

ARTICLE

doi:10.1038/nature25973

Environment dominates over host genetics in shaping human gut microbiota

Daphna Rothschild^{1,2*}, Omer Weissbrod^{1,2*}, Elad Barkan^{1,2*}, Alexander Kurilshikov³, Tal Korem^{1,2}, David Zeevi^{1,2}, Paul I. Costea^{1,2}, Anastasia Godneva^{1,2}, Iris N. Kalka^{1,2}, Noam Bar^{1,2}, Smadar Shilo^{1,2}, Dar Lador^{1,2}, Arnau Vich Vila^{3,4}, Niv Zmora^{5,6,7}, Meirav Pevsner-Fischer⁵, David Israeli⁸, Noa Kosower^{1,2}, Gal Malka^{1,2}, Bat Chen Wolf^{1,2}, Tali Avnit-Sagi^{1,2}, Maya Lotan-Pompan^{1,2}, Adina Weinberger^{1,2}, Zamir Halpern^{7,9}, Shai Carmi¹⁰, Jingyuan Fu^{3,11}, Cisca Wijmenga^{3,12}, Alexandra Zhernakova³, Eran Elinav⁵§ & Eran Segal^{1,2}§

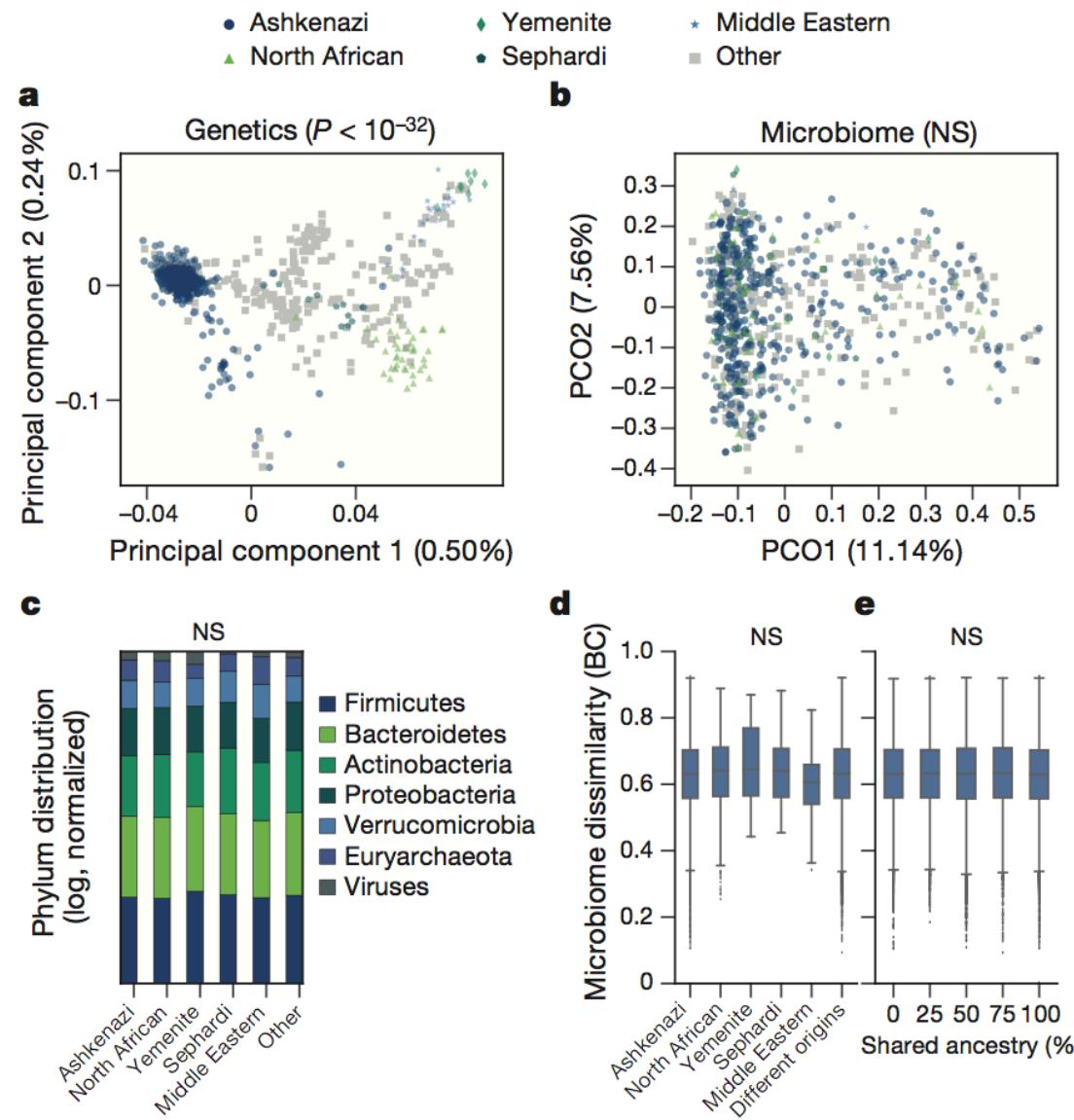


