iScience, Volume *26*

Supplemental information

New insights from GWAS on BMI-related growth traits in a longitudinal cohort of admixed children with Native American and European ancestry

Lucas Vicuña, Esteban Barrientos, Tomás Norambuena, Danilo Alvares, Juan Cristobal Gana, Valeria Leiva-Yamaguchi, Cristian Meza, José L. Santos, Veronica Mericq, Ana Pereira, and Susana Eyheramendy

Supplemental Information

Supplemental Figures

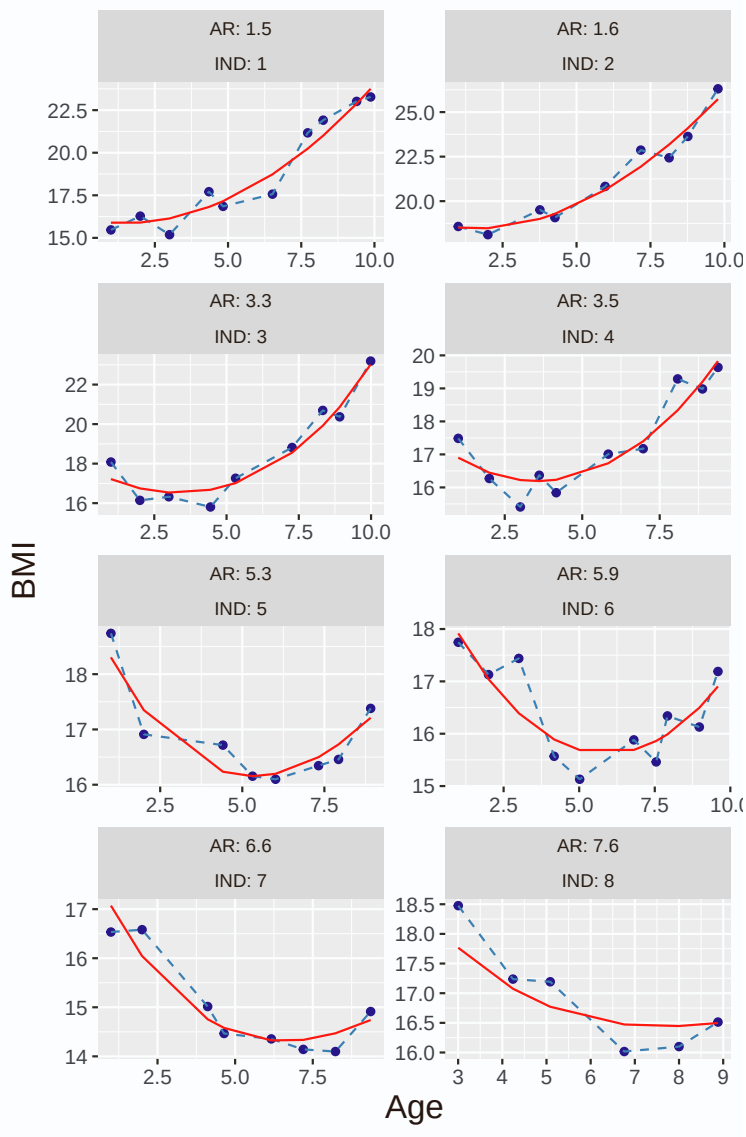


Figure S1. **Individual estimations of Age-AR and BMI-AR, Related to Figure 1.** Shown are examples of BMI trajectories for individuals with Age-AR estimations between 1.5 and 7.6 years old. Blue dotted lines represent the real data, whereas red lines represent the fit curves. Individual (IND) codes do not represent the original codes due to ethical reasons. Age-AR: Age at adiposity rebound.

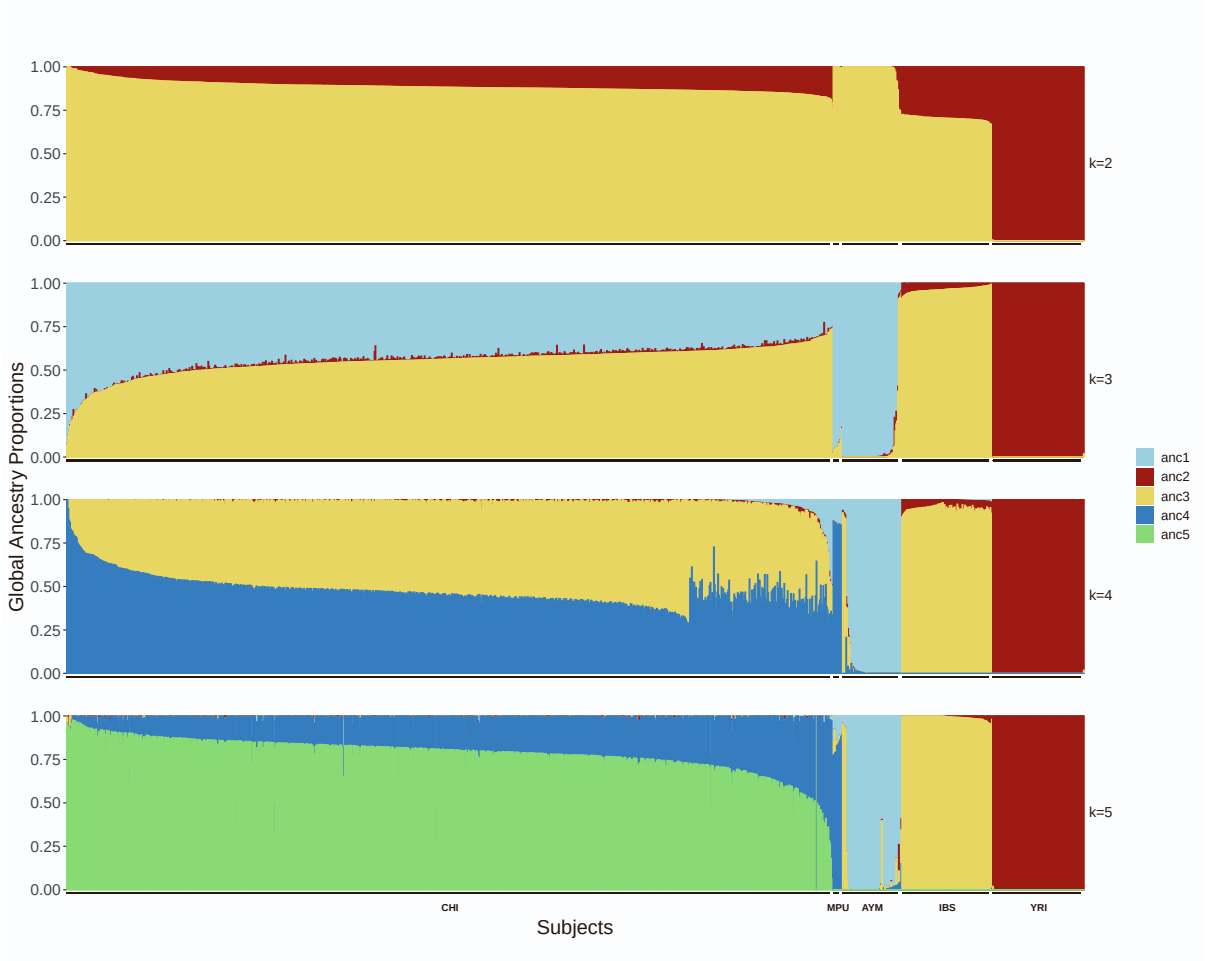


Figure S2. **Global ancestry proportions of Chilean children, Related to STAR Methods.** The number of displayed clusters is K=2 − 5. YRI: Yoruba; AYM: Aymara; CHL: Chilean; IBS: Spaniards; MPU: Mapuche.

