

1 **Highly pleiotropic variants of human traits are enriched in genomic regions**
2 **with strong background selection**

3

4

5 Irene Novo, Eugenio López-Cortegano and Armando Caballero

6

7

8

9

10

11 Corresponding author: Irene Novo, Centro de Investigación Mariña, Universidade
12 de Vigo, Departamento de Bioquímica, Xenética e Inmunoloxía, Facultade de Bioloxía,
13 Edificio CC Experimentais, Campus de Vigo, 36310 Vigo, Spain. Tel: +34 986813828. Fax:
14 +34 986812556. Email: irene.novo.gimenez@uvigo.es

15

16

17 **Table S1.** Number of loci analyzed in this study and pleiotropic loci found for each of 41
 18 diseases and other human traits, grouped by functional domain.

Functional domain	Trait	Loci	Pleio loci
Cancer	Basal cell carcinoma	37	15
	Chronic lymphocytic leukemia	35	19
	Lung cancer	66	20
	Prostate cancer	95	32
	Prostate-specific antigen levels	28	11
	Testicular germ cell tumor	50	7
Cardiovascular	Atrial fibrillation	27	4
	Coronary artery disease	43	29
	Coronary heart disease	35	28
	Myocardial infarction	29	23
Dermatological	Atopic dermatitis	25	5
	Psoriasis	47	16
	Vitiligo	51	26
Endocrine	Body mass index	270	110
	Menarche (age at onset)	103	27
	Obesity	37	32
	Type 2 diabetes	81	40
Gastrointestinal	Digestive disease	248	121
	Ulcerative colitis	107	81
Hematological	Glycated hemoglobin levels	51	22
	Mean platelet volume	229	48
	Monocyte count	193	88
	Neutrophil traits	237	102
	Red blood cell traits	42	23
Immunological	Primary biliary cholangitis	47	23
	Rheumatoid arthritis	92	43
	Systemic lupus erythematosus	75	35
	Type 1 diabetes	36	18
Metabolic	Cholesterol	149	68
	HDL	125	64
	Triglycerides	78	51
	Urate levels	34	11
Neurological / Psychiatric	Migraine	36	8
	Parkinson's disease	58	8
	Schizophrenia	107	20
Skeletal	Bone mineral density	54	11
	Height	322	98
	Waist circumference	49	42
	Waist-related traits	121	92
	Waist-hip ratio	41	35
	Waist-to-hip related traits	81	53

19 **Table S2.** List of the most pleiotropic known genes found in this study, indicating the
 20 chromosome, start and end of gene location, the pleiotropic degree and the value of the
 21 average background statistic B for the whole length of the gene.