Figure 1 | Genetic ancestry is not significantly associated with microbiome composition. a, Genetic principal components are strongly

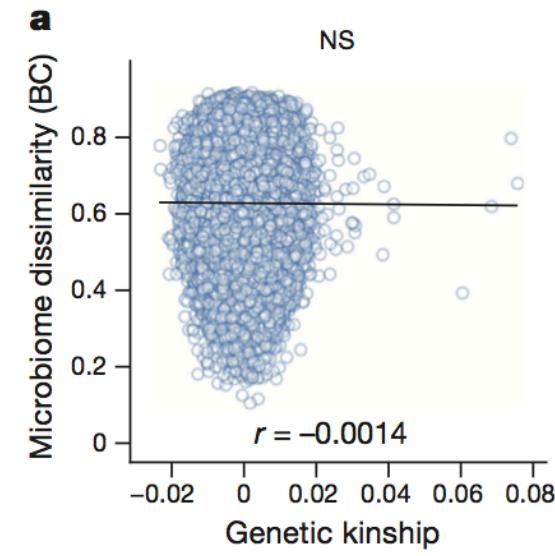
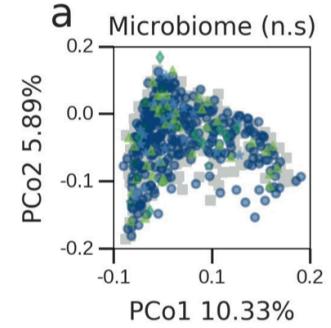


Figure 2 | Genetic kinship is weakly associated with microbiome composition. a, Scatter plot of genetic kinship of pairs of individuals (x

Gene level



b Microbiome dissimilarity (Bray Curtis)

n.s

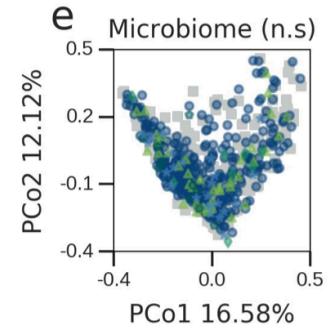
c

n.s

d

n.s

Genus level



f Microbiome dissimilarity (Bray Curtis)

n.s

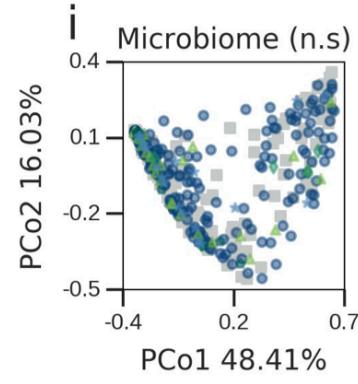
g

n.s

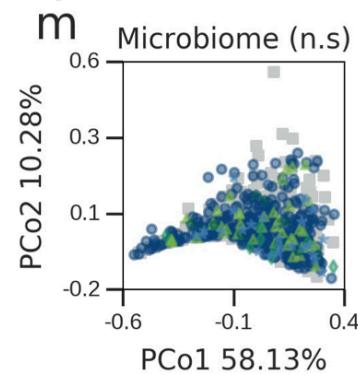
h

n.s

16S Genus level



Phylum level



- Ashkenazi
- ▲ North African
- ◆ Yemenite
- Sephardi
- ★ Middle Eastern
- Other

j Microbiome dissimilarity (Bray Curtis)

