ARTICLES
https://doi.org/10.1038/s42255-022-00549-1

nature metabolism

Characterization of the genetic architecture of infant and early childhood body mass index

Øyvind Helgeland[1,2,15], Marc Vaudel[1,15], Pol Sole-Navais[3], Christopher Flatley[3], Julius Juodakis[3], Jonas Bacelis[3], Ingvild L. Koløen[1,4], Gun Peggy Knudsen[5], Bente B. Johansson[1], Per Magnus[6], Ted Reichborn Kjennerud[7,8], Petur B. Juliusson[9,10,11], Camilla Stoltenberg[5], Oddgeir L. Holmen[12], Ole A. Andreassen[13,14], Bo Jacobsson[2,3], Pål R. Njølstad[1,11] ✉ and Stefan Johansson[1,4] ✉

Early childhood obesity is a growing global concern; however, the role of common genetic variation on infant and child weight development is unclear. Here, we identify 46 loci associated with early childhood body mass index at specific ages, matching different child growth phases, and representing four major trajectory patterns. We perform genome-wide association studies across 12 time points from birth to 8 years in 28,681 children and their parents (27,088 mothers and 26,239 fathers) in the Norwegian Mother, Father and Child Cohort Study. Monogenic obesity genes are overrepresented near identified loci, and several complex association signals near *LEPR*, *GLP1R*, *PCSK1* and *KLF14* point towards a major influence for common variation affecting the leptin–melanocortin system in early life, providing a link to putative treatment strategies. We also demonstrate how different polygenic risk scores transition from birth to adult profiles through early child growth. In conclusion, our results offer a fine-grained characterization of a changing genetic landscape sustaining early childhood growth.

Physical growth is an indicator and predictor of both present and future health. Deviations from a child's growth trajectory may indicate health issues with lifelong implications. Growth in infancy and early childhood is thus monitored closely by parents and health care professionals. Body mass index (BMI) changes substantially with age following a characteristic pattern. From birth, BMI increases rapidly until it reaches a maximum at the age of 9 months, followed by a gradual decline towards a minimum at around 5–6 years of age. These two points are often labelled the adiposity peak (AP) and adiposity rebound (AR)[1,2], respectively. Early increase in BMI is associated with diabetes, earlier puberty, risk of obesity in adolescence and adulthood, a major public health issue worldwide[3–5], and the many complications that follow. Only 38% of adults with class II/III obesity (BMI ≥ 35 kg/m²) have normal weight during childhood[6], and 90% of all children defined as obese at age 3 years remain obese during adolescence[7]. As sustainable weight reduction has proved difficult[8], proactive therapeutic strategies enabling early prevention of obesity are sorely needed; thus, a better understanding of the fundamental mechanisms regulating early growth is needed.

Heritability estimates for BMI in twin studies range from 40% to 70% and vary with age[9,10]. Genetic variants strongly influence the risk of obesity, in a complex relationship with behavioural and lifestyle factors[11]. Common genetic variants explain 17% to 27% of the heritability of BMI[12–14]. The genetics of early weight development is therefore of prime scientific interest for children's health, but also as a predictor for adult obesity. The largest genome-wide association studies (GWAS) on adult BMI identified 941 independent loci in over 700,000 individuals, explaining ~6% of the phenotypic variation[15]. In children, where sample sizes have been much smaller, considerably less is known about the genetics of BMI. Recent meta-analyses suggest an overlap with adult BMI[16–18], while studies estimating age-dependent genetic contribution have revealed low correlation in infancy and early childhood that gradually increases with age[12]. Additionally, transient genetic association with early BMI during infancy and early childhood has been identified by us and others[19,20], suggestive of rapid changes in the genetic architecture of BMI during early growth. Still, how the genetics of BMI transitions from birth to AR, where the genetic signature of an adult-like obesity emerges, remains unknown.

While GWAS performed in very large numbers of adults have been highly successful in discovering common variants of small effect sizes, studies on children with morbid obesity have been more successful at identifying rare genetic variants causing early-onset monogenic and syndromic forms of obesity[21,22]. Recently, there has been a growing recognition that monogenic and polygenic forms of obesity are not discrete entities. Genetic studies point towards shared biological pathways and the influence of both rare

[1]Center for Diabetes Research, Department of Clinical Science, University of Bergen, Bergen, Norway. [2]Department of Genetics and Bioinformatics, Health Data and Digitalization, Norwegian Institute of Public Health, Oslo, Norway. [3]Department of Obstetrics and Gynecology, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden. [4]Department of Medical Genetics, Haukeland University Hospital, Bergen, Norway. [5]Norwegian Institute of Public Health, Oslo, Norway. [6]Centre for Fertility and Health, Norwegian Institute of Public Health, Oslo, Norway. [7]Department of Mental Disorders, Norwegian Institute of Public Health, Oslo, Norway. [8]Institute of Clinical Medicine, University of Oslo, Oslo, Norway. [9]Department of Health Registry Research and Development, National Institute of Public Health, Bergen, Norway. [10]Department of Clinical Science, University of Bergen, Bergen, Norway. [11]Children and Youth Clinic, Haukeland University Hospital, Bergen, Norway. [12]HUNT Research Centre, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway. [13]NORMENT Centre, Institute of Clinical Medicine, University of Oslo, Oslo, Norway. [14]Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway. [15]These authors contributed equally: Øyvind Helgeland, Marc Vaudel. ✉e-mail: pal.njolstad@uib.no; stefan.johansson@uib.no

and polygenic variation to disease risk at both ends of the spectrum[23–25]. A recent investigation of severe childhood obesity found an excess burden of rare, predicted deleterious, variants involving genes near adult obesity loci[26]. Variants with different penetrance were detected in genes in the leptin–melanocortin pathway, a major determinant of satiety and energy expenditure. Interestingly, GWAS suggest that the leptin–leptin receptor (LEP–LEPR) axis is also central to BMI development during infancy and childhood[19,20].

In this study, we investigated the association of common variation with BMI from birth to 8 years of age through a longitudinal analysis in the Norwegian Mother, Father and Child Cohort Study (MoBa)[27]. Using this unique pregnancy-based open-ended cohort with dense harmonized phenotypes and genotypes from both parents and children, we present a detailed characterization of the rapidly changing genetic landscape of BMI during the first 8 years of life.

Results

BMI from 28,681 children was measured at birth, 6 weeks, 3, 6 and 8 months, and 1, 1.5, 2, 3, 5, 7 and 8 years of age (Supplementary Table 1). At each time point, we conducted linear mixed-model regression analyses on standardized BMI under an additive genetic model, followed by approximate conditional and joint (COJO) multiple single-nucleotide polymorphism (SNP) analyses to identify independent signals[28], resulting in 46 independent loci reaching genome-wide significance ($P < 5 \times 10^{-8}$) for at least one time point (Table 1 and Supplementary Table 2). Of these, 29 are novel, that is, they do not have any nearby proxy SNPs ($R^2 > 0.6$) that are genome-wide significant in recent meta-analyses of birth weight and adult BMI[15,29].