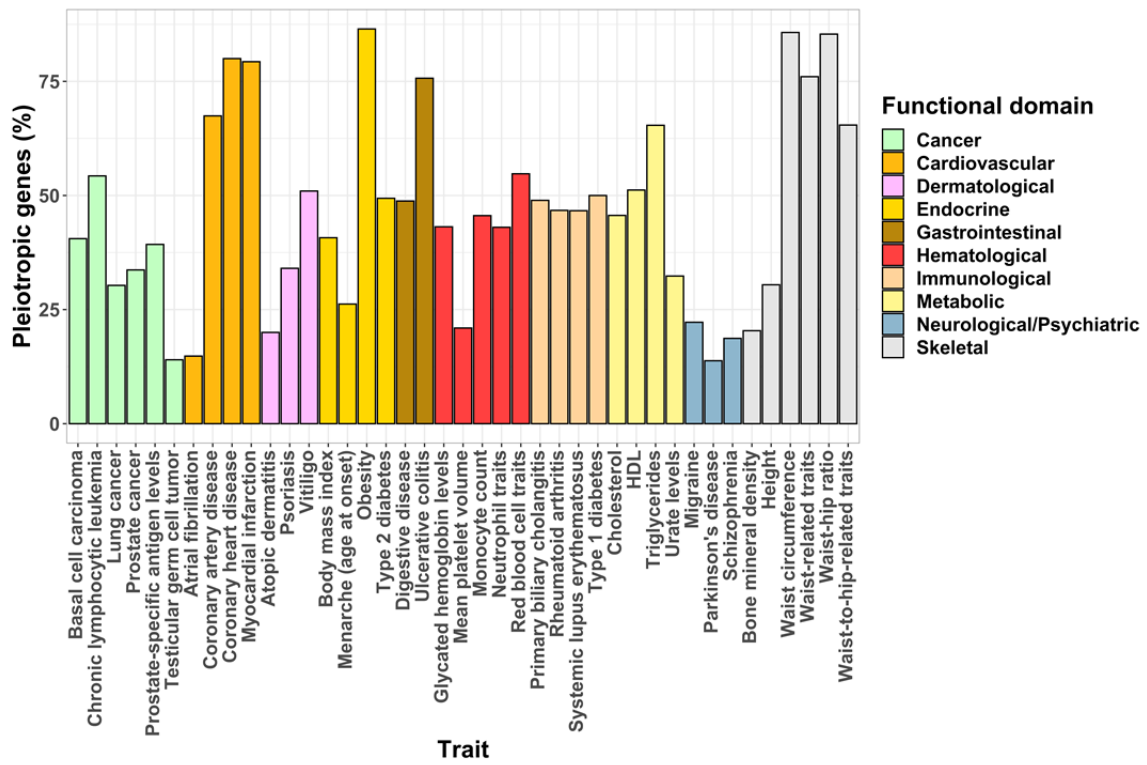
22

Gene	Full name	Ch.	Location	Pleio	B
CDKN2B-AS1 - DMRTA1	CDKN2B antisense RNA 1 - DMRT like family A1	9	21994790 22455739	10	0.918062
FTO	FTO alpha-ketoglutarate dependent dioxygenase	16	53737875 54155853	9	0.867486
HLA-DRB9 - HLA-DRB5	major histocompatibility complex, class II, DR beta 9 - DR beta 5	6	32427597 32498064	8	0.927388
PPARG	peroxisome proliferator activated receptor gamma	3	12328867 12475843	6	0.743163
FADS1	fatty acid desaturase 1	11	61567099 61584475	6	0.498626
GCKR	glucokinase regulator	2	27719706 27746551	6	0.099403
LPA	lipoprotein(a)	6	160952514 161085307	6	0.790012
AP4B1-AS1 PTPN22	AP4B1 antisense RNA 1 - protein tyrosine phosphatase non-receptor type 22	1	114355234 114414381	6	0.126539
ATXN2	ataxin 2	12	111890018 112037477	6	0.116961
HLA-DQB1 - MTCO3P1	major histocompatibility complex, class II, DQ beta 1 - MT-CO3 pseudogene 1	6	32627244 32674580	7	0.940000
JAZF1	JAZF zinc finger 1	7	27870196 28220414	8	0.809438
TYK2	tyrosine kinase 2	19	10461209 10491248	7	0.705571
PEPD	peptidase D	19	33877856 34012697	6	0.829497

23

24

25 **Figure S1.** Proportion of pleiotropic genes found in this study for each of 41 traits grouped in
 26 functional domains.
 27