n.s

k

n.s

l

n.s

n Microbiome dissimilarity (Bray Curtis)

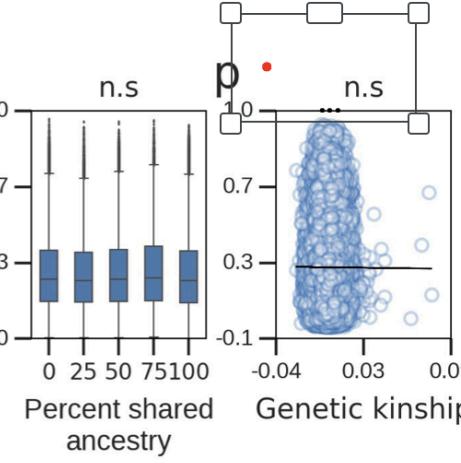
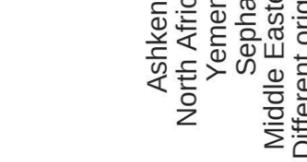
n.s

o

n.s

p

n.s

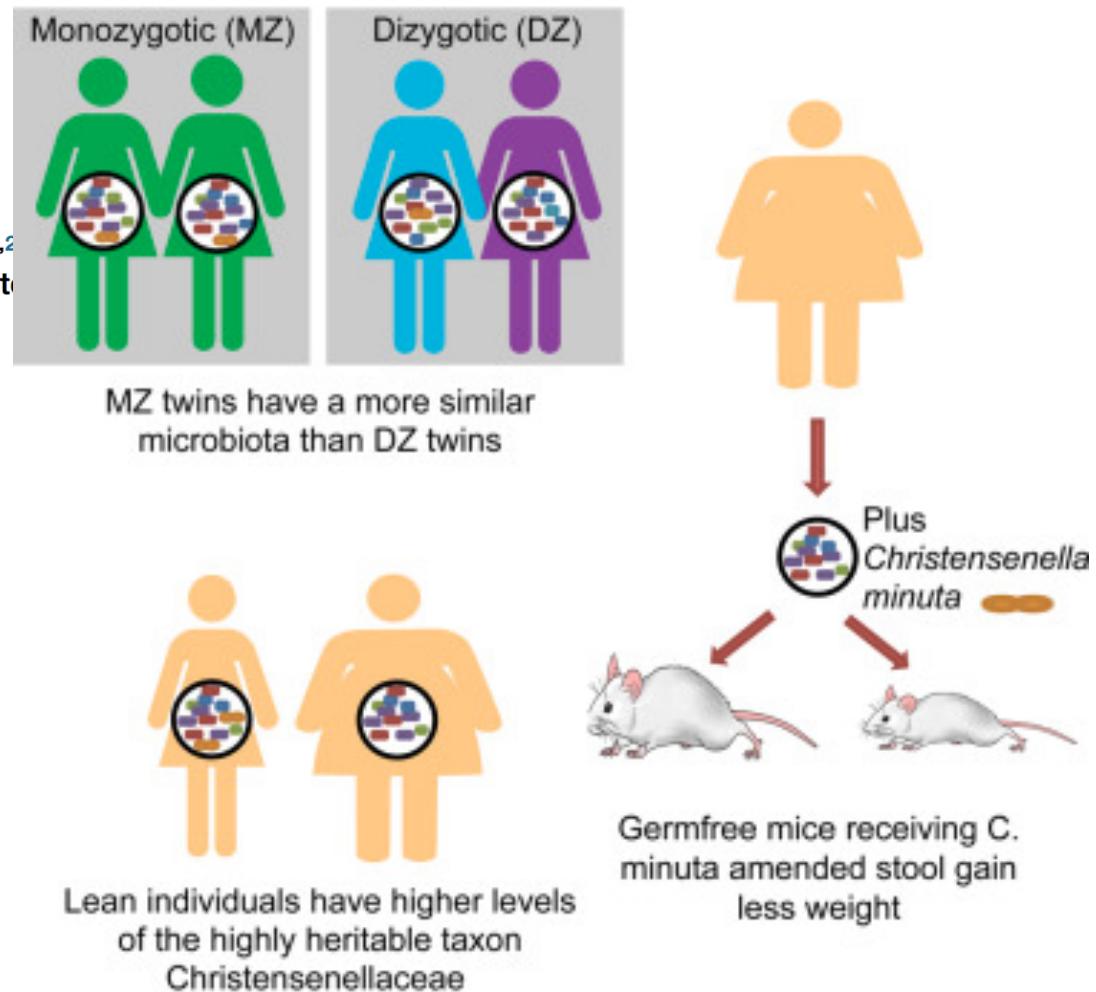


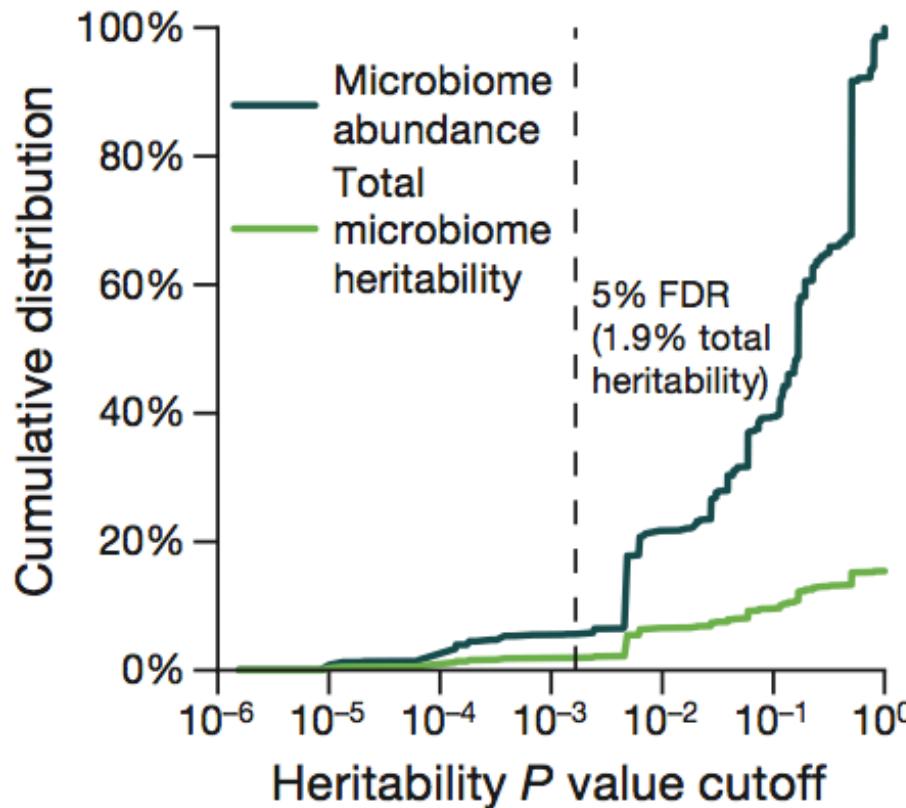
Human Genetics Shape the Gut Microbiome

Julia K. Goodrich,^{1,2} Jillian L. Waters,^{1,2} Angela C. Poole,^{1,2} Jessica L. Sutter,^{1,2} Omry Koren,^{1,2} Michelle Beaumont,³ William Van Treuren,⁴ Rob Knight,^{4,5,6} Jordana T. Bell,³ Timothy D. Spector,³ and Ruth E. Ley^{1,2,*}