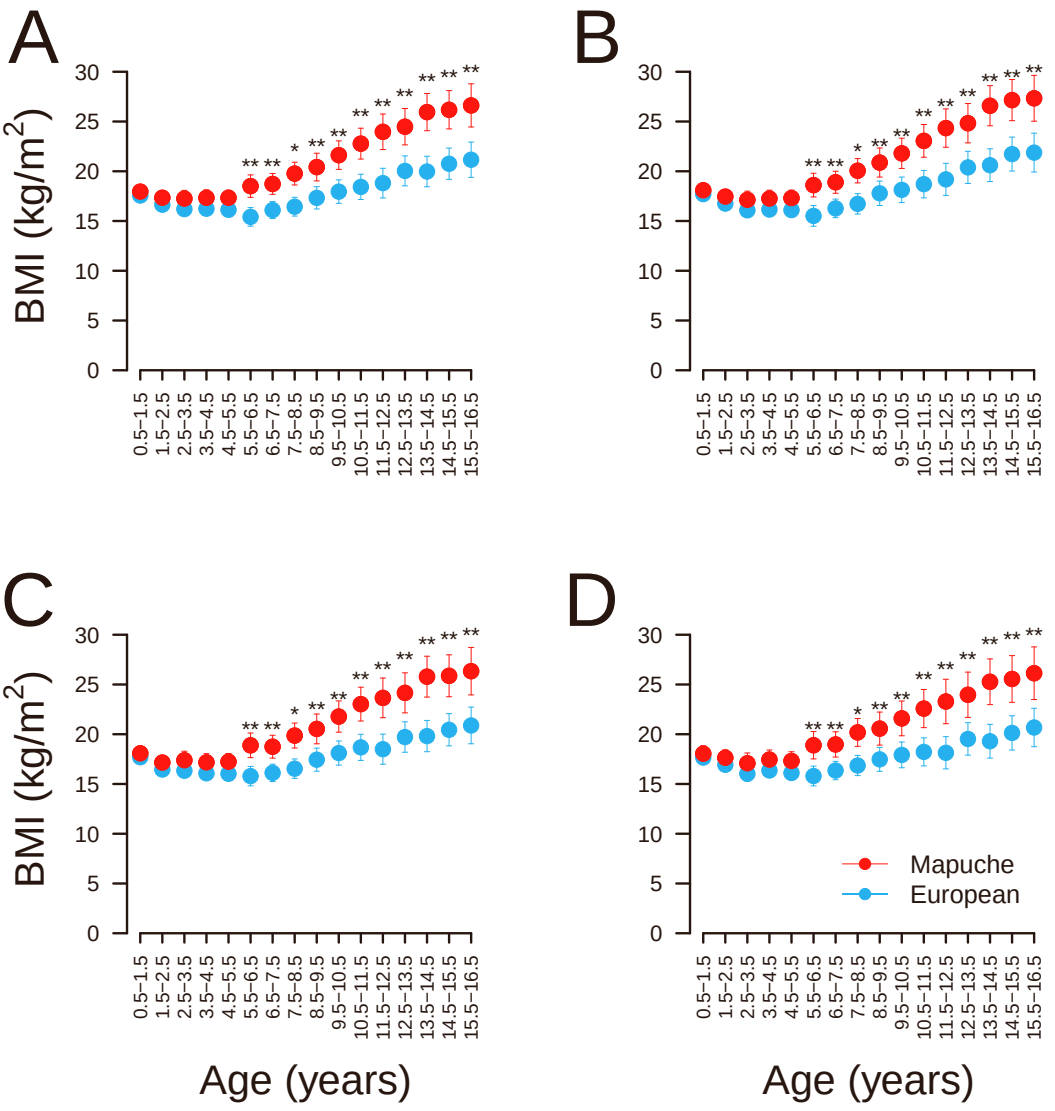


Figure S3. **Effect of global Mapuche and European ancestry on BMI trajectory in individuals grouped by MEL categories, Related to Figure 2A.** BMI (kg/m^2) estimated at each age stratum for predicted individuals with 100% Mapuche (red dots) and 100% European global ancestry (blue dots), grouped by MEL categories. The 4 plots represent the 4 main groups of MEL described in the Methods, namely, Group 1 (A); Group 2 (B); Group 3 (C) and Group 4 (D). Standard error bars are included. $^*P< 0.05$; $^{**}P< 0.01$.

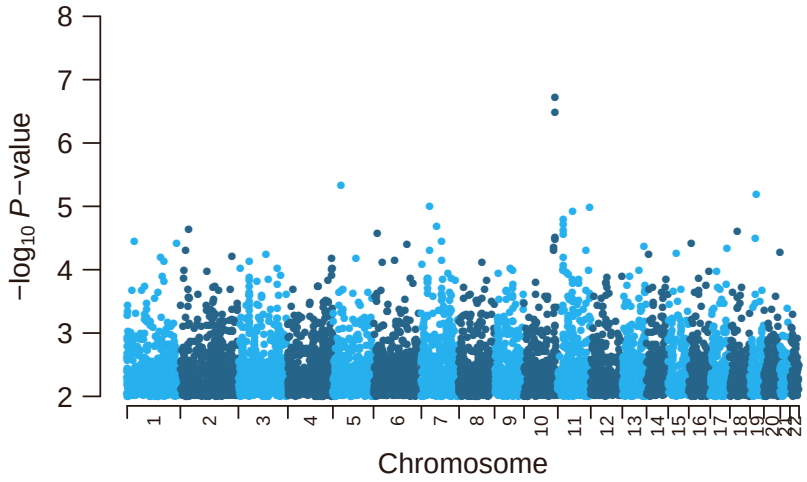


Figure S4. **GWAS of BMI in children aged 0.5-1.5 years old, Related to Figure 3.** Manhattan plot showing genome-wide per-SNP association P-values represented as -log10 P-value, along the 22 autosomes. Variants with log10(P-value) between 0 and 2 are not shown.

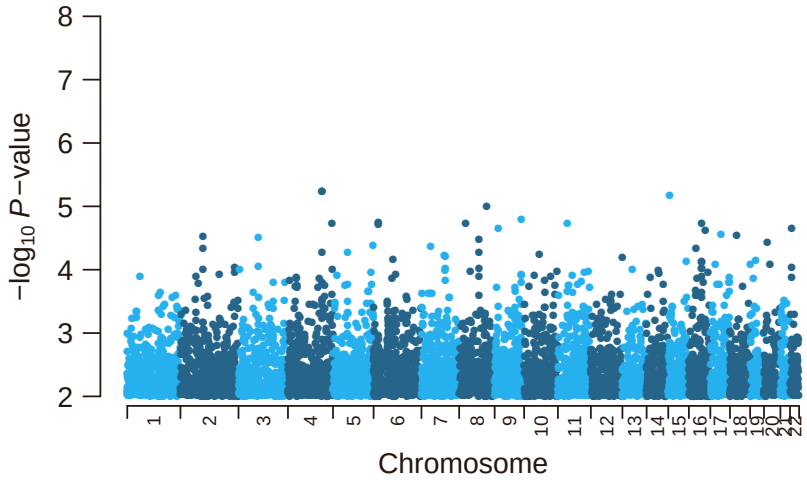


Figure S5. **GWAS of BMI in children aged 2.5-3.5 years old, Related to Figure 3.** Manhattan plot showing genome-wide per-SNP association P-values represented as -log10 P-value, along the 22 autosomes. Variants with log10(P-value) between 0 and 2 are not shown.