Four major association trajectory clusters. We investigated the dynamics of the associations for the 46 loci by projecting their estimated effect sizes over time onto a basis of reference profiles (Fig. 1). We also compared the effect size estimates with published meta-analyses at birth[29] (Fig. 2), investigated the long-term association of the 46 loci during adolescence in the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort[1,30] (Fig. 3), and in adulthood from parents in MoBa and published summary statistics of BMI[15] (Fig. 2). The variants displayed different trajectories (Fig. 1c), demonstrating how the genetics of early childhood BMI is an age-dependent combination of interweaved signals. We define four major clusters of profiles (Fig. 1e and Methods), which we hypothesize to represent distinct biological processes.

The 'birth' cluster represents nine loci previously associated with birth weight[29]. Our longitudinal analysis showed that the association near *SH2B3*, *CCNL1*, *GPSM1*, *GCK* and *DLG4* quickly vanishes after birth, indicating that these loci are conferring pure prenatal influences, while loci near *ESR1*, *DLK1* and *HHEX* seem to influence growth also postnatally (Figs. 1 and 2). The trajectory of *ADCY5*, known primarily as a type 2 diabetes (T2D) locus, is remarkable in presenting a strong association at birth that persists during infancy and childhood, but almost no association with adult BMI[15] (Figs. 2 and 3).

The 'transient' cluster represents 21 independent signals with no effect at birth, a peak association during infancy or early childhood, and little or no effect after the AR. None of the SNPs in this cluster reach genome-wide significance ($P < 5 \times 10^{-8}$) in the largest adult BMI meta-analysis to date[15] (Figs. 1 and 2), and among SNPs that reached $P < 1 \times 10^{-5}$, three of four have an opposite direction of effect on BMI in adulthood compared to infancy (*LEPR* (rs10493377), *MLXIPL* (rs17145750) and *KLF14* (rs287621)). Conversely, of the variants previously implicated in birth weight, only one (*PTCH1* (rs28457693)) is present in this cluster. Thus, this cluster represents biological mechanisms with distinct effects on BMI development in infancy and childhood. The other phenotypes associated with the loci in this cluster are primarily anthropometric traits (Fig. 1 and Supplementary Table 3), yet the majority (11 of 21) are not known to be associated with adult traits.

The 'early rise' cluster represents 12 loci showing a gradually stronger association with BMI from infancy into childhood, plateauing around AR and maintaining some effect until age 7 to 8 years. This cluster includes variants associated with self-estimated comparative height and size at age 10 years in the UK Biobank, as well as traits related to adult body composition, which supports the hypothesis of a more persisting effect. However, while the effect sizes for approximately half of the variants in this cluster are consistent throughout adolescence and towards adulthood, the effect vanishes for the other half. Eventually, only two SNPs in this cluster (*ADCY3* (rs11676272) and *TNNI3K* (rs10493544)) reach genome-wide significance in the largest adult BMI study[15], while one SNP (AC105393.2 (rs77165542)) with no proxy in Yengo et al.[15] showed an association with BMI in the parents in MoBa, and the nine others showed no association with adult BMI per se (Figs. 1 and 2).

The 'late rise' cluster, represents four loci (*FTO* (rs17817288), *MC4R* (rs78263856), *SEC16B* (rs545608) and *FAIM2* (rs7132908)) that show little to no association before AR where they exhibit a rapid increase contrasting with the other clusters. The variants in the late rise cluster are in high linkage disequilibrium (LD) with loci reported in a previous study on childhood BMI consisting mainly of children measured at ages 6 to 10 years[16] and with adult BMI[15] (Fig. 2). The observed upward trajectory thus yields effects that seem to persist during adolescence and remain significant into adulthood.

Effect trajectories of known birth weight and adult body mass index SNPs. The density of the overall distribution of trajectory profiles in MoBa for all previously detected birth weight and adult BMI SNPs[15,29] along with the density of the 46 early growth loci detected in this study is depicted in Fig. 1h: the trajectories for birth weight and adult BMI segregate to the left and right sides of the space defined by the reference profiles, respectively, while the early growth BMI is dominated by transient profiles. In contrast to the association profiles in the birth cluster, the birth weight variants mostly display trajectories persisting or rising throughout childhood. Conversely, variants associated with adult BMI presented a strong concentration of late rising profiles, suggesting that better power at late ages would provide a higher number of variants in this cluster.

Trajectory agreement between the MoBa and ALSPAC cohorts. The trajectories of the effect size estimates are generally consistent between the two cohorts (Fig. 3 and Extended Data Fig. 1). However, ALSPAC estimates in early childhood present high s.e. values due to smaller sample sizes especially during infancy and early childhood (Supplementary Table 1). Despite this modest power, a sign test showed that the directions of effect at the peak-effect time points from the MoBa cohort are highly consistent between MoBa and ALSPAC, both cumulatively ($n = 40/45$ consistent, $P < 10^{-7}$) and for each of the four clusters (birth, 9/9; transient, 18/20; early rise, 9/12; late rise, 4/4). A considerably larger sample size is needed to enable formal replication at individual loci.

SNP heritability and genetic correlation. We estimated SNP-based heritability and genetic correlation between various traits and BMI at all time points using LD score regression. The heritability estimates vary with age in a pattern mirroring childhood BMI curves (Extended Data Fig. 2 and Supplementary Table 4). Overall, the phenotypes assessed displayed age-dependent genetic correlation patterns with BMI, with lower correlation from 6 months to 3 years (Extended Data Fig. 3 and Supplementary Table 4). Birth weight adjusted for maternal effect presented a high genetic correlation

