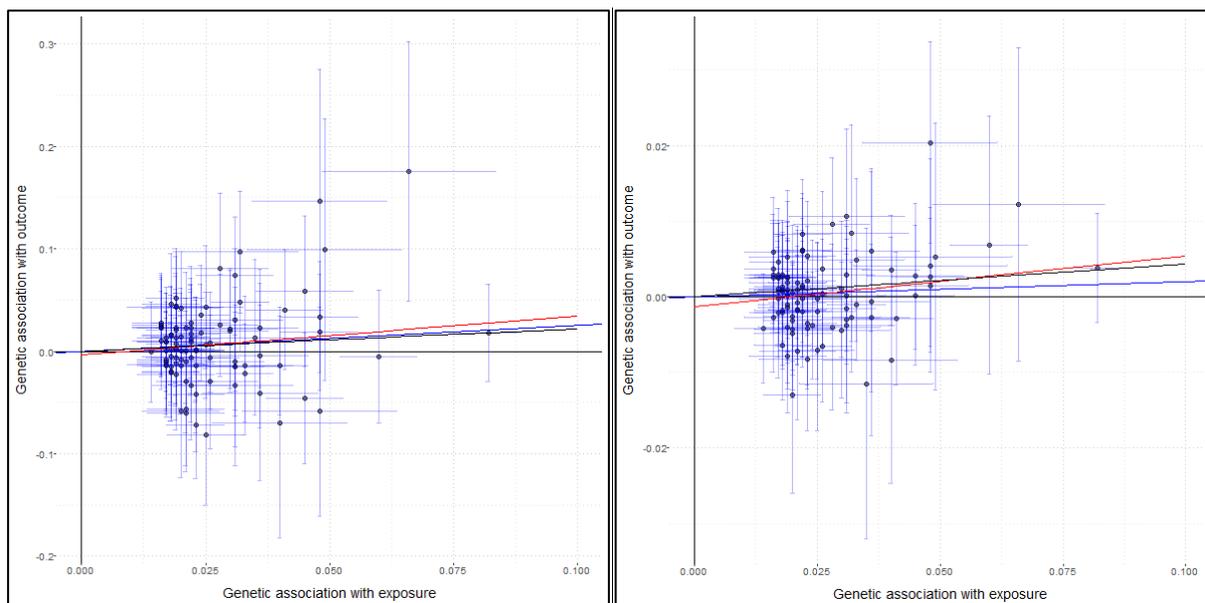
Supplemental Figure 1C Scatter plots of genetic association between exposure (BMI) and outcome (creatine) across the combined sample of three Swedish cohorts (PIVUS, ULSAM and TwinGene) (left) and KORA/TwinsUK (right). Regression lines indicate instrumental variable estimates for Mendelian randomization analyses using the inverse variance weighted method (blue), weighted median method (black), and MR-Egger method (red).