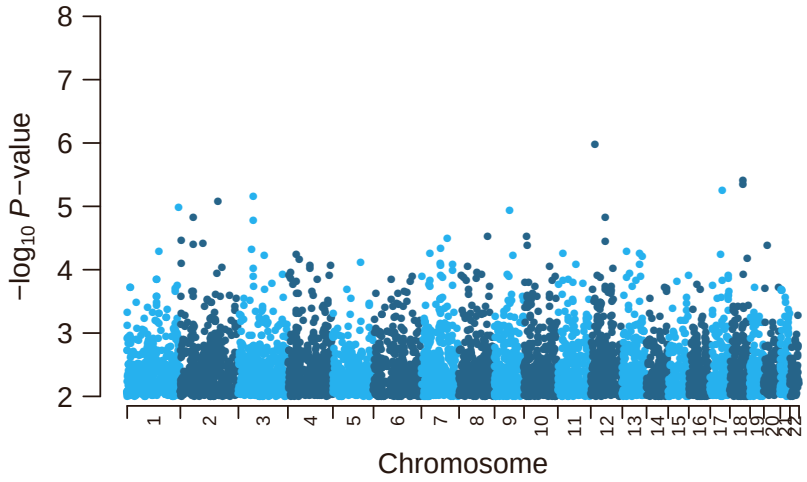


Figure S6. **GWAS of BMI in children aged 3.5-4.5 years old, Related to Figure 3.** Manhattan plot showing genome-wide per-SNP association P-values represented as -log10 P-value, along the 22 autosomes. Variants with log10(P-value) between 0 and 2 are not shown.

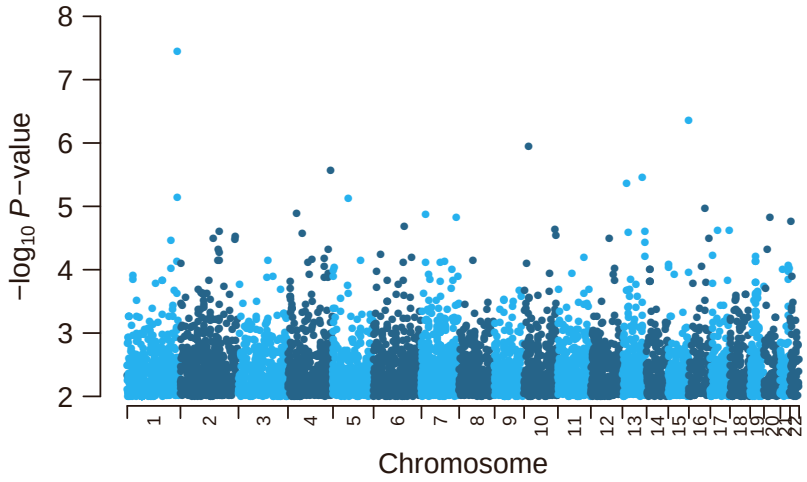


Figure S7. **GWAS of BMI in children aged 4.5-5.5 years old, Related to Figure 3.** Manhattan plot showing genome-wide per-SNP association P-values represented as -log10 P-value, along the 22 autosomes. Variants with log10(P-value) between 0 and 2 are not shown.

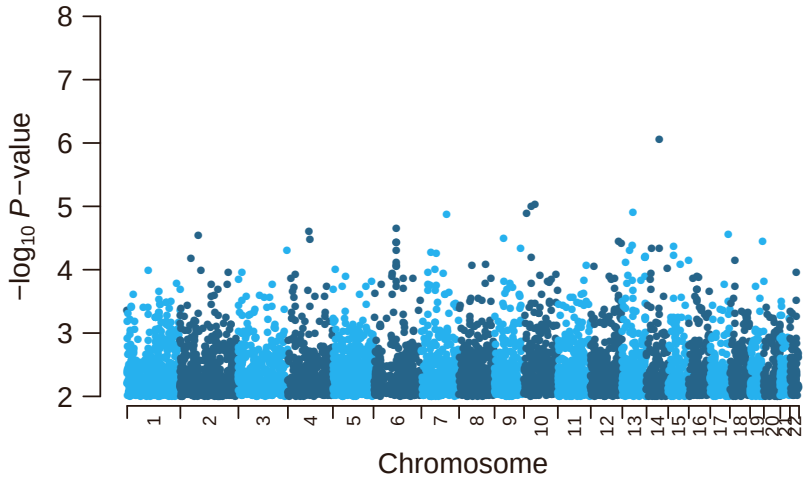


Figure S8. **GWAS of BMI in children aged 5.5-6.5 years old, Related to Figure 3.** Manhattan plot showing genome-wide per-SNP association P-values represented as -log10 P-value, along the 22 autosomes. Variants with log10(P-value) between 0 and 2 are not shown.

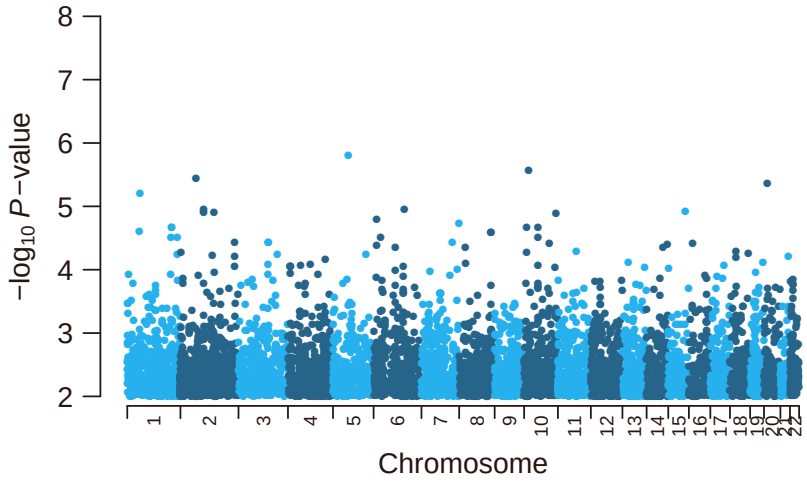


Figure S9. **GWAS of BMI in children aged 6.5-7.5 years old, Related to Figure 3.** Manhattan plot showing genome-wide per-SNP association P-values represented as -log10 P-value, along the 22 autosomes. Variants with log10(P-value) between 0 and 2 are not shown.

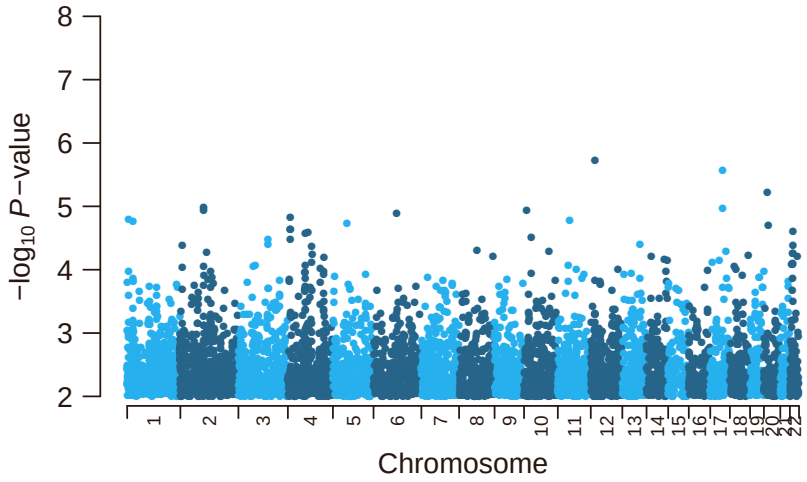


Figure S10. **GWAS of BMI in children aged 7.5-8.5 years old, Related to Figure 3.** Manhattan plot showing genome-wide per-SNP association P-values represented as -log10 P-value, along the 22 autosomes. Variants with log10(P-value) between 0 and 2 are not shown.