Table 1 | Association summary statistics for the top hits

Name	SNP	Chr	Position	EA	OA	EAF	Age	Beta	s.e.	*P* value	Cluster
LEPR	rs10493377	1	65,879,252	A	G	53%	1.5 y	0.057	0.010	2.20×10^{-9}*	Transient
LEPR	rs10889551	1	65,906,137	G	A	65%	1 y	0.088	0.010	5.20×10^{-19}*	Transient
LEPR	rs2767486	1	65,991,203	G	A	16%	6 m	0.143	0.012	6.40×10^{-34}*	Transient
TNNI3K	rs10493544[aBMI]	1	74,983,835	T	C	43%	8 m	0.054	0.009	1.40×10^{-8}	Early rise
SEC16B	rs545608[aBMI]	1	177,899,121	C	G	23%	8 y	0.088	0.016	3.20×10^{-8}	Late
NR5A2	rs2816985	1	200,072,966	G	A	45%	3 m	0.059	0.009	5.40×10^{-11}	Transient
AC105393.2	rs77165542	2	430,975	C	T	98%	1.5 y	0.187	0.032	3.50×10^{-9}	Early rise
ADCY3	rs11676272[aBMI]	2	25,141,538	G	A	49%	1 y	0.089	0.009	2.80×10^{-22}	Early rise
ADCY5	rs11708067[BW]	3	123,065,778	G	A	23%	Birth	0.079	0.010	5.20×10^{-16}	Birth
CCNL1	rs1482853[BW]	3	156,798,473	C	A	60%	Birth	0.099	0.008	5.90×10^{-32}	Birth
LCORL	rs2610989[BW]	4	18,022,834	T	C	26%	1.5 y	0.060	0.011	5.50×10^{-8}*	Early rise
HHIP	rs1032296	4	145,434,688	T	C	38%	6 m	0.052	0.009	1.10×10^{-8}	Transient
PCSK1	rs6899303	5	95,650,975	C	A	63%	6 m	0.057	0.009	5.30×10^{-11}*	Transient
PCSK1/CAST	rs263377	5	95,884,775	A	G	41%	1 y	0.054	0.010	2.90×10^{-8}*	Transient
GLP1R	rs2268657	6	39,020,542	T	C	51%	3 m	0.056	0.009	8.40×10^{-10}*	Transient
GLP1R	rs2268647	6	39,043,178	T	C	50%	1 y	0.048	0.009	2.60×10^{-7}*	Transient
GLP1R	rs1820721	6	39,110,046	A	C	49%	6 m	0.061	0.009	7.20×10^{-12}*	Transient
UBE3D	rs209421	6	83,523,684	G	T	26%	6 m	0.073	0.010	5.40×10^{-13}	Transient
ESR1	rs7772579[BW]	6	152,042,502	A	C	70%	Birth	0.065	0.009	5.90×10^{-13}	Birth
OPRM1	rs1772945	6	154,312,285	A	G	56%	8 m	0.056	0.009	3.20×10^{-9}	Transient
GCK	rs78412508[BW]	7	44,223,858	G	A	99%	Birth	0.376	0.047	4.00×10^{-15}	Birth
MLXIPL	rs17145750	7	73,026,378	C	T	84%	6 m	0.070	0.012	6.80×10^{-9}	Transient
LEP	rs10487505	7	127,860,163	C	G	49%	1.5 y	0.056	0.009	3.20×10^{-9}	Early rise
KLF14	rs287621	7	130,435,181	T	C	26%	6 m	0.064	0.010	3.70×10^{-10}*	Transient
KLF14	rs12672489	7	130,483,555	C	T	75%	1.5 y	0.067	0.011	2.10×10^{-9}*	Early rise
HNF4G	rs117212676	8	76,632,003	A	G	2%	6 m	0.166	0.030	1.80×10^{-7}*	Early rise
PTCH1	rs28457693[BW]	9	98,217,348	G	A	13%	6 m	0.073	0.013	2.40×10^{-8}	Transient
GPSM1	rs28642213[BW]	9	139,248,082	A	G	27%	Birth	0.062	0.010	4.70×10^{-11}	Birth
HHEX	rs11187129[BW]	10	94,429,907	C	T	46%	Birth	0.047	0.008	2.10×10^{-8}	Birth
PLCE1	rs1830890	10	96,019,501	G	A	32%	3 y	0.067	0.012	1.30×10^{-8}	Early rise
SCGB1A1	rs1985927	11	62,193,537	C	T	73%	8 m	0.060	0.011	6.80×10^{-9}	Early rise
EHBP1L1	rs2298615	11	65,352,062	T	C	23%	6 w	0.071	0.012	5.40×10^{-9}	Transient
RP11-405A12.2	rs2728641	12	20,111,569	C	T	48%	3 m	0.050	0.009	1.90×10^{-8}	Transient
FAIM2	rs7132908[aBMI]	12	50,263,148	A	G	40%	8 y	0.081	0.014	3.30×10^{-9}	Late rise
RP11-690J15.1	rs6538845	12	98,544,888	C	T	48%	3 m	0.055	0.009	1.50×10^{-9}	Early rise
SH2B3	rs7310615[Both]	12	111,865,049	G	C	55%	Birth	0.050	0.009	6.50×10^{-9}	Birth
NCOR2	rs3741508	12	124,812,678	T	G	86%	8 m	0.083	0.013	1.20×10^{-9}	Transient
DLK1	rs75806555	14	101,189,448	C	T	86%	Birth	0.074	0.012	2.10×10^{-9}	Birth
SH3GL3	rs2585058	15	84,284,552	G	A	53%	8 m	0.063	0.009	8.60×10^{-12}	Transient
FTO	rs17817288[aBMI]	16	53,807,764	G	A	49%	8 y	0.095	0.013	1.30×10^{-12}	Late rise
KIAA0895L	rs111810144	16	67,216,110	T	C	3%	8 m	0.147	0.025	5.20×10^{-9}	Early rise
DLG4	rs739669[BW]	17	7,122,377	A	G	62%	Birth	0.072	0.009	4.70×10^{-17}	Birth
MC4R	rs78263856[aBMI]	18	58,042,821	T	C	95%	7 y	0.150	0.027	3.80×10^{-8}	Late rise
RIN2	rs148252705	20	17,851,179	T	C	97%	3 m	0.157	0.029	2.60×10^{-8}	Transient
EFCAB8	rs13038017	20	31,467,551	C	T	53%	1 y	0.054	0.009	1.20×10^{-8}	Early rise
PTCHD1-AS	rs5926278	X	23,296,291	T	C	2%	3 m	0.149	0.027	4.80×10^{-8}	Transient

An asterisk denotes a locus with multiple signals, independent and significant after COJO analysis. [BW]Variant associated with birth weight (BW) according to Warrington et al.[29]. [aBMI] Variant associated with adult BMI (aBMI) according to Yengo et al.[15]. [Both] Variant associated with both birth weight and adult BMI according to Warrington et al.[29] and Yengo et al.[15], respectively. Loci are ordered according to chromosomal position. SNP, rsID of the SNP with lowest *P* value at age at peak association. Chr and Position, chromosome and position of the SNP in GRCh37 coordinates. EA, effect allele; OA, other allele; EAF, effect allele frequency estimate in MoBa, where the effect allele is the BMI-raising allele at age of peak association. Age, age at peak association defined as the age with the lowest association *P* value. Name, locus name based on the nearest gene or previous naming in the literature. Beta, s.e. and *P* value denote the effect size, standard error and unadjusted *P*-value estimates for the association with standardized BMI at age at the peak, respectively. Cluster, cluster corresponding to the effect size profile over time. Membership to multiple-signal loci and previous association of the lead SNP with BW according to Warrington et al.[29], aBMI from Yengo et al.[15] or both are annotated with superscripts. See Supplementary Table 1 for the number of samples at each time point and the Methods for the statistical analysis.