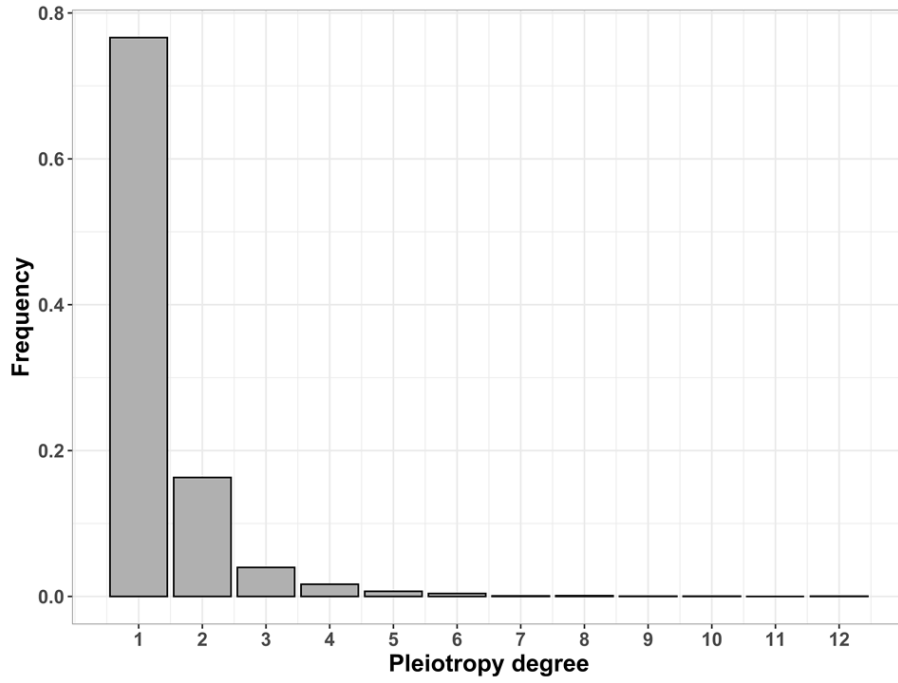
28
 29
 30

31 **Figure S2.** Number of genes found in this study for each degree of pleiotropy.