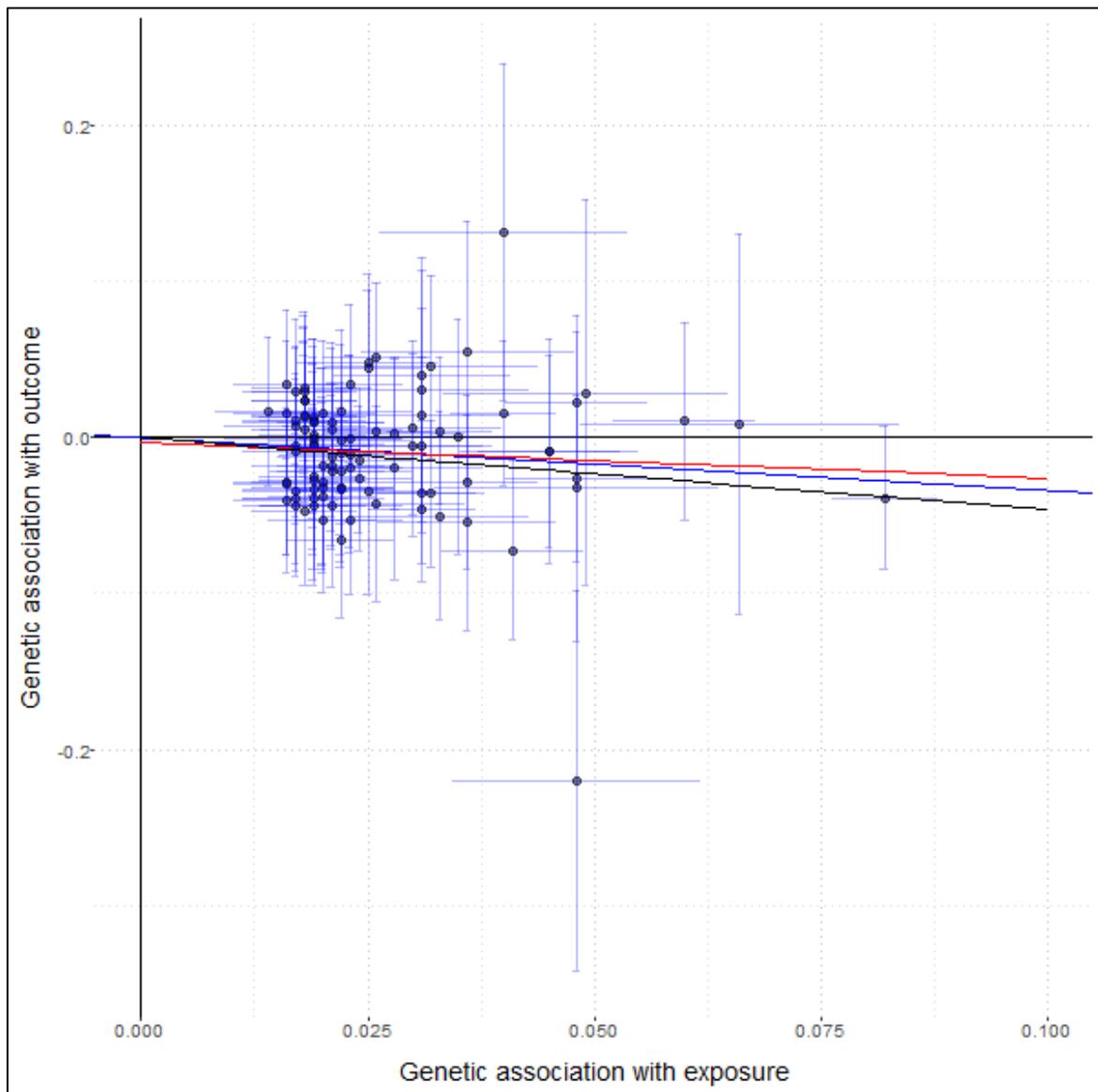
Supplemental Figure 1D Scatter plots of genetic association between exposure (BMI) and outcome (Lysophosphatidylcholine(P-16:0)) across the combined sample of three Swedish cohorts (PIVUS, ULSAM and TwinGene). Regression lines indicate instrumental variable estimates for Mendelian randomization analyses using the inverse variance weighted method (blue), weighted median method (black), and MR-Egger method (red).