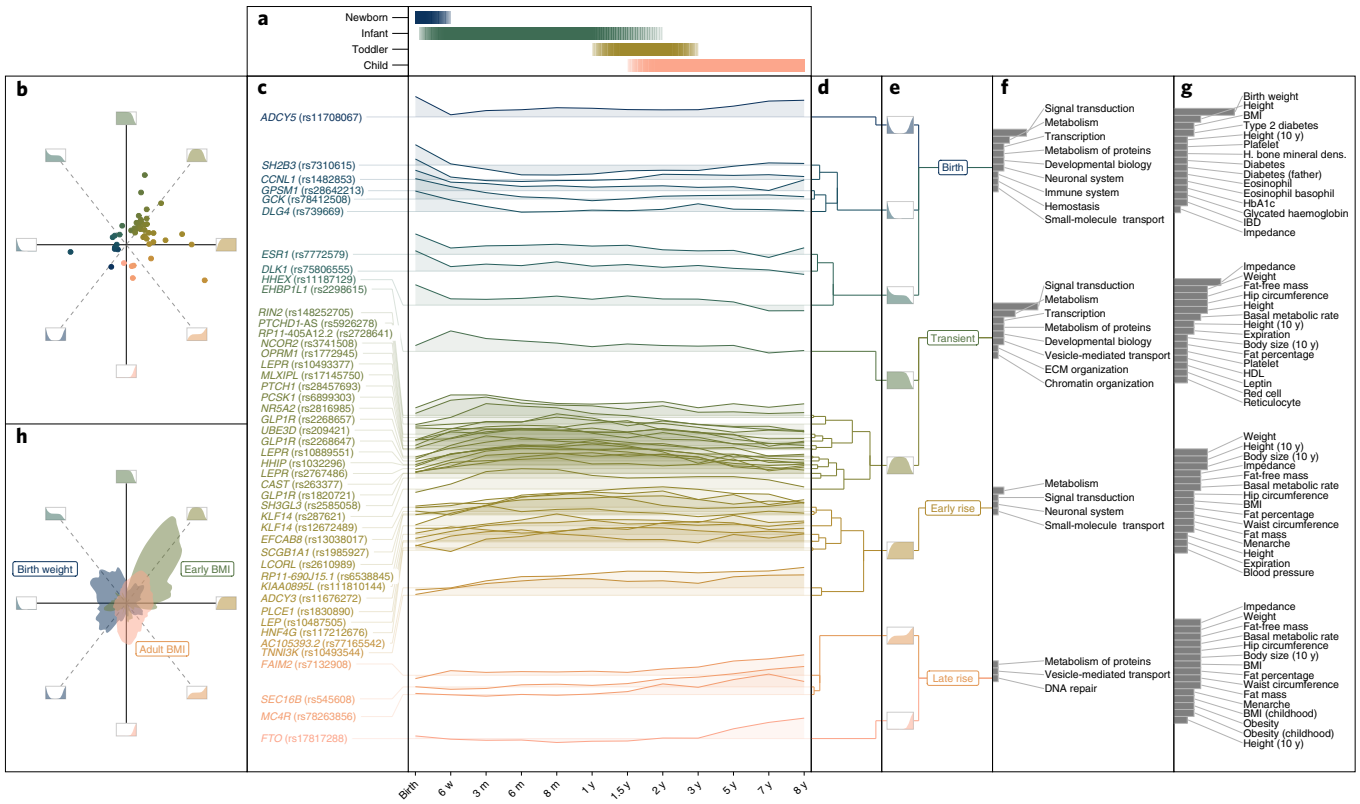


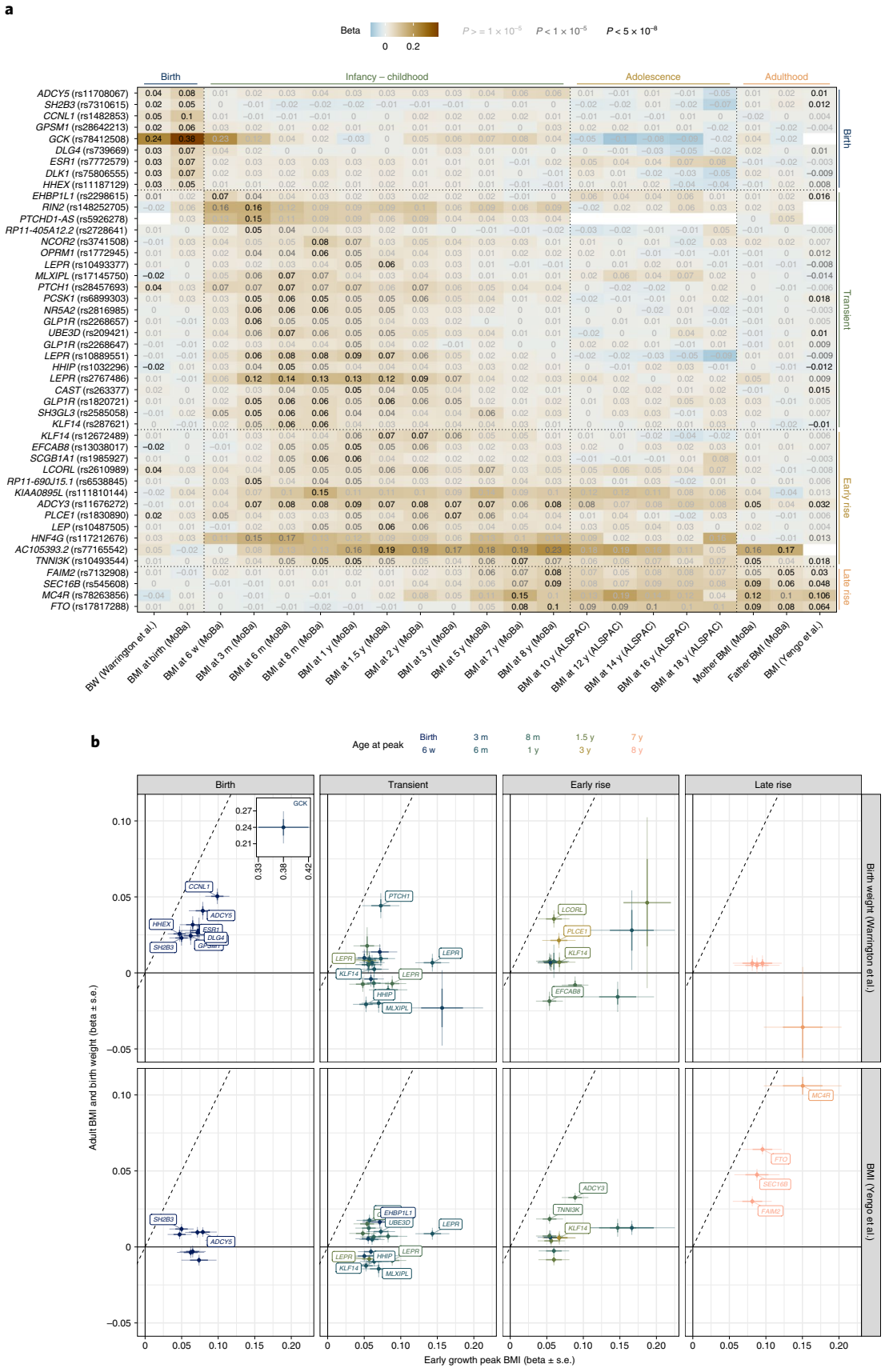
Fig. 1 | Longitudinal association effect size profiles for the 46 top hits. a, Age span of the early childhood developmental stages covered by this association study with BMI. **b**, Quadrant plot of the 46 top hits where the radial and angular coordinates of a SNP indicate the magnitude and shape of the effect size profile over time, respectively. Inserts at cardinal and intercardinal directions indicate the association profile represented by a given angular coordinate. **c**, Effect sizes at the different time points grouped by profile similarity, with the vertical position of the profile corresponding to the angular position in **b**. **d**, Dendrogram of the effect size profile clustering. **e**, Grouping of effect size profiles into four main clusters: birth, transient, early rise and late rise. Inserts to the left indicate the association profiles in each cluster. **f**, Overlap with top-level biological pathways. Bars represent the number of variants in a cluster mapping a given pathway. ECM, extracellular matrix. **g**, Comparisons with other GWAS present in PhenoScanner. Bars represent the number of variants associated with a trait (P value $= 5 \times 10^{-8}$). IBD, inflammatory bowel disease. **h**, Angular density of beta profiles for variants associated with birth weight (blue) and adult BMI (red), compared to early BMI (green) according to refs. [15,29] and this study, respectively, and processed as in **b**. See 'Clustering of association profiles' for details on how the different panels are built. See Supplementary Table 1 for the number of samples at each time point.