32

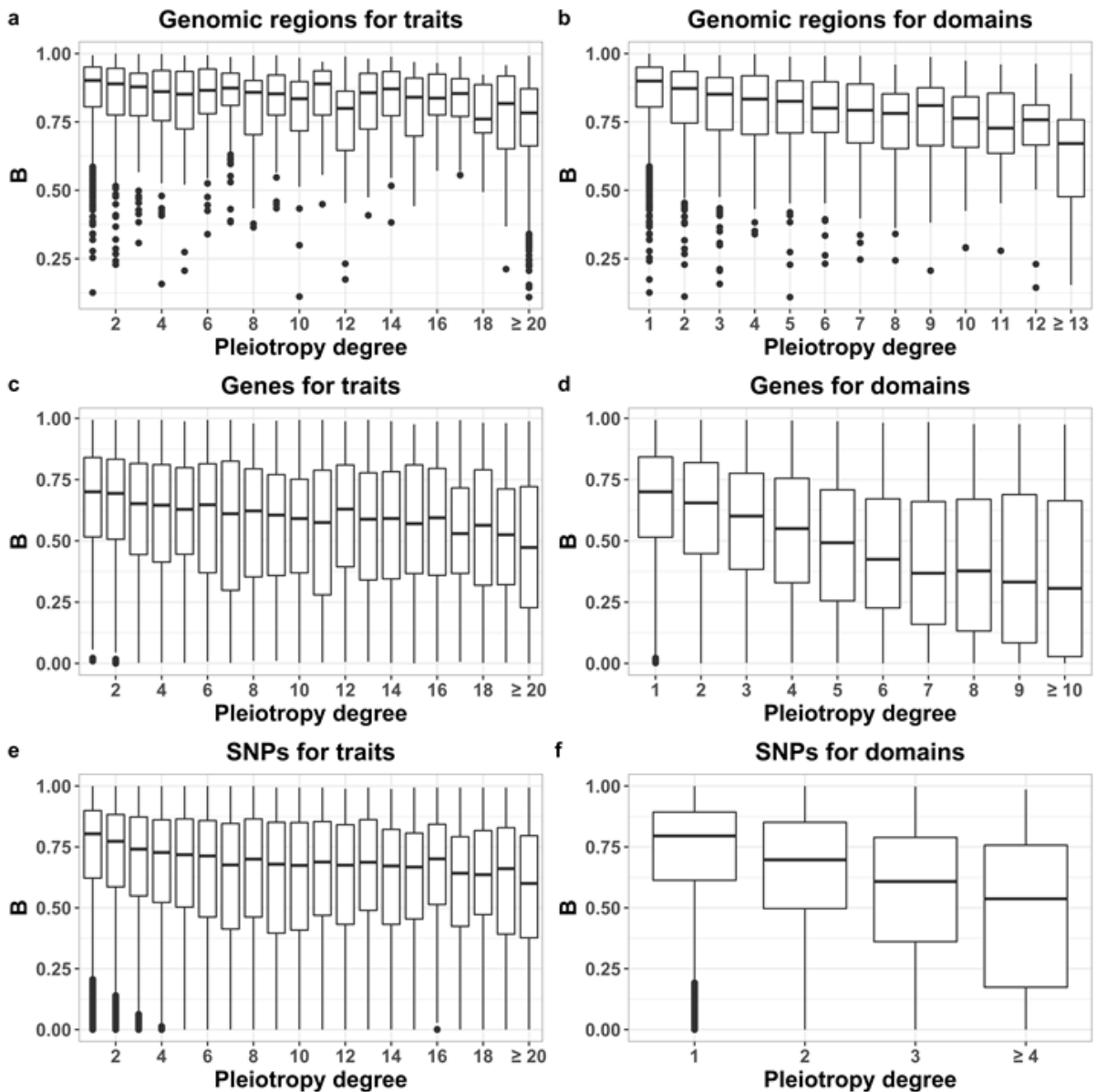
33

34



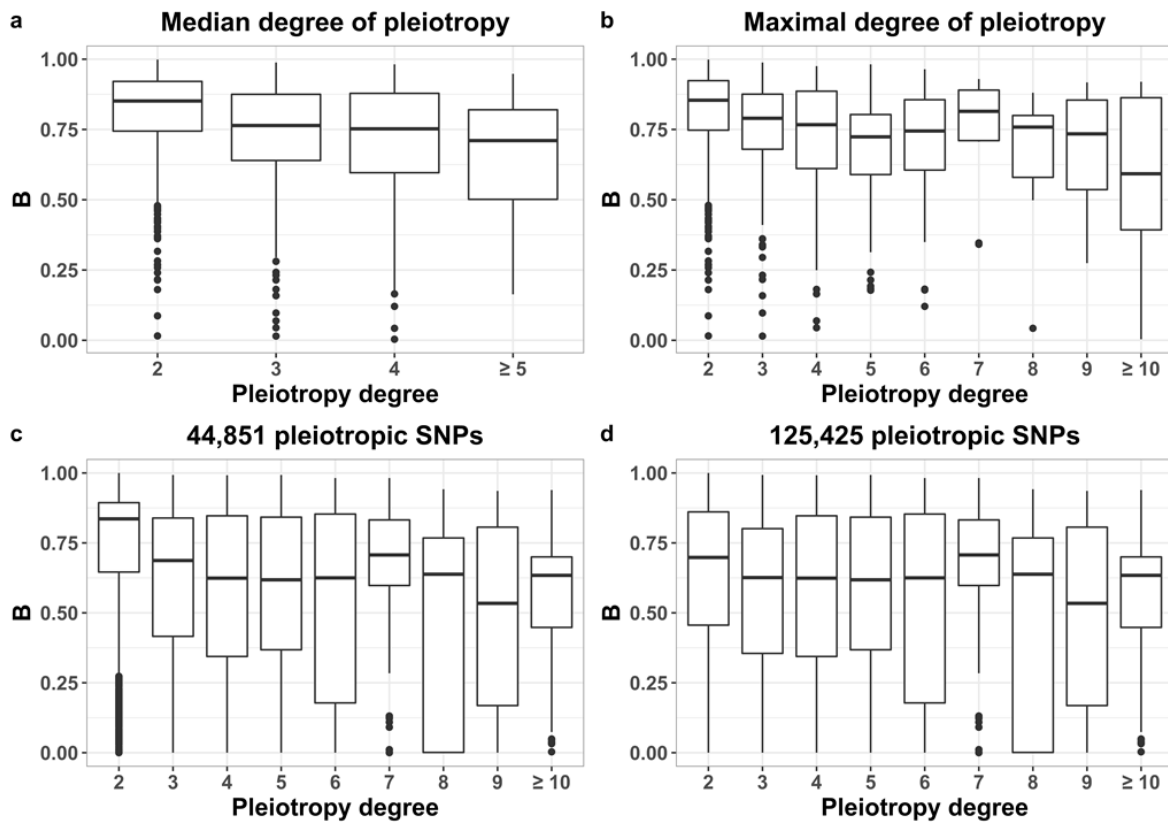
35

36 **Figure S3.** Relationship between the background selection statistic (B) and the degree of
 37 pleiotropy of data found by Watanabe et al. (2019). (a) Average B value for genomic regions
 38 considering pleiotropy across traits ($b = -0.003$, $p < 2 \times 10^{-16}$). (b) Average B value for
 39 genomic regions considering pleiotropy across domains ($b = -0.018$, $p < 2 \times 10^{-16}$) (presented
 40 also in Figure 3c of the main text). (c) Average B value for genes considering traits ($b = -$
 41 0.003 , $p < 2 \times 10^{-16}$); (d) Average B value for genes considering pleiotropy across domains ($b = -$
 42 0.024 , $p < 2 \times 10^{-16}$); (e) Average B value for SNPs considering traits ($b = -0.005$, $p < 2 \times$
 43 10^{-16}); (f) Average B value for SNPs considering pleiotropy across domains ($b = -0.046$, $p < 2$
 44 $\times 10^{-16}$).
 45