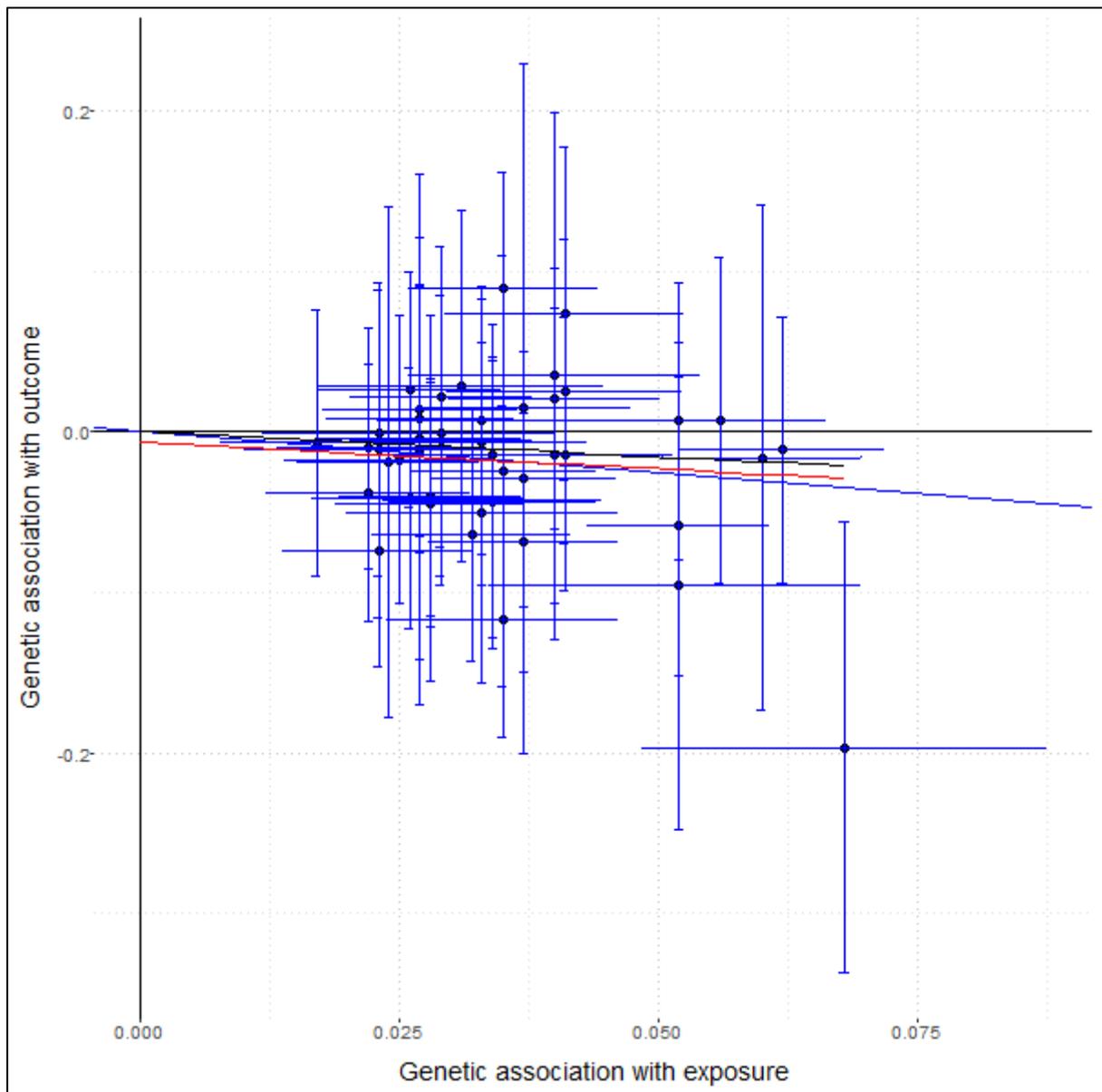
Supplemental Figure 2A Scatter plots of genetic association with exposure (WHRadjBMI) and outcome (Sphingomyelin (32:2)) across the combined sample of three Swedish cohorts (PIVUS, ULSAM and TwinGene). Regression lines indicate instrumental variable estimates for Mendelian randomization analysis using the inverse variance weighted (blue), weighted median method (black), and MR-Egger method (red) Mendelian randomization analysis.

Supplemental Figure 1. a) Dodecanedioic acid was annotated at metabolomics standards initiative (MSI 1) using an external standard and detected in 5% of TwinGene samples, 3% of PIVUS samples, and 2% of ULSAM samples. b) EIC of M+H = 253.141 m/z and RT across the three cohorts is shown to the left and the corresponding mass spectra to the right. c) EIC of M+H = 253.141 and RT from a standard injection to the left and the corresponding mass spectra to the left. d) Mass spectral matching between metabolic feature and external standard feature. Due to the low abundance of Dodecanedioic acid in samples we are unable to produce idMSMS mass spectral matching.