with BMI at birth ($r_g = 0.89$, s.e. $= 0.061$, $P < 1 \times 10^{-47}$) that decreased quickly in infancy and throughout childhood, whereas for indirect maternal effects, the correlation was initially lower but increased from 1 year onwards. While obesity-related traits in general show constant correlation levels before accelerating at 3 years, comparative body size at age 10 in the UK Biobank, in which participants reported being thinner or plumper than average at age 10 years, presented a rapid linear increase throughout development from birth to 7 years ($r_g = 0.86$, s.e. $= 0.06$, $P < 9 \times 10^{-53}$), which is in line with the observed overlap of this phenotype with the early and late rise clusters. Higher childhood BMI correlates with younger age of menarche and taller stature in early puberty, indicating a strong genetic correlation between childhood BMI and early pubertal development. Despite 11 of the 46 top hits having previously been associated with adult height (Supplementary Table 3), the overall genetic correlation with adult height is close to zero for all time points. The well-known inverse relationship of T2D with fetal growth vanishes quickly after birth and the genetic correlation of BMI with glycaemic traits varies rapidly throughout childhood.

Monogenic obesity and the leptin–melanocortin pathway. We further investigated whether genes involved in monogenic obesity

Fig. 2 | Comparison with previous studies on birth weight and adult BMI. a, Heat map of the effect size for the 46 top hits from birth to adulthood. Variants are ordered vertically according to Fig. 1c. The estimated effect size for association with birth weight (Warrington et al.[29]; column 1), BMI during early growth (this study, MoBa cohort, columns 2–12), BMI during preadolescence and adolescence (this study, ALSPAC cohort, columns 13–17) and adult BMI (this study, mothers and fathers of the MoBa cohort (columns 18–19) and Yengo et al.[15] (column 20)) is displayed in each cell. The cell colour represents the estimated effect size and the text colour represents the unadjusted P value. Empty cells indicate that no proxy could be found for the given variant in the given study (Methods). **b**, Estimated effect size for association with birth weight[29] and adult BMI[15] plotted against the estimated effect size at the age of peak association during early growth (this study). The dashed lines indicate equal effect sizes in both studies. The colour represents the age of peak association, as defined as the age with the lowest P value. Variants are grouped according to their profile cluster as defined in Fig. 1e. Thick and thin error bars represent one s.e. value on each side of the effect size estimate and 95% confidence intervals, respectively. For the sake of readability, *GCK* at birth is plotted in an insert with a different scale, and the axes might crop the 95% confidence intervals. See Supplementary Table 1 for the number of samples at each time point and Methods for the statistical analysis.

are overrepresented in the vicinity of the loci. Of 42 genes used in routine testing for monogenic and severe early-onset obesity, seven resided within 250 kb of one of the 46 top hits (overrepresentation $P < 1.01 \times 10^{-7}$; Supplementary Table 5). Six of these seven genes encode proteins participating in the leptin–melanocortin pathway (*LEP*, *LEPR* (three signals), *PCSK1* (two signals), *POMC*, *ADCY3*



and *MC4R*) providing compelling support for the importance of this pathway also in normal growth. Apart from *MC4R*, the associated variants belong to the transient and early rise clusters, showing that mechanisms at play act very early after birth, some of which in a narrow age window (Extended Data Fig. 4).

Key roles for variants in the *LEP* and *LEPR* loci. The strongest association with BMI across all time points was the intronic variant rs2767486 with peak association at 6 months in the *LEPR* locus (transient cluster; EAF = 16%; $\beta = 0.14$, s.e. = 0.012, $P < 6.4 \times 10^{-34}$), presenting a transient association profile that peaked at 6 months, in agreement with previous reports[19,20]. COJO multiple-SNP analysis revealed two additional independent signals in this locus (Supplementary Table 6 and Extended Data Fig. 5). The previously described association with rs10487505 in *LEP*[19] was assigned to the early rise cluster. Its child BMI-increasing allele is associated with lower plasma leptin levels adjusted for BMI in adults[31], and our results suggest that the association with BMI is specific to childhood.

Established BMI variants near *ADCY3* and *MC4R*. Both *ADCY3* and *MC4R* are implicated in Mendelian forms of obesity and polygenic BMI in adults and children and expressed in the hypothalamus where they are important for central regulation of energy homeostasis[15,32–34]. The well-known nonsynonymous variant rs11676272 in *ADCY3* was the second strongest locus overall for infant and childhood BMI, peaking at 1 year (early rise cluster). The variant rs78263856 upstream of *MC4R* belongs to the late rise cluster, with effects on BMI appearing from 2 years of age, peaking at 7 years and lasting into adult life (Fig. 2).