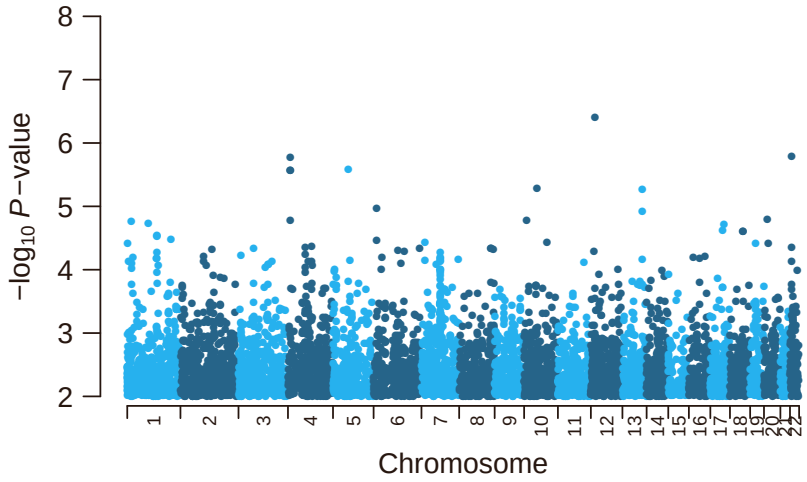


Figure S11. **GWAS of BMI in children aged 8.5-9.5 years old, Related to Figure 3.** Manhattan plot showing genome-wide per-SNP association P-values represented as -log10 P-value, along the 22 autosomes. Variants with log10(P-value) between 0 and 2 are not shown.

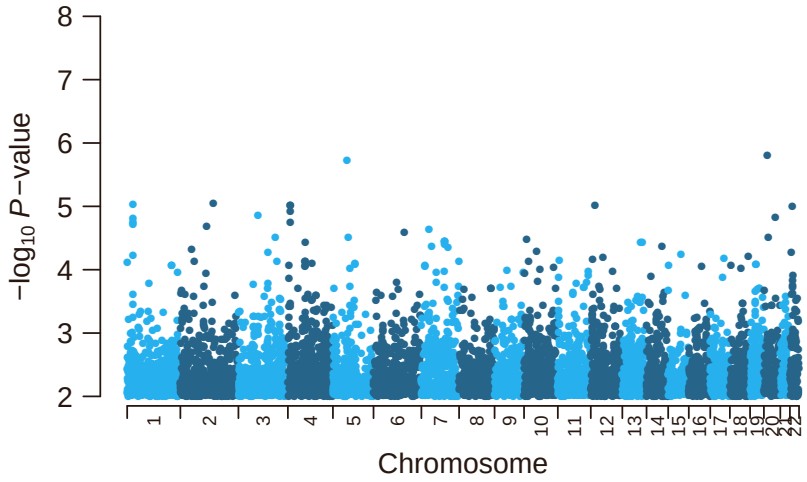


Figure S12. **GWAS of BMI in children aged 9.5-10.5 years old, Related to Figure 3.** Manhattan plot showing genome-wide per-SNP association P-values represented as -log10 P-value, along the 22 autosomes. Variants with log10(P-value) between 0 and 2 are not shown.

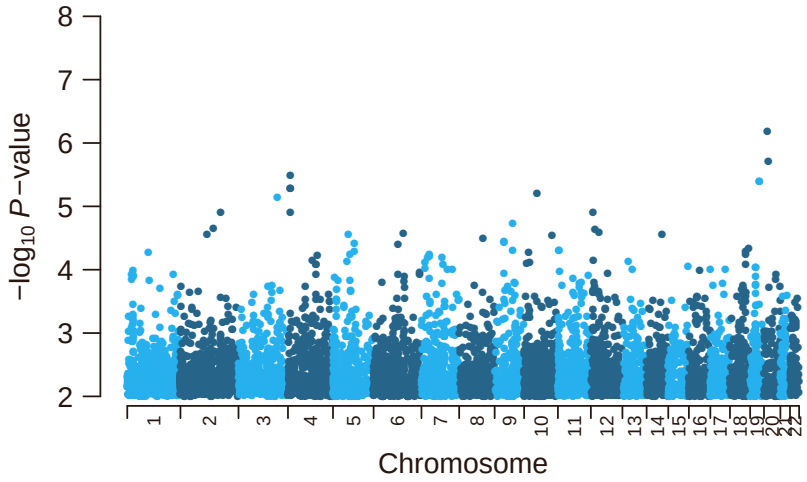


Figure S13. **GWAS of BMI in children aged 10.5-11.5 years old, Related to Figure 3.** Manhattan plot showing genome-wide per-SNP association P-values represented as -log10 P-value, along the 22 autosomes. Variants with log10(P-value) between 0 and 2 are not shown.

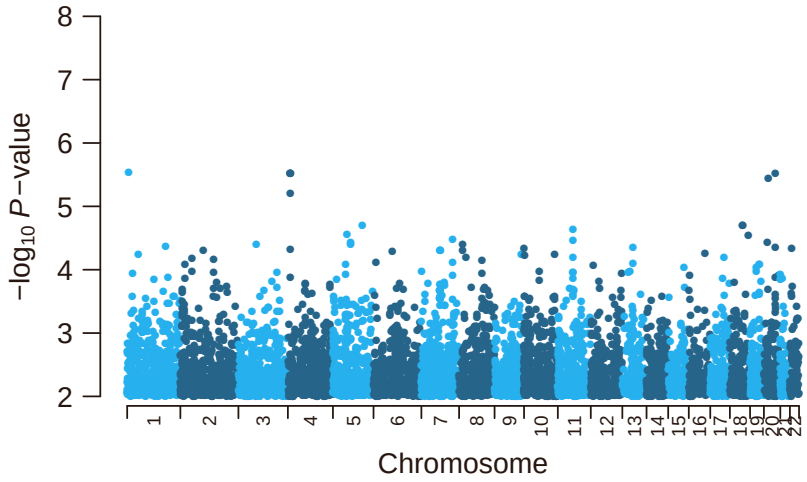


Figure S14. **GWAS of BMI in children aged 11.5-12.5 years old, Related to Figure 3.** Manhattan plot showing genome-wide per-SNP association P-values represented as -log10 P-value, along the 22 autosomes. Variants with log10(P-value) between 0 and 2 are not shown.

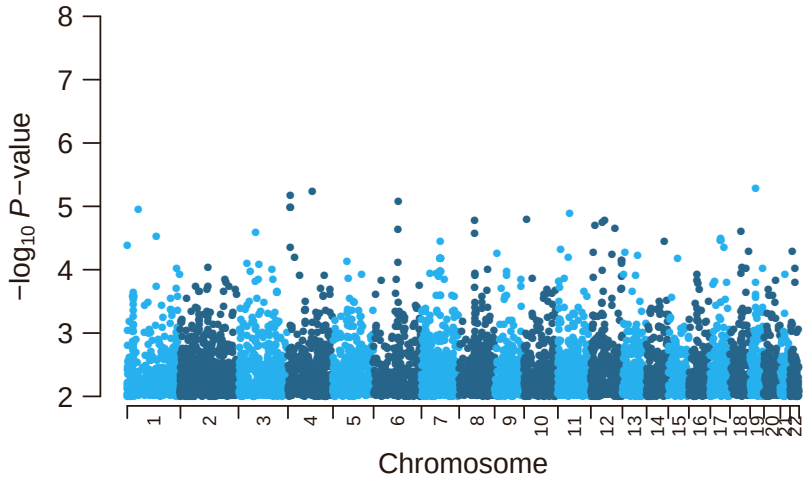


Figure S15. **GWAS of BMI in children aged 12.5-13.5 years old, Related to Figure 3.** Manhattan plot showing genome-wide per-SNP association P-values represented as -log10 P-value, along the 22 autosomes. Variants with log10(P-value) between 0 and 2 are not shown.