Supplemental Figure 3. a) Dodecanedioic acid was annotated at metabolomics standards initiative (MSI 1) using an external standard and detected in 5% of TwinGene samples, 3% of PIVUS samples, and 2% of ULSAM samples. b) EIC of M+H = 253.141 m/z and RT across the three cohorts is shown to the left and the corresponding mass spectra to the right. c) EIC of M+H = 253.141 and RT from a standard injection to the left and the corresponding mass spectra to the left. d) Mass spectral matching between metabolic feature and external standard feature. Due to the low abundance of Dodecanedioic acid in samples we are unable to produce idMSMS mass spectral matching.
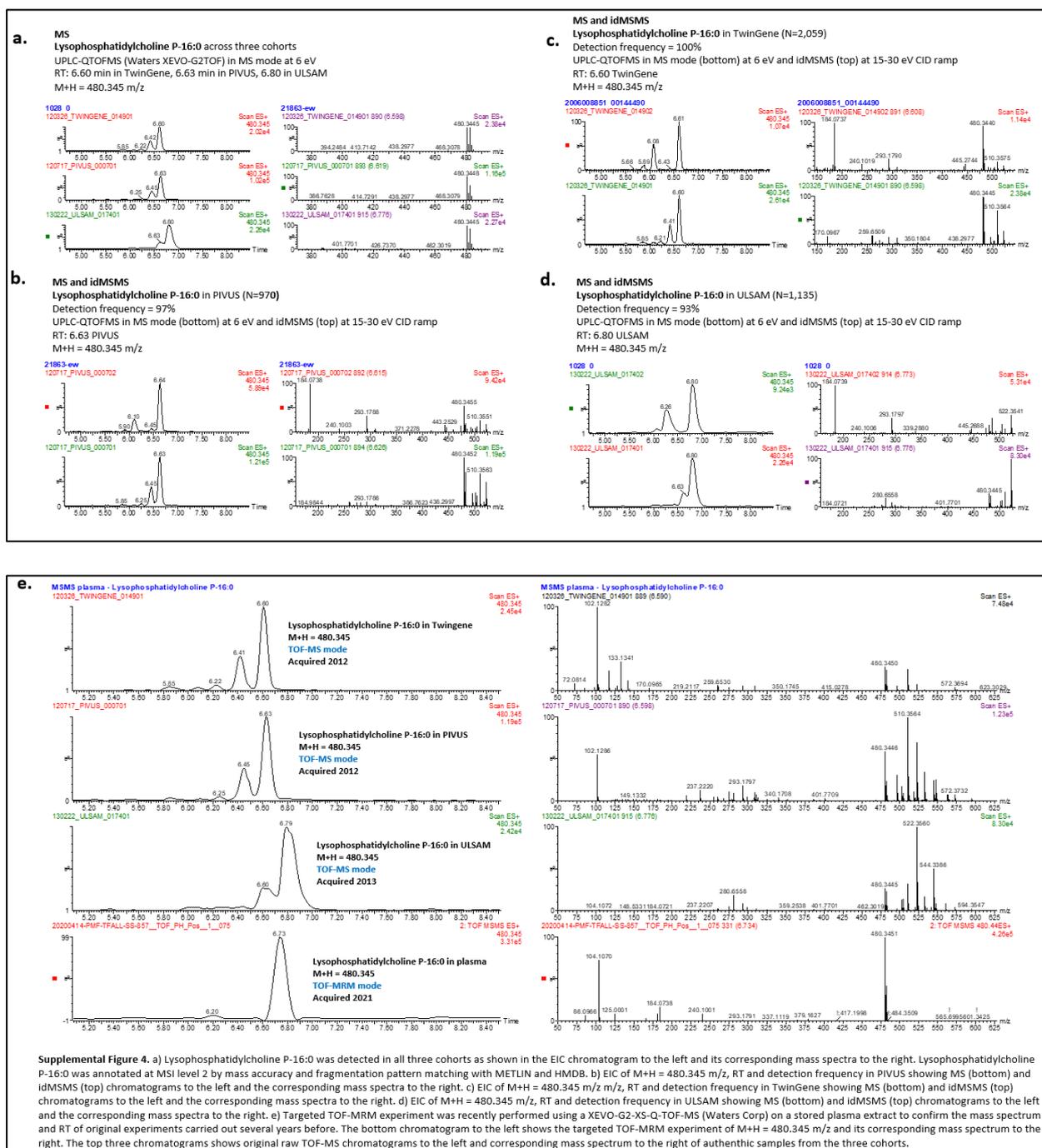
Supplemental Figure 4. a) Lysophosphatidylcholine P-16:0 was detected in all three cohorts as shown in the EIC chromatogram to the left and its corresponding mass spectra to the right. Lysophosphatidylcholine P-16:0 was annotated at MSI level 2 by mass accuracy and fragmentation pattern matching with METLIN and HMDB. b) EIC of M+H = 480.345 m/z, RT and detection frequency in PIVUS showing MS (bottom) and idMSMS (top) chromatograms to the left and the corresponding mass spectra to the right. c) EIC of M+H = 480.345 m/z m/z, RT and detection frequency in TwinGene showing MS (bottom) and idMSMS (top) chromatograms to the left and the corresponding mass spectra to the right. d) EIC of M+H = 480.345 m/z, RT and detection frequency in ULSAM showing MS (bottom) and idMSMS (top) chromatograms to the left and the corresponding mass spectra to the right.

