Supplementary table 1. Phenotypic characteristics of analyzed samples. (NA: Not available)

Sample origin	Number of families	Comorbid OCD (%)	Comorbid ADHD (%)	Autism Spectrum Disorder (%)	Male (%)	Female (%)
Greek	10	16.67	42.86	0	85.71	14.29
Albanian	6	NA	NA	NA	NA	NA
Hungarian	84	9.64	34.94	7.23	81.93	18.07
Polish	36	6.06	33.33	3.13	75.76	24.24
Spanish	19	66.67	41.18	0	80.00	20.00
Italian	50	79.07	69.05	0	86.05	13.95
German (Munich)	69	28.36	4.48	0	78.57	21.43
German (GCTS)	96	33	NA	NA	80	20
Canadian	150	28.28	57.24	NA	83.45	16.55

Supplementary table 2. Selection of tagging SNPs (tSNPs) at *HDC* using the HapMap CEPH European population as reference (http://hapmap.ncbi.nlm.nih.gov/). The r^2 threshold for tSNP selection was set to 0.8. Variation across *HDC* (chr15:48316000-48356000) was captured by the selected tSNPs with a mean r^2 of 0.968.

Selected tSNP	captured SNPs
rs1365503	rs1365503,rs2187576,rs2114447,rs10220755
rs854157	rs854160,rs860526,rs854157
rs854163	rs854159,rs854163
rs854158	rs854158,rs2238292
rs1894236	rs1894236,rs8029889
rs1677254	rs2853766,rs1677254
rs854151	rs854151
rs854150	rs854150
rs9920021	rs9920021
rs2070595	rs2070595
rs2070596	rs2070596
rs7182203	rs7182203

Supplementary Table 3. Transmission test for linkage disequilibrium in studied populations (total of 520 families) for single markers tested at the *HDC* gene. Analysis for the Greek and Albanian samples is not shown separately, because individual sample sizes are too small. (test implemented and P values determined by Haploview; T:U=Transmitted:Untransmitted)

Population	Rs#	Over-transmitted allele	T:U*	Chi square	P value	Permutation P value
Hungarian	rs854150	G	47:36	1.458	0.2273	
(84 families)	rs854151	G	33:30	0.143	0.7055	
	rs2070596	T	29:17	3.13	0.0768	
	rs2070595	G	35:27	1.032	0.3096	
	rs2853766	A	36:30	0.545	0.4602	
	rs1894236	C	29:15	4.455	0.0348	
	rs854157	T	44:42	0.047	0.8292	
	rs7182203	G	34:27	0.803	0.3701	
	rs854158	G	40:32	0.889	0.3458	
	rs854163	A	38:30	0.941	0.332	
	rs9920021	C	40:29	1.754	0.1854	
	rs1365503	G	36:34	0.057	0.8111	
Polish	rs854150	G	20:19	0.026	0.8728	
(36 families)	rs854151	A	16:15	0.032	0.8575	
	rs2070596	A	13:5	3.556	0.0593	
	rs2070595	T	18:9	3.0	0.0833	
	rs2853766	-	17:17	0	1.0	
	rs1894236	T	12:8	0.8	0.3711	
	rs854157	T	26:16	2.381	0.1228	
	rs7182203	A	17:10	1.815	0.1779	
	rs854158	-	20:20	0	1.0	
	rs854163	A	17:15	0.125	0.7237	
	rs9920021	A	15:13	0.143	0.7055	
	rs1365503	A	12:14	1.778	0.1824	
Spanish	rs854150	G	5:4	0.111	0.7389	
(19 families)	rs854151	G	7:2	2.778	0.0956	
	rs2070596	T	12:3	5.4	0.0201	
	rs2070595	G	13:6	2.579	0.1083	
	rs2853766	A	7:2	2.778	0.0956	
	rs1894236	C	13:3	6.25	0.0124	
	rs854157	C	11:5	2.25	0.1336	
	rs7182203	G	16:5	5.762	0.0164	
	rs854158	G	6:4	0.4	0.5271	
	rs854163	A	7:2	2.778	0.0956	
	rs9920021	-	8:8	0	1.0	
	rs1365503	G	10:9	0.053	0.8185	