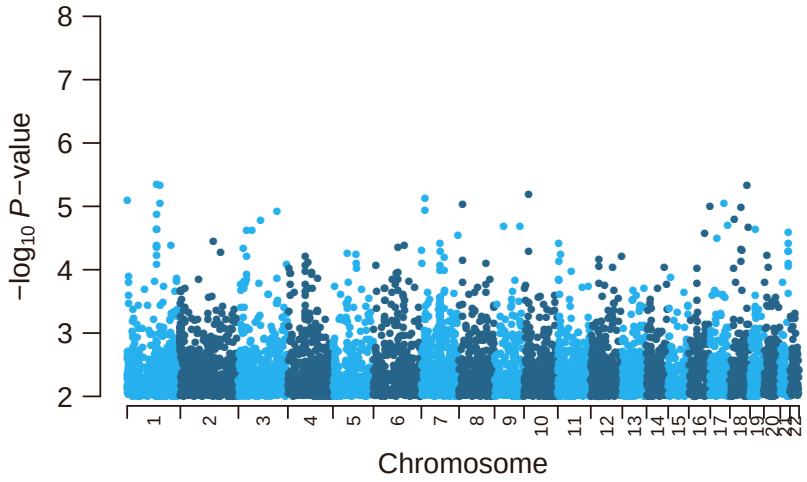


Figure S16. **GWAS of BMI in children aged 13.5-14.5 years old, Related to Figure 3.** Manhattan plot showing genome-wide per-SNP association P-values represented as -log10 P-value, along the 22 autosomes. Variants with log10(P-value) between 0 and 2 are not shown.

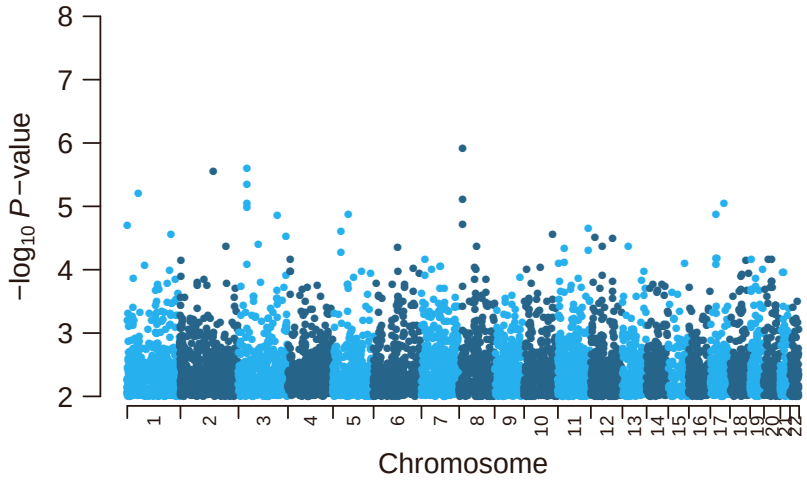


Figure S17. **GWAS of BMI in children aged 14.5-15.5 years old, Related to Figure 3.** Manhattan plot showing genome-wide per-SNP association P-values represented as -log10 P-value, along the 22 autosomes. Variants with log10(P-value) between 0 and 2 are not shown.

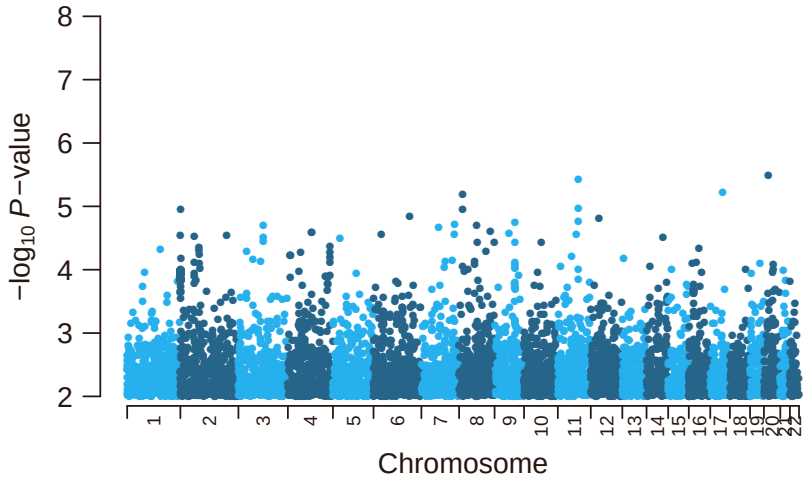


Figure S18. **GWAS of BMI in children aged 15.5-16.5 years old, Related to Figure 3.** Manhattan plot showing genome-wide per-SNP association P-values represented as -log10 P-value, along the 22 autosomes. Variants with log10(P-value) between 0 and 2 are not shown.

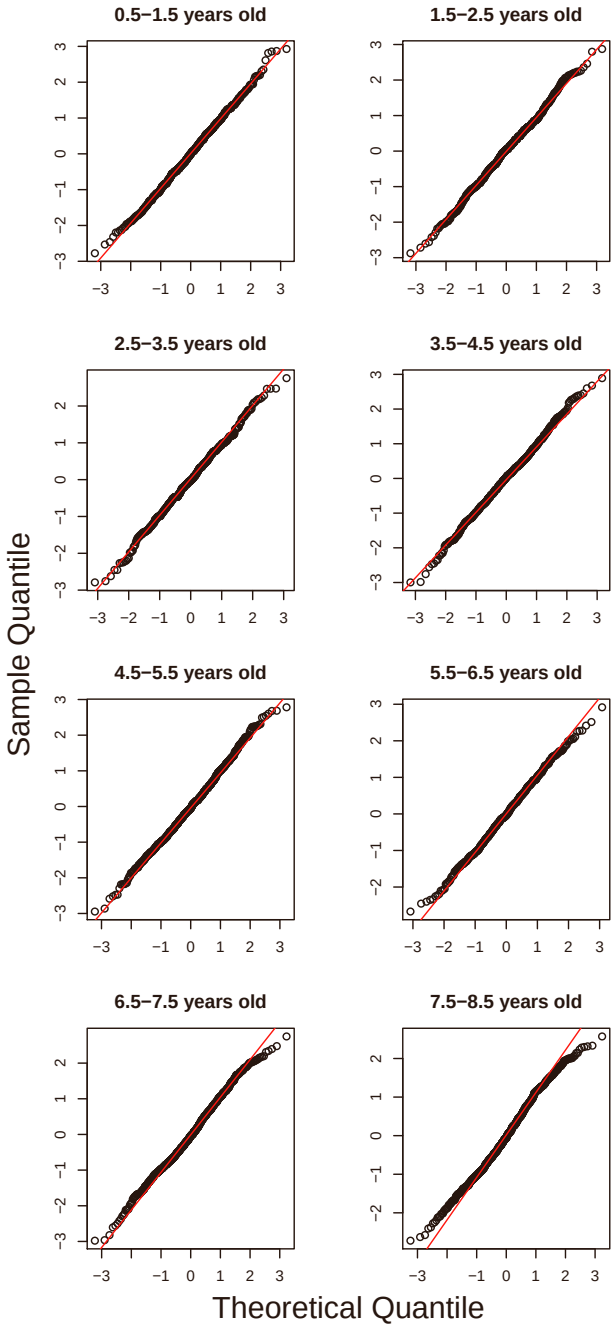


Figure S19. **QQ-plots of cross-sectional GWAS across 0.5-1.5 years old to 7.5-8.5 years old strata, Related to Figure 3.** The ordinate and abscissa represent the sample quantiles and the theoretical quantiles, respectively.