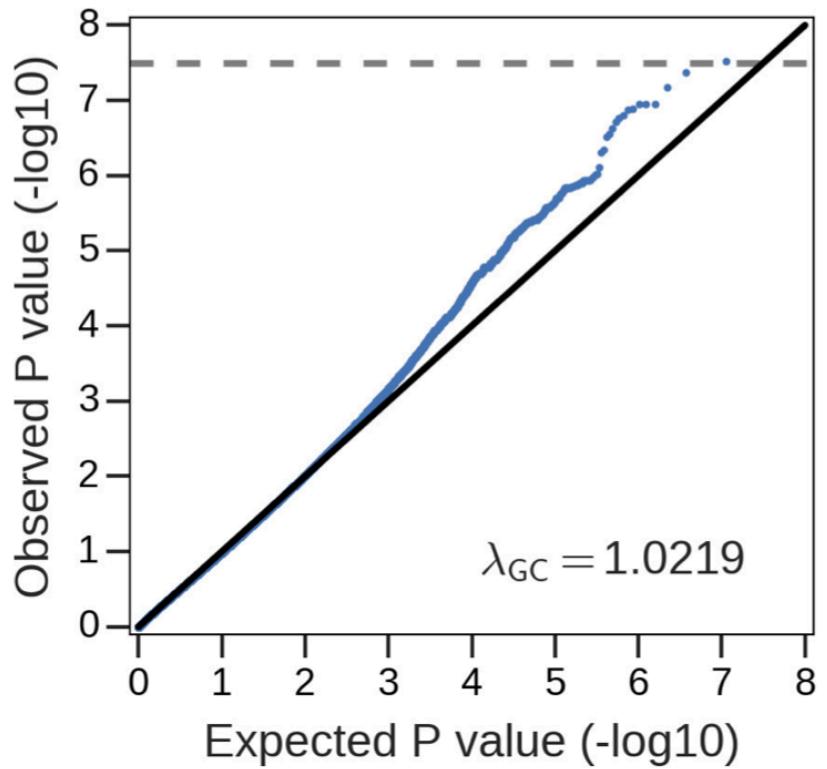
Highlights

- Host genetics influence abundances of health-associated gut bacteria
- Many heritable taxa co-occur
- The most heritable, Christensenellaceae, associates with a lean BMI
- Heritable taxa reduce weight gains in germ-free transplant experiments



b

- Re-analyze the 2,252 twins to quantify microbiome heritability.
- The sum of relative abundances of all 33 taxa reported as significantly heritable is 5.6%.
- **Their total heritability is only 1.9%.**



- Discovered only 2 SNPs that are significantly associated with beta-diversity
- Cannot replicate the 2 SNPs in other cohorts

Extended Data Figure 2 | Limited evidence for associations between microbiome β -diversity and specific SNPs. The quantile-quantile plot shows that only two SNPs are significantly associated with microbiome β -diversity at $P < 5 \times 10^{-8}$, computed using a distance-based F test with $n = 715$ unrelated genotyped individuals. λ_{GC} , genomic inflation factor.

Genome-wide association analysis identifies variation in vitamin D receptor and other host factors influencing the gut microbiota