Supplemental Figure 5. a) Arachidonic acid was detected in all three cohorts and external standard as shown in the EIC chromatogram to the left and its corresponding mass spectra to the right. Arachidonic acid was annotated at MSI level 1 by mass spectral and RT matching. b) EIC of M+H = 305.248 m/z, RT and detection frequency in PIVUS showing MS (bottom) and idMSMS (top) chromatograms to the left and the corresponding mass spectra to the right. c) EIC of M+H = 305.248 m/z, RT and detection frequency in TwinGene showing MS (bottom) and idMSMS (top) chromatograms to the left and the corresponding mass spectra to the right. d) EIC of M+H = 305.248 m/z, RT and detection frequency in ULSAM showing MS (bottom) and idMSMS (top) chromatograms to the left and the corresponding mass spectra to the right. e) Targeted TOF-MRM experiment was recently performed using a XEVO-G2-XS-Q-TOF-MS (Waters Corp) on a stored plasma extract to confirm the mass spectrum and RT of original experiments carried out several years before. The bottom chromatogram to the left shows the targeted TOF-MRM experiment of M+H = 305.248 m/z and its corresponding mass spectrum to the right. The top three chromatograms shows original raw TOF-MS chromatograms to the left and corresponding mass spectrum to the right of authentic samples from the three cohorts.
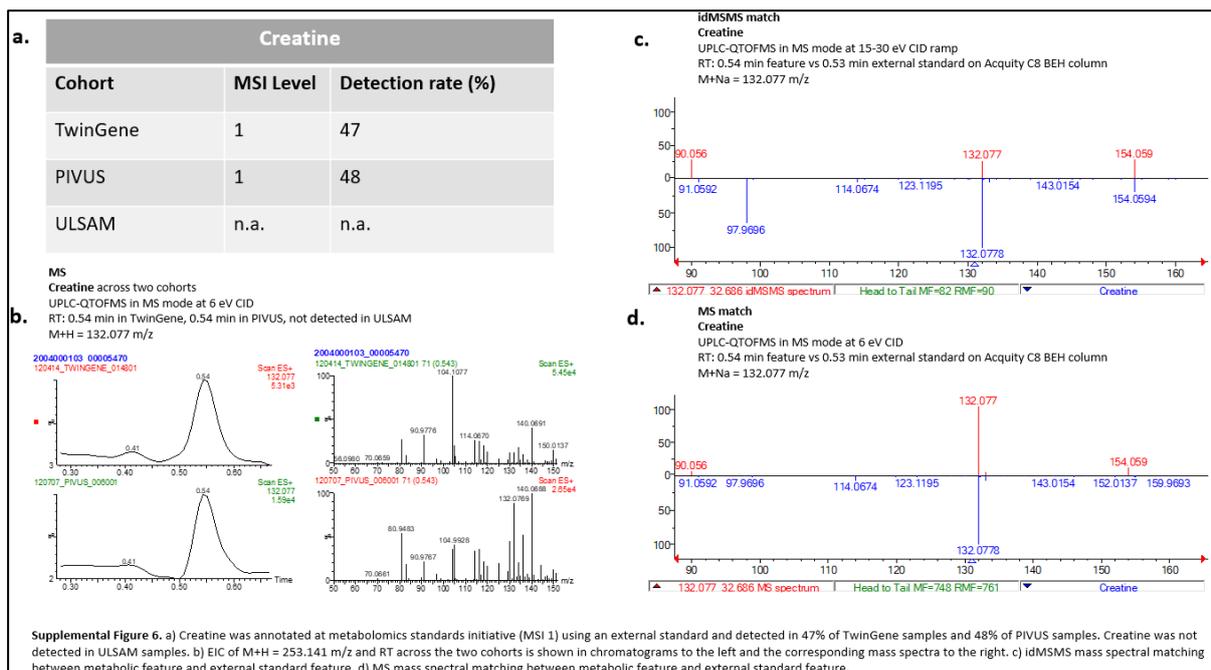