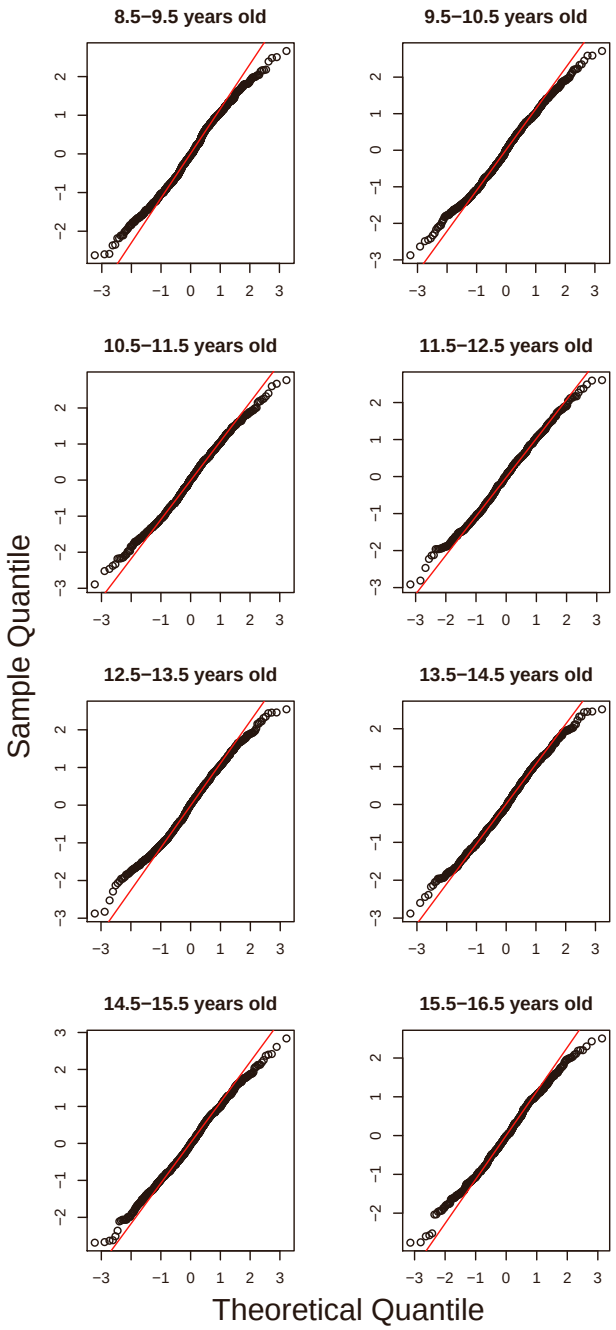


Figure S20. **QQ-plots of cross-sectional GWAS across 8.5-9.5 years old to 15.5-16.5 years old strata, Related to Figure 3.** The ordinate and abscissa represent the sample quantiles and the theoretical quantiles, respectively.

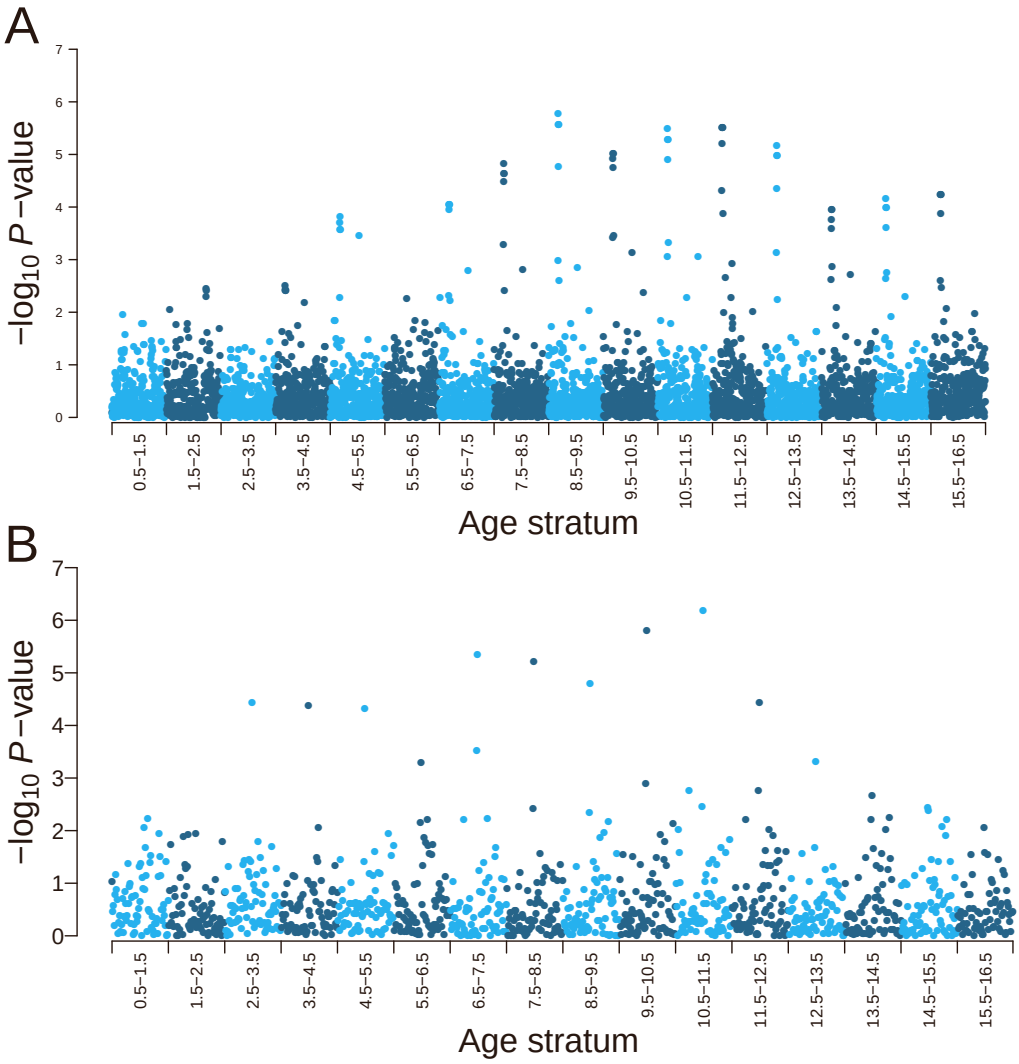


Figure S21. **Trajectories of association peaks from chromosomes** 4 **and** 20 **along age strata, Related to Figure 3.** Manhattan plots showing per-SNP association *P*-values represented as -log10 *P*-value. A: Chromosome 4. The strongest association corresponds to the *rs12501266* variant of gene *SORCS2*. Shown is the region with physical coordinates 4:7440183-7310183. B: Chromosome 20. The strongest association corresponds to the *rs474169* variants of gene *SNAP25-AS1*. The region shown has physical coordinates 20:10016800-10181800.