Nature Genetics, 2016

Table 1 Summary of loci significantly associated with β diversity

Locus	SNP ID	Chr.	A1	A2	Locus start	Locus end	Nearest gene	Genes in locus	Effect size (%)
1	rs804427	1	A	C	33,538,964	33,623,510	AK2	ADC, TRIM62, AK2	0.79
2	rs1288616	1	G	A	53,885,577	53,965,248	DMRTB1	DMRTB1	0.76
3	rs1102737	1	G	A	172,700,868	172,779,833	FASLG		0.66
4	rs72853661	2	T	C	25,323,083	25,453,968	POMC	POMC, EFR3B	0.79
5	rs7567349	2	A	G	61,384,324	61,853,037	XPO1	AHSA2, USP34, XPO1, KIAA1841	0.76
6	rs2010917	2	T	C	135,172,338	135,197,891	MGAT5	MGAT5	0.74
7	rs71415332	2	G	A	102,309,520	102,616,128	—	IL1R2, MAP4K4	0.68
8	rs4670302	2	T	G	33,808,725	34,068,392	FAM98A	FAM98A	0.92
9	rs6711771	2	C	G	34,339,420	34,491,584	—	—	0.71
10	rs13099587	3	G	A	146,250,561	146,275,555	PLSCR1	PLSCR1	0.70
11	rs9647379	3	G	C	171,759,410	171,833,266	FNDC3B	FNDC3B	0.75
12	rs143050036	3	C	T	49,898,318	50,208,819	SEMA3F	RBM5, MST1R, CAMKV, MON1A, RBM6, SEMA3F	0.75
13	rs60500975	4	A	T	102,769,693	102,929,034	—	BANK1	0.82
14	rs62367773	5	A	G	74,171,398	74,220,999	FAM169A		0.67
15	rs1292672	6	C	T	87,217,958	87,509,434	HTR1E		0.70
16	rs35148810	7	C	T	151,515,842	151,530,983	—	PRKAG2	0.83
17	rs12705241	7	A	C	104,219,681	104,381,102	—	LHPL3	0.76
18	rs13260600	8	C	T	3,705,807	3,713,004	CSMD1	CSMD1	0.77
19	rs138022915	8	T	C	19,815,256	19,939,049	LPL	LPL	0.73
20	rs11986935	8	T	A	10,576,753	10,732,050	SOX7	SOX7, PINX1	0.97
21	rs7818750	8	G	A	135,273,640	135,299,611	ZFAT		0.74
22	rs1325919	9	C	T	37,626,956	37,650,386	FRMPD1		0.67
23	rs7082134	10	A	G	87,865,009	87,884,110	GRID1	GRID1	0.84
24	rs2251536	11	G	C	8,852,239	8,853,177	—	ST5	0.76
25	rs4472950	11	C	T	120,798,714	120,853,675	—	GRIK4	0.69
26	rs7974353	12	T	C	48,256,280	48,270,596	VDR	VDR	0.75
27	rs4760399	12	T	C	93,011,759	93,081,307	C12orf74		0.67
28	rs6573564	14	T	A	65,119,676	65,157,187	PLEKHG3		0.73
29	rs12910631	15	G	T	26,603,288	26,622,999	—		0.79
30	rs8040493	15	T	G	101,414,167	101,418,682	—		0.65
31	rs293377	15	G	C	89,623,490	89,635,268	ABHD2	ABHD2	0.70
32	rs8055365	16	T	C	84,566,729	84,581,275	KIAA1609	KIAA1609	0.70
33	rs59986499	16	G	A	3,065,924	3,097,940	CLDN6	MMP25, TNFRSF12A, CLDN6, CCDC64B, HCFC1R1, THOC6	0.68
34	rs12931878	16	A	G	11,031,741	11,207,817	CLEC16A	DEX1, CLEC16A	0.65
35	rs62085746	17	T	C	66,166,300	66,213,540	AMZ2		0.69
36	rs16969051	17	C	T	32,248,813	32,258,877	ACCN1	ACCN1	0.65
37	rs12601692	17	A	G	782,416	794,333	—	NXN	0.68
38	rs2267922	19	C	G	18,217,350	18,289,634	IFI30	MAST3, IFI30, PIK3R2	0.77
39	rs273647	19	C	G	51,739,767	51,766,748	C19orf75	CD33, C19orf75	0.84
40	rs4809760	20	A	G	48,428,863	48,591,125	SLC9A8	RNF114, SLC9A8, SPATA2	0.85
41	rs2835692	21	A	G	38,657,572	38,704,886	DSCR3		0.68
42	rs9917541	22	C	A	31,520,338	31,531,133	PLA2G3	PLA2G3, INPP5J	0.71

All loci have effect sizes greater than the significance threshold calculated from the null distribution ($P < 5 \times 10^{-8}$; Online Methods). The name of the lead SNP, chromosome, position, nearest gene and genes within a locus were determined using DEPICT software. A1 and A2 are the reference/alternative alleles based on the 1000 Genome Project. Chr., chromosome.