Population	Rs#	Over-transmitted allele	T:U*	Chi square	P value	Permutation P value
Italian	rs854150	G	29:22	0.961	0.327	
(50 families)	rs854151	A	20:18	0.105	0.7456	
	rs2070596	A	21:14	1.4	0.2367	
	rs2070595	T	26:11	6.081	0.0137	
	rs2853766	A	18:17	0.029	0.8658	
	rs1894236	T	21:14	1.4	0.2367	
	rs854157	T	32:16	5.333	0.0209	
	rs7182203	**A**	**26:10**	**7.111**	**0.0077**	**0.0480**
	rs854158	-	20:20	0	1.0	
	rs854163	G	20:17	0.243	0.6219	
	rs9920021	A	22:21	0.023	0.8788	
	rs1365503	G	23:20	0.209	0.6473	
German	rs854150	G	91:66	3.981	0.046	
(165 families)	rs854151	G	61:60	0.008	0.9276	
	rs2070596	T	61:56	0.214	0.6439	
	rs2070595	T	74:67	0.348	0.5555	
	rs2853766	A	63:54	0.692	0.4054	
	rs1894236	C	63:52	1.052	0.305	
	rs854157	T	82:71	0.791	0.3738	
	rs7182203	A	73:63	0.735	0.3912	
	rs854158	G	72:69	0.064	0.8005	
	rs854163	A	66:58	0.516	0.4725	
	rs9920021	A	74:64	0.725	0.3946	
	rs1365503	G	79:60	2.597	0.1071	
Canadian	rs854150	G	87:66	2.882	0.0896	
(150 families)	rs854151	G	66:55	1.0	0.3173	
	rs2070596	T	65:40	5.952	0.0147	0.1100
	rs2070595	G	91:68	3.327	0.0681	
	rs2853766	A	70:50	3.333	0.0679	
	rs1894236	**C**	**68:37**	**9.152**	**0.0025**	**0.0240**
	rs854157	-	83:83	0.0	1.0	
	rs7182203	G	79:64	1.573	0.2097	
	rs854158	G	82:66	1.73	0.1884	
	rs854163	A	73:58	1.718	0.19	
	rs9920021	**A**	**75:42**	**9.308**	**0.0023**	**0.0210**
	rs1365503	**A**	**100:55**	**13.065**	**0.0003**	**0.0010**

Supplementary Table 4. Transmission test for linkage disequilibrium in studied populations (total of 520 trios) for SNP haplotypes encompassing the identified haplotype blocks in each population across the *HDC* gene. Haplotype blocks are shown in figure 1 of the main text. Analysis for the Greek and Albanian samples is not shown separately, because individual sample sizes are too small. (test implemented and P values determined by Haploview; T:U=Transmitted:Untransmitted; Permutation P value was calculated after 1,000 permutations).

Population	Haplotype	Frequency	T:U*	Chi square	P value	Permutation P value
Hungarian	Block 1					
(84 families)	CA	0.585	35.3 : 45.6	1.317	0.2512	
	GG	0.294	38.5 : 33.0	0.419	0.5173	
	GA	0.108	18.5 : 12.4	1.186	0.2761	
	CG	0.014	2.0 : 3.2	0.283	0.5946	
	Block 2					
	TGGCCG	0.395	43.0 : 41.0	0.048	0.8262	
	TGACTG	0.269	36.0 : 28.0	1.006	0.316	
	ATGTTA	0.185	17.0 : 28.0	2.684	0.1013	
	TTGCTA	0.103	17.0 : 13.0	0.53	0.4665	
	TTGCTG	0.026	5.0 : 4.0	0.111	0.7389	
	TTGCCG	0.010	0.0 : 2.0	2.013	0.156	
	Block 3					
	AG	0.646	32.0 : 40.0	0.888	0.346	
	GA	0.295	38.0 : 31.0	0.71	0.3994	
	GG	0.054	8.0 : 7.0	0.067	0.7963	
	Block 4					
	AA	0.402	36.4 : 37.5	0.017	0.8949	
	AG	0.306	32.5 : 42.9	1.428	0.232	
	CG	0.292	41.0 : 29.5	1.88	0.1703	
Polish	Block 1					
(36 families)	CA	0.551	20.0 : 20.0	0	1.0	
	GG	0.319	16.0 : 18.5	0.187	0.6655	
	GA	0.130	8.0 : 5.5	0.48	0.4886	
	Block 2					
	TGGCCG	0.392	14.0 : 24.0	2.632	0.1048	
	TGACTG	0.297	17.0 : 16.0	0.03	0.8618	
	ATGTTA	0.139	12.0 : 7.0	1.316	0.2513	
	TTGCTA	0.107	8.0 : 6.0	0.286	0.593	
	TTGCCG	0.024	2.0 : 1.7	0.017	0.8957	
	ATGCTA	0.013	2.0 : 0.0	2.0	0.1573	
	Block 3					
	AG	0.645	20.1 : 20.0	0	0.9929	
	GA	0.316	17.5 : 16.5	0.029	0.8638	
	GG	0.038	2.5 : 3.6	0.184	0.6678	