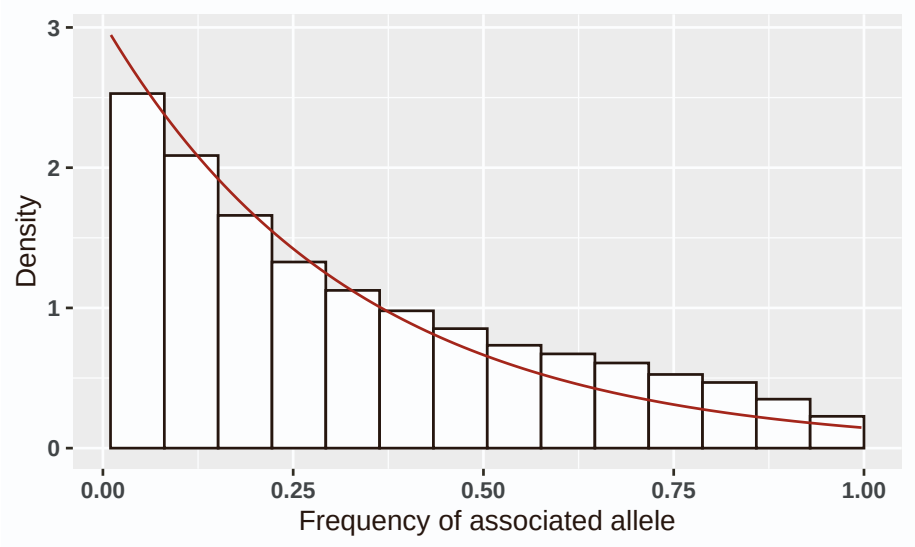


Figure S22. **Distribution of allele frequencies for the GWAS on age stratum 1.5 - 2.5 years old, Related to STAR Methods.** The histogram shows the real distribution of the allele frequencies of all SNPs. The red curve represents the exponential model fit with parameter $\lambda = 3.041907$. The truncated exponential distribution from which the allele frequencies were generated in the simulations, slightly under-represent frequencies between 0.5 and 0.9. We are confident that this is likely not affecting the simulations for the following reasons: i) most of the observed allele frequencies are below 0.5 (roughly 75%); ii) small P-values are more likely to be generated from extreme allele frequencies (very small or very large), and we are excluding frequencies below 0.01%; iii) frequencies at the same distance from 0.5 (to the right or left) generate the same P-values.

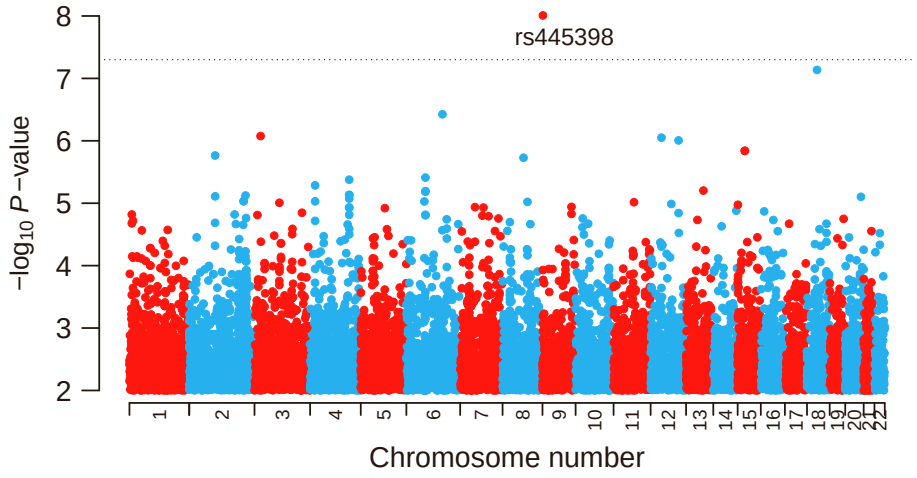


Figure S23. **GWAS of Age-AR, Related to Table 2.** Manhattan plot showing genome-wide association P-values represented as -log10 P-value. Variants achieving the genome wide threshold of $P < 5 \text{ x } 10^{-8}$ are labeled. Variants with log10(P-value) between 0 and 2 are not shown. The black dotted line represents the genome wide significance threshold.

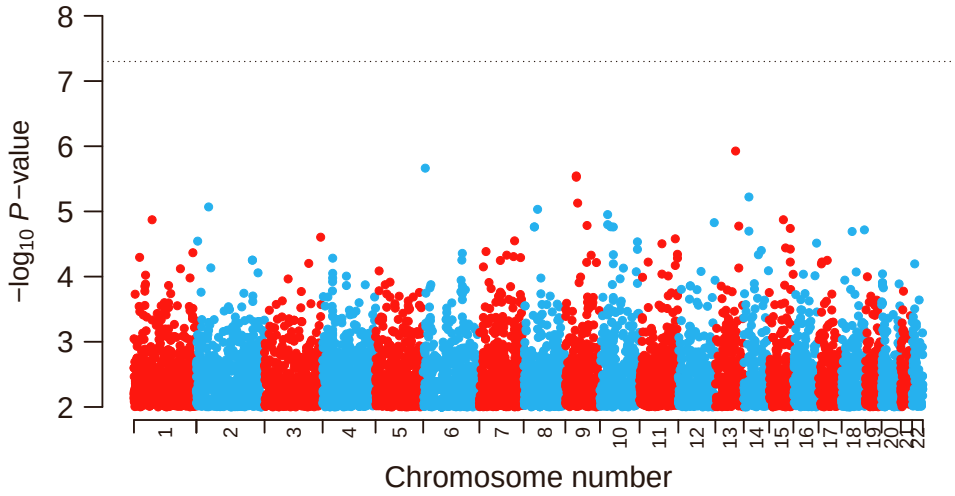
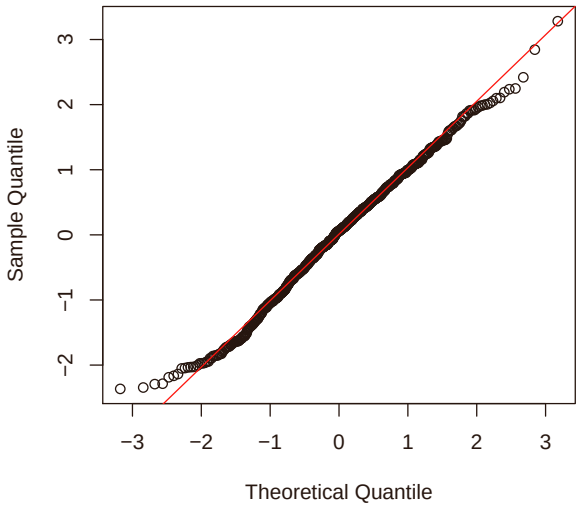


Figure S24. **GWAS of BMI-AR, Related to Table 2.** Manhattan plot showing genome-wide association P-values represented as -log10 P-value. Variants with log10(P-value) between 0 and 2 are not shown. The black dotted line represents the genome wide significance threshold.



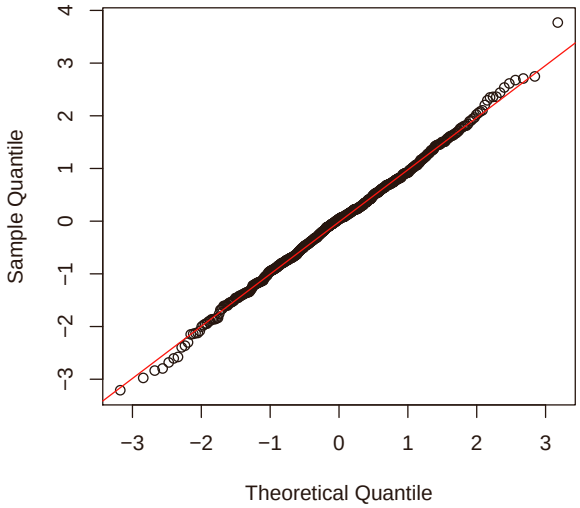


Figure S25. **QQ-plots of Age-AR and BMI-AR GWAS, Related to Table 2.** The top panel shows the QQ-plot for the GWAS on Age-AR, whereas the bottom panel shows the QQ-plot for the GWAS on BMI-AR GWAS. The ordinate and abscissa represent the sample quantiles and the theoretical quantiles, respectively.