Novel variants near *PCSK1*. We identified two independent loci near the monogenic obesity gene *PCSK1* (refs. [35,36]; belonging to the transient cluster) (Extended Data Fig. 5). *PCSK1* encodes the prohormone convertase 1/3 (PC1/3), highly expressed in the hypothalamic arcuate nucleus regulating food intake and body weight[37]. No previous phenotypic associations are reported for the lead SNP rs6899303, but the variant is a strong protein quantitative trait locus for PC1/3 (ref. [38]). The second signal, tagged by rs263377, displays its strongest association at 1 year (transient cluster), and associates with multiple adult anthropometric traits including fat-free body mass in the UK Biobank ($P < 1.84 \times 10^{-9}$). None of the two variants are in LD with the *PCSK1* missense variant rs6235 associated with insulin and adult BMI-related traits[23]. The hypothalamic PC1/3 expression is high in two leptin-sensitive neuronal populations: proopiomelanocortin (POMC)-expressing neurons, and neuropeptide Y (NPY) and agouti-related peptide (AgRP)-expressing neurons. In the periphery, PC1/3 is highly expressed in specific ghrelin-expressing endocrine cells in the stomach, the alpha and beta cells of the islets of Langerhans in the pancreas, and various intestinal enteroendocrine cells. These play an important role in appetite, glucose homeostasis and nutrient assimilation by secreting several PC1/3 products including ghrelin, insulin and proglucagon-derived peptides such as the hormone glucagon-like peptide 1 (GLP-1).

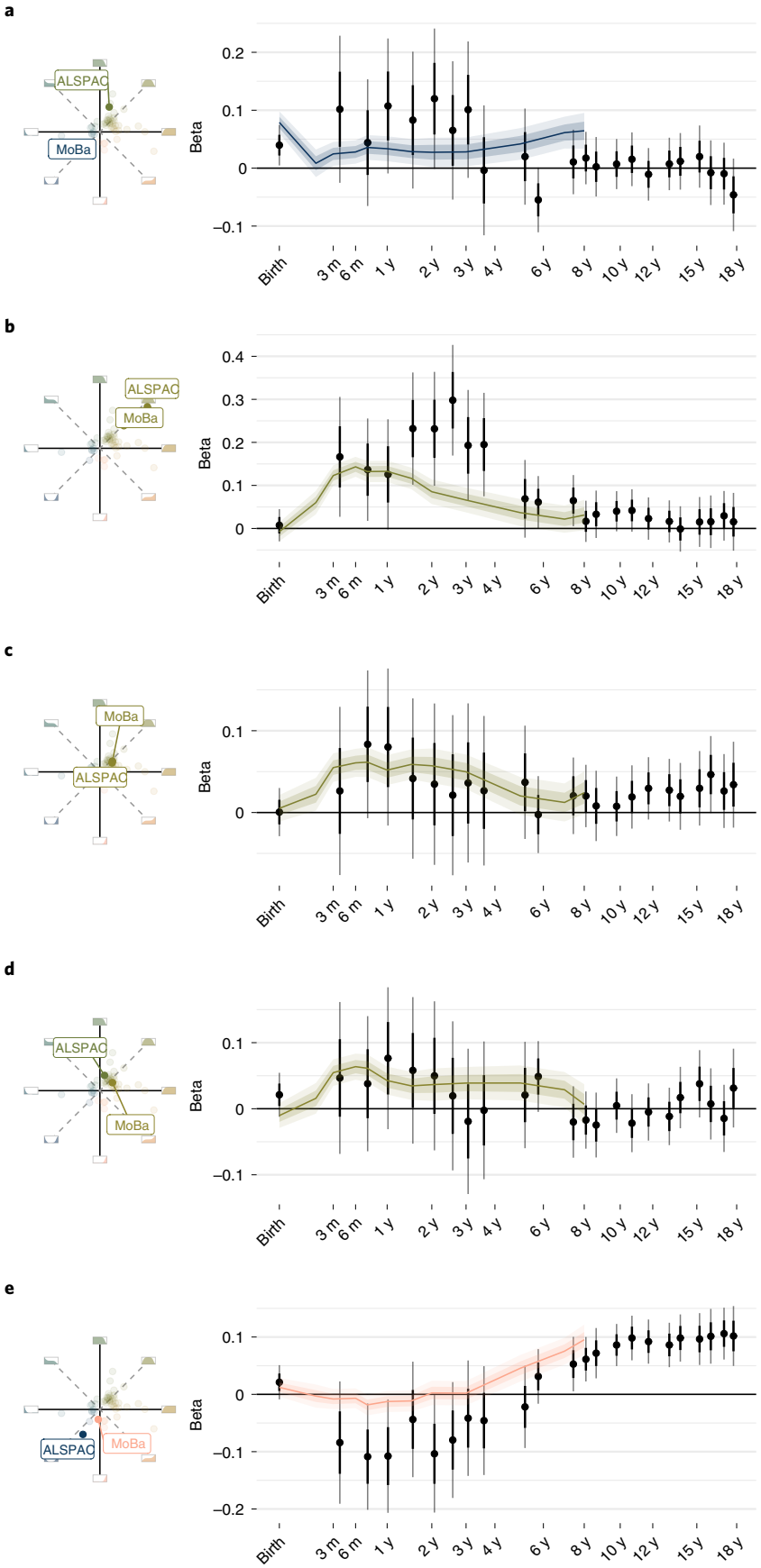
Three novel variants in *GLP1R* with effects on infant body mass index. GLP-1 is released in the small intestines in response to food intake. It interacts with glucagon-like peptide 1 receptor (GLP1R), abundant in hypothalamic regions regulating feeding behavior[39], inducing satiety. It is an incretin with insulinotropic effects in response to oral food intake. GLP-1 improves glucose-stimulated insulin secretion by interacting with the beta cell GLP1R. We identified three independent signals at the *GLP1R* locus belonging to the transient cluster (Extended Data Fig. 5). The strength of association increased for all three variants when analysed together, in particular for rs1820721 (at 6 months, $P_{COJO} < 5.3 \times 10^{-21}$). None of the three SNPs have been associated with childhood or adult BMI. However, the BMI-increasing alleles at rs2268657 and rs2268647 are both associated with lower *GLP1R* expression in stomach, pancreas, and adipose tissues (Genotype-Tissue Expression). Interestingly, the BMI-increasing allele at rs2268657 has previously been associated with faster gastric emptying rate[40], suggesting that *GLP1R* variants may affect childhood BMI through higher digestion rate, in line with its function in the treatment of T2D.