Supplemental Figure 6. a) Creatine was annotated at metabolomics standards initiative (MSI 1) using an external standard and detected in 47% of TwinGene samples and 48% of PIVUS samples. Creatine was not detected in ULSAM samples. b) EIC of M+H = 253.141 m/z and RT across the two cohorts is shown in chromatograms to the left and the corresponding mass spectra to the right. c) idMSMS mass spectral matching between metabolic feature and external standard feature. d) MS mass spectral matching between metabolic feature and external standard feature.
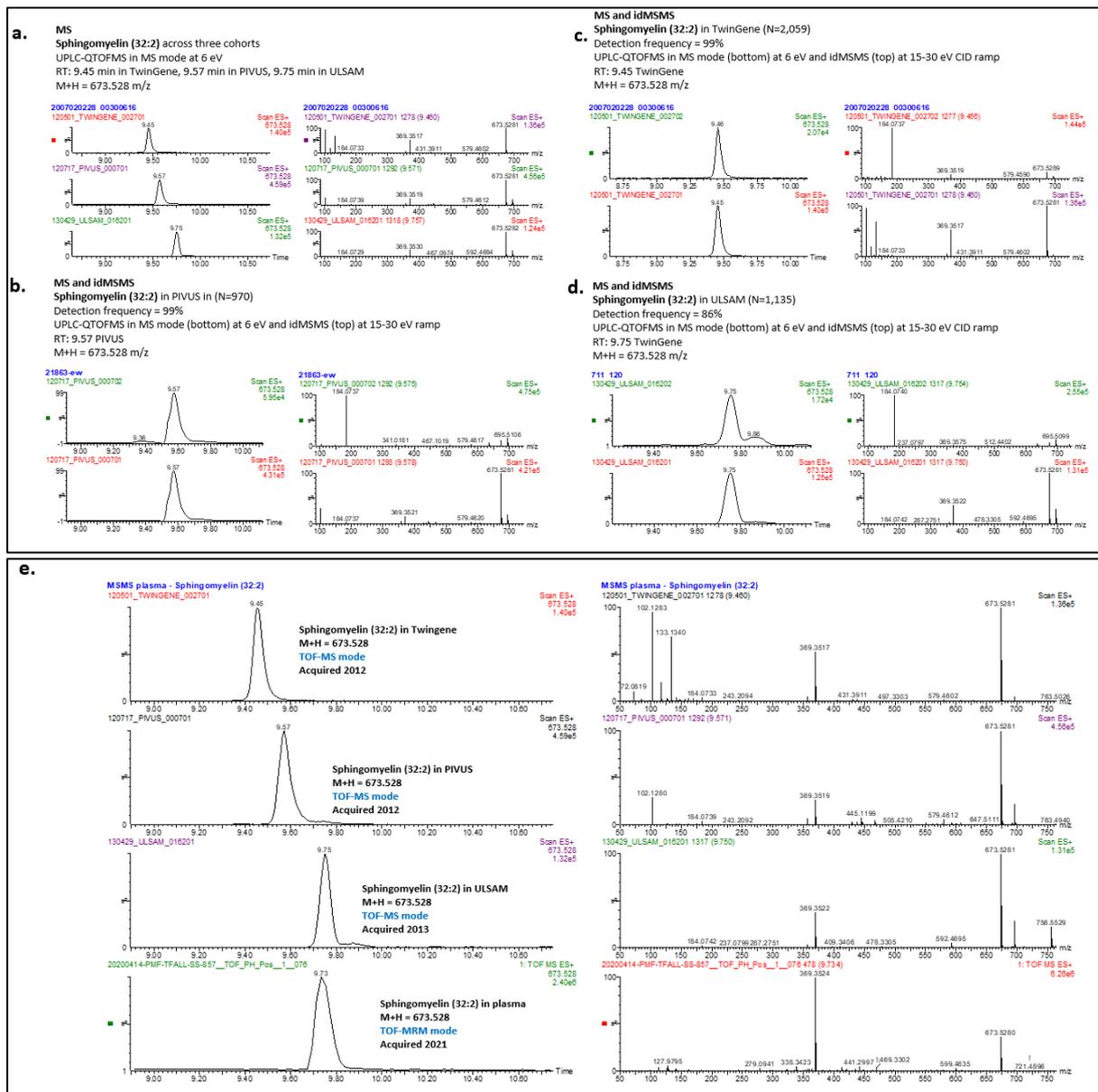
Supplemental Figure 7. a) Sphingomyelin (32:2) was detected in all three cohorts as shown in the EIC chromatogram to the left and its corresponding mass spectra to the right. Sphingomyelin (32:2) was annotated at MSI level 2 by mass accuracy and fragmentation pattern matching with METLIN and HMDB. b) EIC of M+H = 673.528 m/z, RT and detection frequency in PIVUS showing MS (bottom) and idMSMS (top) chromatograms to the left and the corresponding mass spectra to the right. c) EIC of M+H = 673.528 m/z, RT and detection frequency in TwinGene showing MS (bottom) and idMSMS (top) chromatograms to the left and the corresponding mass spectra to the right. d) EIC of M+H = 673.528 m/z, RT and detection frequency in ULSAM showing MS (bottom) and idMSMS (top) chromatograms to the left and the corresponding mass spectra to the right. e) Targeted TOF-MRM experiment was recently performed using a XEVO-G2-XS-Q-TOF-MS (Waters Corp) on a stored plasma extract to confirm the mass spectrum and RT of original experiments carried out several years before. The bottom chromatogram to the left shows the targeted TOF-MRM experiment of M+H = 673.528 m/z and its corresponding mass spectrum to the right. The top three chromatograms shows original raw TOF-MS chromatograms to the left and corresponding mass spectrum to the right of authentic samples from the three cohorts.