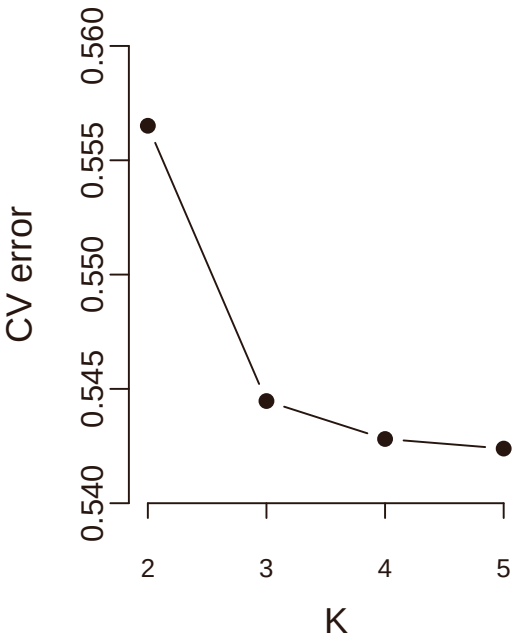


Figure S26. **Cross-validation error across K ancestral populations, Related to STAR Methods.** Cross-validation errors estimated by ADMIXTURE, using K= $2-5$ ancestral populations.

Supplemental Tables

Maternal Educational Level	Maternal Educational Level (Chilean education system; in Spanish)	Number of individuals
Incomplete middle school	Básica incompleta	59
Complete middle school	Básica completa	465
Incomplete high school	Media incompleta	137
Complete high school	Media completa	2
Incomplete technical education	Instituto profesional incompleto	46
Complete technical education	Instituto profesional completo	101
Incomplete university education	Universitaria incompleta	9
Complete university education	Universitaria completa	62
Graduate studies	Postgrado	2
Special education	Educación especial	1
No studies	Sin estudios	4
Does not know or does not reply	No sabe o no responde	9
No classification	NA	6
Other	Otro	1

Table S1. **Maternal education level, Related to Figure 2A.** Number of individuals in each education category. The name of each category was translated into Spanish to better reflect the nomenclature of the Chilean education system.

Age stratum	Sample size	Lambda
0.5-1.5	718	-1.0
1.5-2.5	684	-1.3
2.5-3.5	504	-1.6
3.5-4.5	642	-2.0
4.5-5.5	788	-2.0
5.5-6.5	483	-1.8
6.5-7.5	803	-1.7
7.5-8.5	829	-1.4
8.5-9.5	828	-1.0
9.5-10.5	841	-0.6
10.5-11.5	795	-0.5
11.5-12.5	667	-0.7
12.5-13.5	748	-0.7
13.5-14.5	782	-1.1
14.5-15.5	780	-1.1
15.5-16.5	587	-1.1

Table S2. **BMI estimation by age stratum in Mapuche and European individuals, Related to Figure 2A.** Shown are the age stratum; number of individuals; and λ used in Box-Cox transformations for each age stratum.

Significance	Observed	Simulated 95% CI [L, U]
P < 1E-1	115595	115893 [115691, 116096]
P < 1E-2	29334	28349 [28261, 28437]
P < 1E-3	4640	4167 [4134, 4199]
P < 1E-4	635	525 [508, 542]
P < 1E-5	62	60 [53, 68]
P < 1E-6	11	7 [5, 8]
P < 1E-7	0	0 [0, 1]
P < 1E-8	0	0 [0, 0]

Table S3. **Simulations results of cross-sectional GWAS, Related to Table 1.** SNP counts for the observed and expected (simulated) associated SNPs across different P-value ranges at any age stratum. Shown are the lower (L) and upper (U) limits of the 95% confidence intervals for the simulated data under the null hypothesis of no association.

SNP ID-Allele	Age stratum	Sample size	Effect size	Statistical power
rs269511-G	(1.5,2.5]	684	-0.272	0.655
rs9275582-T	(1.5,2.5]	684	0.297	0.662
rs9275593-A	(1.5,2.5]	684	0.297	0.662
rs9275595-C	(1.5,2.5]	684	0.297	0.662
rs7134291-A	(8.5,9.5]	828	-0.332	0.831
rs7896870-C	(0.5,1.5]	718	0.250	0.843
rs474169-T	(10.5,11.5]	795	-0.285	0.755
rs1495271-T	(4.5,5.5]	788	0.256	0.476
rs11244839-A	(0.5,1.5]	718	0.248	0.824
rs13257360-A	(14.5,15.5]	780	0.280	0.992

Table S4. **Statistical power of associations from cross-sectional GWAS, Related to Table 1.** Shown are the SNP ID and their associated alleles, age stratum, sample size, effect size and the statistical power. The calculations are described in the Methods.

SNP ID-Allele	Location	Consequence	Gene	ß-GT	*P*-GT	Frequency
rs4655426-T	1:215893063	intron	*USH2A*	-4.93	9.8E-07	0.162
rs79316274-C	13:57529593	intergenic	-	4.74	2.4E-06	0.013
rs75032037-T	13:57426033	intergenic	-	4.72	2.8E-06	0.014
rs11545026-T	15:57842525	3'UTR	*CGNL1*	4.66	3.6E-06	0.014
rs1058511-C	19:55859439	3'UTR	*SUV420H2*	-4.58	5.4E-06	0.470
rs6770886-G	3:152675646	intergenic	-	-4.57	5.7E-06	0.289
rs45520932-T	11:115103597	intron	*CADM1*	4.48	8.3E-06	0.162
rs1403501-C	3:152674921	intergenic	-	-4.48	8.4E-06	0.287
rs79523360-A	4:73424464	intron	*ADAMTS3*	4.41	1.2E-05	0.017
rs7822543-C	8:65186350	intron/NCT	*RP11-32K4.1*	4.41	1.2E-05	0.014

Table S5. **Strongest associations in the GWAS for longitudinal BMI, Related to STAR Methods.** Shown is the SNP rs ID with the associated allele, physical position in the chromosome, SO consequence type, gene, effect size of the genotype (β_{GT}) with the corresponding association P-value (P_{GT}), and the frequency of the associated allele. NCT: noncoding transcript variant.

SNP.ID-Allele	Location	Consequence	Gene	ß-GTxSex	*P*-GTxSex	Frequency
rs2183606-A	13:92804506	intron	*GPC5*	-4.90	1.2E-06	0.453
rs6926791-A	6:5169584	intron	*LYRM4*	4.78	2.2E-06	0.040
rs10812580-A	9:27383756	intron	*MOB3B*	4.72	2.9E-06	0.259
rs13283804-G	9:27391922	intron	*MOB3B*	4.71	3.0E-06	0.284
rs76077232-C	14:34930273	intron	*SPTSSA*	4.56	6.0E-06	0.013
rs141739085-C	9:32332364	downstream	*SLC25A5P8*	4.52	7.4E-06	0.137
rs1464005-G	2:34238100	intron/NCT	*AC009499.1*	-4.49	8.5E-06	0.346
rs73230790-A	8:31052352	intergenic	-	-4.47	9.3E-06	0.220
rs10734038-T	10:18436515	intron	*CACNB2*	-4.43	1.1E-05	0.436
rs72731396-T	15:72321999	intron	*MYO9A*	-4.39	1.3E-05	0.236

Table S6. **Strongest associations for BMI-AR GWAS, Related to STAR Methods.** Shown is the SNP rs ID with the associated allele, physical position in the chromosome, SO consequence type, gene, effect size of the interaction between genotype and female sex (β_{GTxSex}) with the corresponding association P-value (P_{GTxSex}), and the frequency of the associated allele. NCT: noncoding transcript variant.