Maternal influences at birth for *SH2B3*, *HHEX* and *ADCY5*. For each of the 46 independent loci, we extended the association model using the parental genotypes, and conducted child–mother–father trio-resolved and haplotype-resolved analyses. For most loci, the child effect at peak association remains after conditioning on the maternal and paternal genotypes, with no noticeable parental effect (Fig. 4). However, for five variants, different patterns emerged: three loci from the birth cluster *SH2B3* (rs7310615), *HHEX* (rs11187129) and *ADCY5* (rs11708067), and two from the transient cluster near *KLF14* (rs287621 and rs12672489).

For the *ADCY5* and *HHEX* loci, associated with T2D and birth weight, respectively, the trio analysis demonstrated opposing fetal and maternal effects, as already observed for birth weight[29], and no effect from the father (Supplementary Table 7). This differs from the *SH3B2* locus, where the trio analysis indicated a dual and directionally consistent effect from both maternal and fetal alleles on birth BMI. The association trajectory of these three birth weight loci illustrates how the maternal genome provides heterogeneous indirect effects on fetal growth that vanish after birth with different dynamics (Fig. 4 and Supplementary Table 7).

Age-dependent association with imprinting patterns near *KLF14*. We identified two variants associated with childhood BMI upstream of *KLF14*, rs287621 and rs12672489, separated by a recombination hotspot. Maternal imprinting has been demonstrated for *KLF14* in T2D[41], with risk alleles associated with increased fasting insulin, reduced high-density lipoprotein cholesterol and decreased expression in adipocyte in adults, only when inherited from mothers[41,42]. Our haplotype analysis revealed that the association for both variants is driven by the maternally inherited allele throughout infancy, with little to no contribution from the paternal allele and the non-transmitted alleles (Fig. 4 and Supplementary Table 7), consistent with imprinting effects. While rs287621 is associated with several adult phenotypes, the strongest known association for rs12672489 is comparative body size at age 10 in the UK Biobank ($P < 3.5 \times 10^{-7}$), showing that this variant influences childhood growth despite residing outside the region critical for adult traits. Expression quantitative trait locus studies have linked variants to the abundance of *KLF14* transcripts in adipose tissue[43] and a variant near *KLF14* has been associated with lower plasma leptin levels[44], offering a mechanistic hypothesis and yet another putative link between leptin regulation and weight gain in infancy.

Fig. 3 | MoBa effect trajectories overlayed with association profiles obtained from ALSPAC. a–e, Effect size estimates for rs11708067 in *ADCY5* (**a**), rs2767486 in *LEPR* (**b**), rs1820721 in *GLP1R* (**c**), rs287621 in *KLF14* (**d**) and rs17817288 in *FTO* (**e**) obtained in the MoBa and ALSPAC cohorts. The quadrant plots (left) display the shape of the effect size estimate over time as obtained in Fig. 1b, for both cohorts, between birth and 8 years of age. The effect size estimates are plotted at each age (right) using the line and ribbons for MoBa and the point and error bars for ALSPAC. To maintain readability of earlier time points, the scale of the *x* axis is not linear. Thick and thin error bars/ribbons represent one s.e. value on each side of the effect size estimates and 95% confidence intervals, respectively. See Supplementary Table 1 for the number of samples at each age bin.


a
ALSPAC
MoBa
Beta
0.2
0.1
0
−0.1
Birth
3 m
6 m
1 y
2 y
3 y
4 y
6 y
8 y
10 y
12 y
15 y
18 y
b
ALSPAC
MoBa
Beta
0.4
0.3
0.2
0.1
0
c
MoBa
ALSPAC
Beta
0.1
0
d
ALSPAC
MoBa
Beta
0.1
0
−0.1
e
ALSPAC
MoBa
Beta
0.1
0
−0.1
−0.2