Supplementary Table 4 continued

Population	Haplotype	Frequency	T:U*	Chi square	P value	Permutation P value
Spanish	Block 1					
(19 families)	CCG	0.447	12.0 : 5.0	2.882	0.0896	
	TTA	0.327	3.0 : 13.9	6.993	0.0082	
	CTG	0.171	6.0 : 2.0	2.0	0.1573	
	CTA	0.055	3.0 : 3.1	0.003	0.954	
Italian	Block 1					
(50 families)	ATGGC	0.375	14.1 : 28.0	4.602	0.0319	
	AATGT	0.227	21.0 : 14.0	1.4	0.2367	
	GTGAC	0.193	18.0 : 18.0	0	1.0	
	ATTGC	0.170	18.0 : 10.0	2.286	0.1306	
	GTGGC	0.034	3.0 : 4.1	0.166	0.6838	
	Block 2					
	CG	0.447	16.0 : 32.0	5.333	0.0209	0.0460
	TA	**0.343**	**27.0 : 10.6**	**7.131**	**0.0076**	**0.0350**
	TG	0.210	20.0 : 20.4	0.004	0.9524	
	Block 3					
	AG	0.738	21.8 : 22.1	0.003	0.9592	
	GA	0.200	17.8 : 20.0	0.126	0.7231	
	GG	0.055	7.2 : 5.0	0.379	0.5383	
	Block 4					
	AG	0.364	24.0 : 20.0	0.364	0.5465	
	CG	0.348	21.0 : 22.0	0.023	0.8788	
	AA	0.288	20.0 : 23.0	0.209	0.6473	
German	Block 1					
(165 families)	ATGGC	0.354	71.0 : 83.4	1.0	0.3173	
	GTGAC	0.247	67.0 : 54.0	1.4	0.2367	
	AATGT	0.229	56.0 : 66.0	0.809	0.3684	
	ATTGC	0.132	44.0 : 31.3	2.153	0.1423	
	GTGGC	0.029	6.0 : 11.3	1.598	0.2062	
	Block 2					
	AAG	0.332	77.7 : 70.8	0.33	0.5657	
	GAG	0.332	69.9 : 79.7	0.644	0.4222	
	GGA	0.263	66.5 : 59.3	0.414	0.5199	
	GGG	0.073	16.0 : 20.4	0.531	0.4661	
	Block 3					
	AA	0.369	66.0 : 84.1	2.204	0.1377	
	CG	0.329	64.5 : 74.5	0.719	0.3963	
	AG	**0.302**	**92.5 : 64.4**	**5.064**	**0.0244**	**0.0480**

Supplementary table 4 continued

Population	Haplotype	Frequency	T:U*	Chi square	P value	Permutation P value
Canadian	Block 1					
(150 families)	ATGGC	0.393	88.7 : 83.1	0.183	0.6691	
	GTGAC	0.248	71.9 : 49.9	3.976	0.0461	0.3160
	AATGT	**0.199**	**39.0 : 67.0**	**7.397**	**0.0065**	**0.0430**
	ATTGC	0.126	46.9 : 44.3	0.072	0.7886	
	GTGGC	0.013	2.0 : 5.2	1.397	0.2373	
	Block 2					
	GAG	0.394	83.9 : 84.1	0.0	0.9863	
	AAG	0.295	65.0 : 80.0	1.552	0.2128	
	GGA	0.269	72.0 : 57.8	1.558	0.2119	
	GGG	0.039	13.1 : 12.1	0.04	0.842	
	Block 3					
	AA	**0.371**	**100.0 : 55.5**	**12.713**	**0.0004**	**0.0010**
	AG	0.326	65.0 : 77.5	1.09	0.2964	
	CG	**0.303**	**45.5 : 77.5**	**8.325**	**0.0039**	**0.0370**

Supplementary Figure 1. Rare allele frequencies of studied SNPs across the *HDC* region, in each analyzed population.