- 42 SNPs discovered to be *significantly* associated with beta-diversity, and they collectively explained 10% of the beta-diversity variance.
- Replicated none of the 42 SNPs... (no longer significant with permutation)
- If using the top 42 SNPs identified in the Israeli datasets, 12.1% could be explained...

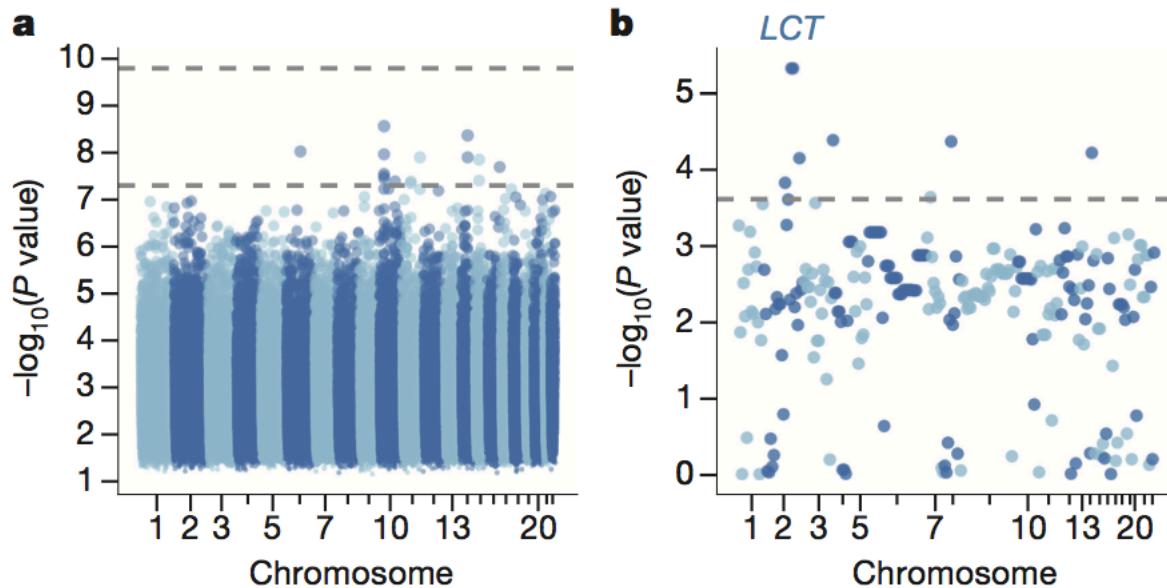


Figure 3 | Limited evidence for microbiome associations with specific SNPs. **a**, Manhattan plot showing the lowest P value obtained at every SNP tested for association with 313 taxa (computed by FaST-LMM³⁸ using $n = 814$ individuals) and with microbiome β -diversity (computed with a distance-based F test using $n = 715$ non-related individuals). The dashed lines represent a genome-wide significant P value corrected (top line) and not corrected (bottom line) for testing 313 taxa. **b**, The lowest P value obtained across 313 taxon association tests for each of 225 SNPs in 211 loci previously reported to be significantly associated with the microbiome^{5,6,8–10} (computed by FaST-LMM using $n = 814$ individuals). Seven SNPs are successfully replicated at $P < 0.05/211$ (dashed line; rs4988235, rs6730157, rs7656342, rs10112815, rs11626933, rs56006724 and rs7782745), two of which reside near the *LCT* gene.

- Only 7 of the 211 genes previously identified to be associated with specific taxa were replicated...

Tweet



Ruth Ley @RuthLeyMicro · 3/3/18

Replying to @segal_eran

Has anyone claimed Genetics as main determinant? More like, a determinant of a few members.