46
 47
 48

49 **Figure S4.** Relationship between the background selection statistic (B) and the degree of
50 pleiotropy of data found by Shikov et al. (2020). (a) Average B value for genomic regions
51 considering the median degree of pleiotropy for each region ($b = -0.021$, $p < 2 \times 10^{-16}$). (b)
52 Average B value for genomic regions considering the maximal degree of pleiotropy for each
53 region ($b = -0.022$, $p < 2 \times 10^{-16}$) (presented also in Figure 3d of the main text). (c) Average
54 B value for non-MHC SNPs considering the 44,851 high confidence pleiotropic SNPs ($b = -$
55 0.008 , $p = 3 \times 10^{-6}$); (d) Average B value for non-MHC SNPs considering all 125,425
56 pleiotropic SNPs found ($b = -0.015$, $p < 2 \times 10^{-16}$).
57

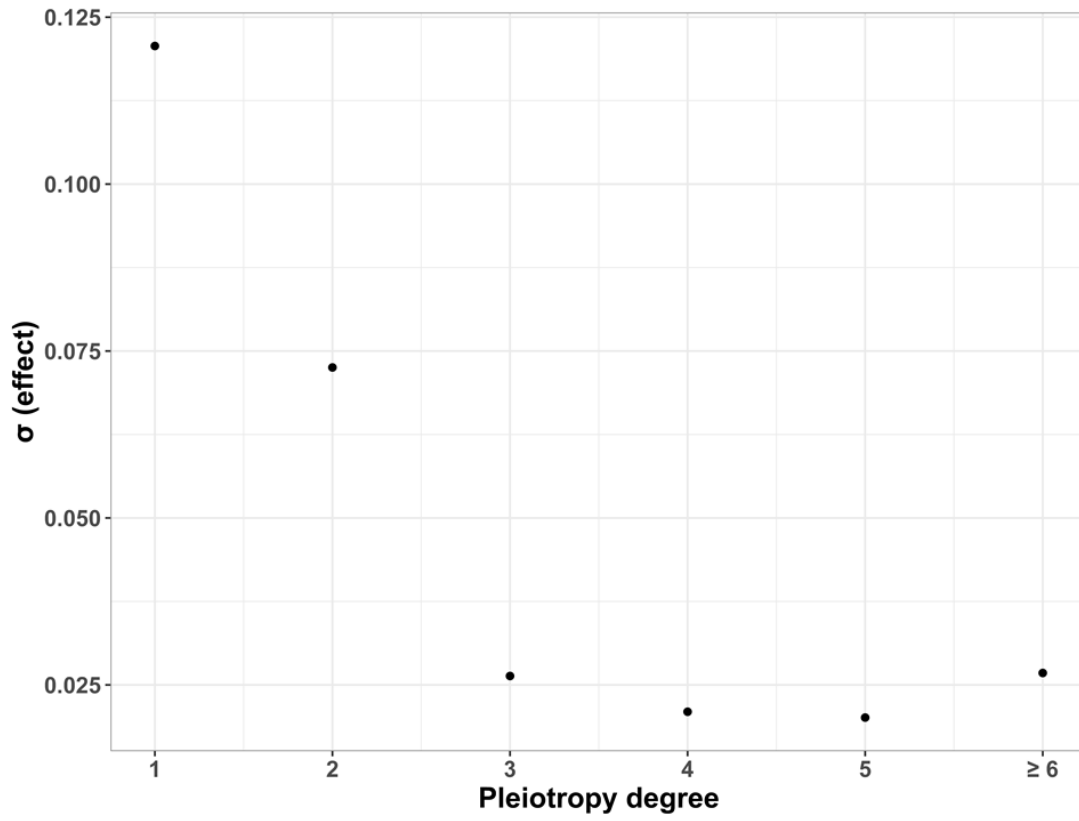


58

59

60

61 **Figure S5.** Relationship between the standard deviation of the average effect sizes of
62 pleiotropic variants and the degree of pleiotropy ($b = -0.008$, $p = 0.008$).
63



64
65
66