REFERENCES

1. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015; **518**: 197-206.
2. Shungin D, Winkler TW, Croteau-Chonka DC, et al. New genetic loci link adipose and insulin biology to body fat distribution. *Nature* 2015; **518**: 187-96.
3. Sumner LW, Amberg A, Barrett D, et al. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* 2007; **3**: 211-21.
4. Ganna A, Fall T, Salihovic S, et al. Large-scale non-targeted metabolomic profiling in three human population-based studies. *Metabolomics* 2015; **12**: 4.
5. Ganna A, Salihovic S, Sundstrom J, et al. Large-scale metabolomic profiling identifies novel biomarkers for incident coronary heart disease. *PLoS Genet* 2014; **10**: e1004801.
6. Fall T, Salihovic S, Brandmaier S, et al. Non-targeted metabolomics combined with genetic analyses identifies bile acid synthesis and phospholipid metabolism as being associated with incident type 2 diabetes. *Diabetologia* 2016; **59**: 2114-24.
7. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res Synth Methods* 2010; **1**: 97-111.
8. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG. Sensitivity Analyses for Robust Causal Inference from Mendelian Randomization Analyses with Multiple Genetic Variants. *Epidemiology* 2017; **28**: 30-42.
9. Bowden J, Holmes MV. Meta-analysis and Mendelian randomization: A review. *Res Synth Methods* 2019; **10**: 486-96.
10. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Method Estimator. *Genet Epidemiol* 2016; **40**: 304-14.