1 1 5 1



Eran Segal @segal_eran · 3/3/18

Agree, from our intro: "Previous studies have identified several heritable bacterial taxa". One paper (Wang et al., Nat. Gen. 2016) claimed that 42 SNPs explain 10% of the beta diversity. But AFAIK, the global effect of genetics on biome was not evaluated. We tried to answer that

1 1 4 1



Ruth Ley @RuthLeyMicro · 3/4/18

Ours and other calculated heritability at every branch in tree...as well as beta diversity and so on. That not global?

1 1 1 1



Eran Segal @segal_eran · 3/4/18

Tweet your reply

Tweet



Eran Segal @segal_eran · 3/4/18

Few points:

1. You asked if anyone claimed genetics as a main determinant. Goodrich (Cell 2014) title was "Human Genetics Shape the Gut Microbiome" and highlights included "Host genetics influence abundances of health-associated gut bacteria" and "Many heritable taxa co-occur"

2 1 5 1



Eran Segal @segal_eran · 3/4/18

2. Others indeed studied (a) association of b-diversity and top PCs/SNPs; or (b) heritability of individual taxa. Rather than trying to find which taxa are heritable, we wanted to arrive at a single number that quantifies the overall microbiome heritability

1 1 4 1



Eran Segal @segal_eran · 3/4/18

3. Our quantification uses not only the estimated heritability of each taxon, but

Tweet your reply

Tweet



Eran Segal @segal_eran · 3/4/18

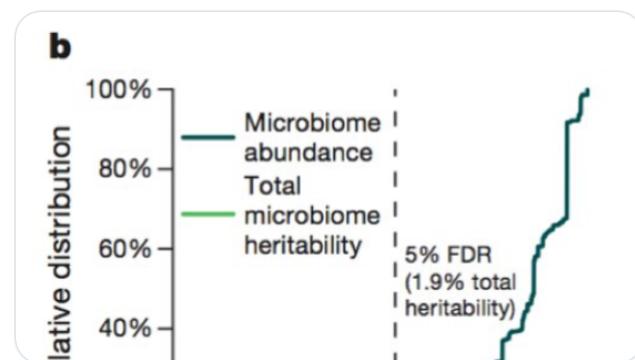
3. Our quantification uses not only the estimated heritability of each taxon, but also relative abundance. Thus, rare taxa do not contribute much to the overall heritability, even if they're very heritable

2 1 3 1



Eran Segal @segal_eran · 3/4/18

Point 3 is the main difference between our approach and yours in our re-analysis of your twins data (which was instrumental in our study, btw). This was our Fig. 2b, attached here. Accounting for abundance, we estimate overall 2% heritability based on your data+analysis+pvalues



Tweet your reply





Extended Data Figure 4 | The gut microbiome is significantly associated with multiple environmental factors. The fraction of variance of the microbiome β -diversity matrix that can be inferred from different categories of environmental factors is shown. $n = 715$ individuals (Supplementary Table 17); numbers in parentheses indicate the number of

features in each category. The fraction of inferred variance can reflect both the information that the category conveys on the microbiome as well as the number of factors in the category, which depends on the questionnaire used in the study.

b

Phenotype	Microbiome association index		Genetic heritability (literature)
	Israeli cohort	LLD cohort	
HDL	35.9%***	27.9%***	23.9%–48%
Lactose cons.	35.5%***	N/A	N/A
Waist circ.	28.8%***	26%***	15%–24%
Hip circ.	27.1%***	28%***	10.6%–27%
Glycaemic status	24.5%***	N/A	N/A
BMI	24.5%***	27.8%***	14%–32%
WHR	23.9%***	6.9%*	12%–14%
Fasting glucose	21.9%***	8%**	9%–33%
HbA1c%	16.1%*	8.4%	21%–32%
Creatinine	12.3%*	6.7%	19%–25%
Height	3.2%	25.9%***	33%–68%
Total cholesterol	0%	13.5%	14%–53%