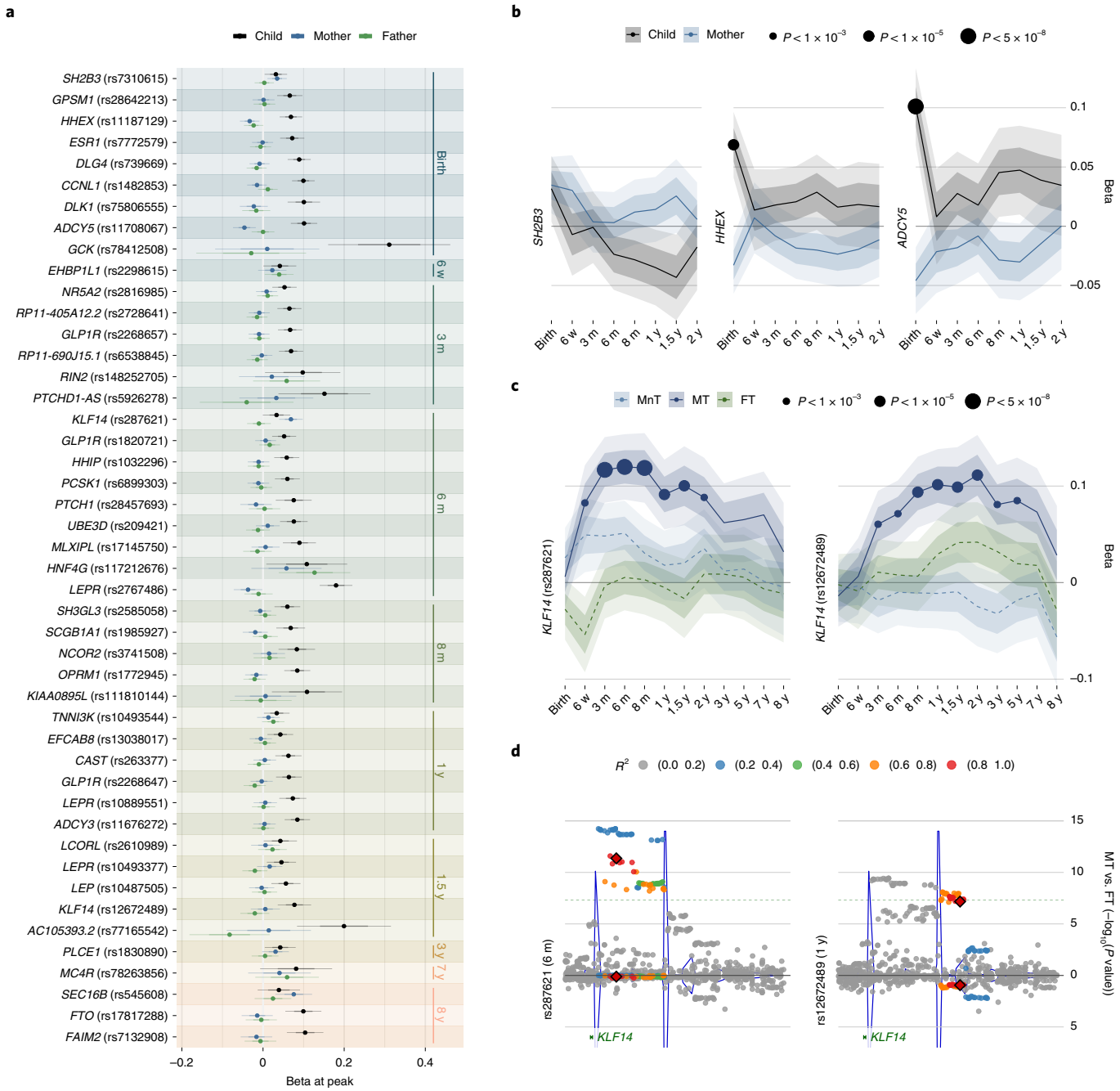


Fig. 4 | Trio-resolved and haplotype-resolved association profiles. a, Effect size estimate for the conditional allelic association of the child, mother and father with child standardized BMI for each of the 46 variants at age at peak association. Here, child, mother and father genotypes are conditioned on each other (Methods). **b**, Association profile for birth weight loci known to present both maternal and fetal effects on birth weight. Effect size estimates of the association with child standardized BMI are represented for the child and the mother from birth to 2 years of age (fathers were included in the model, but not displayed for readability). Unadjusted *P* values represent the significance of the association with the number of effect alleles in the child, mother and father in a joint model, and thus differ from the *P* values of the GWAS. **c**, Association profiles with child standardized BMI from birth to 8 years of age for two variants upstream of *KLF14* in a model combining the child, mother and father alleles into four haplotypes: non-transmitted allele from mother to child (MnT); allele transmitted from mother to child (MT); allele transmitted from father to child (FT); and non-transmitted allele from father to child (FnT). FnT is not represented here for readability; all results are available in Supplementary Table 4. Unadjusted *P* values represent the significance of the association with the number of effect alleles for each haplotype in a joint model. **d**, Regional plots for the unadjusted *P* values of association with the MT and FT haplotypes, top and bottom, respectively, in the haplotype-resolved model. The first and second locus, to the left and right, respectively, are annotated with a red diamond and SNPs coloured according to the LD R^2. The coordinates of the nearest exon encoding KLF14 are annotated at the bottom. Thick and thin error bars and ribbons represent one s.e. value on each side of the effect size estimates and 95% confidence intervals, respectively. See Supplementary Table 1 for the number of samples at each time point and Methods for the statistical analysis.

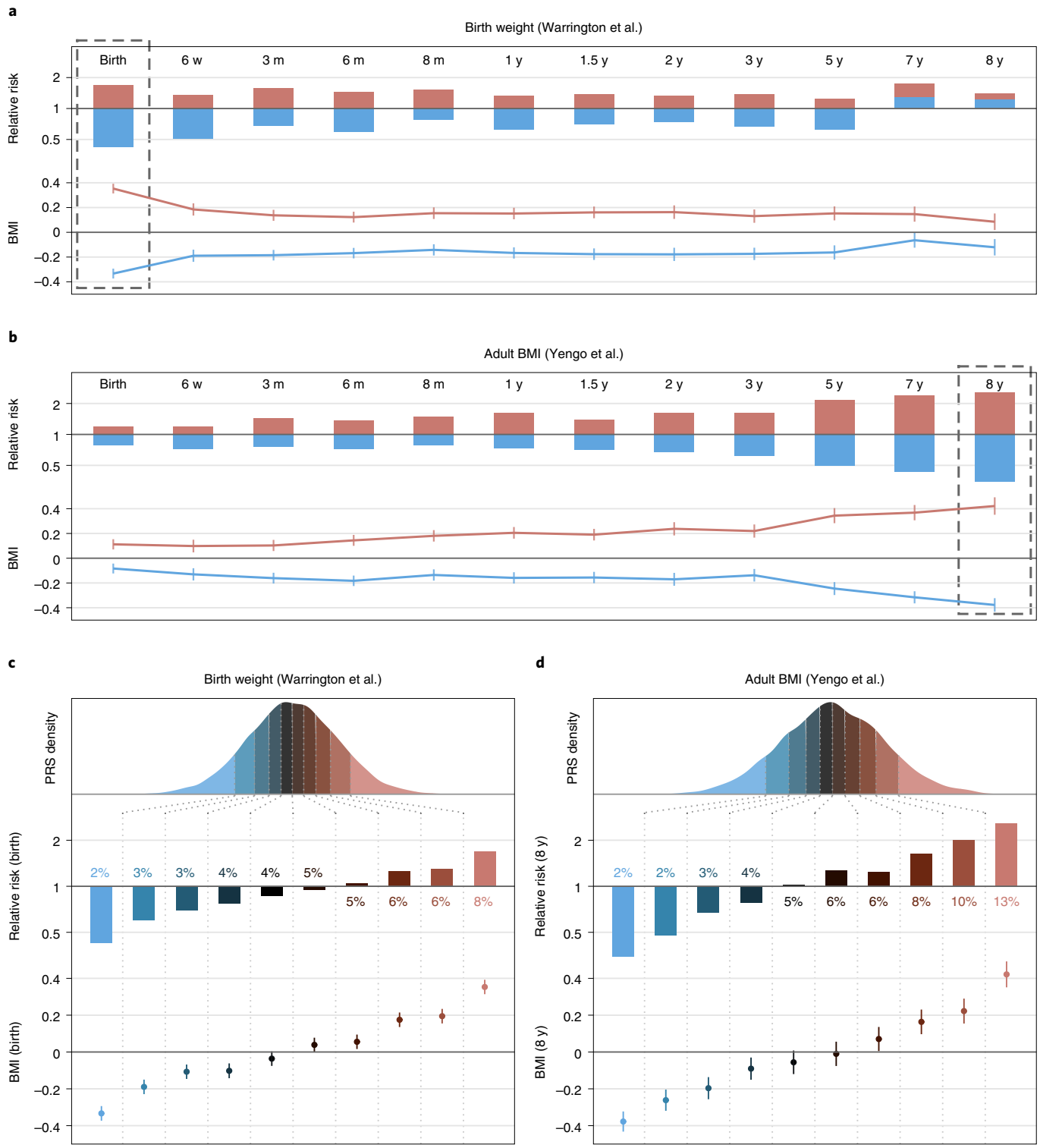


Fig. 5 | Polygenic risk score analyses. a,**b**, Mean standardized BMI of children in this study at each time point after stratification in PRS deciles using PRSs trained using summary statistics from meta-analyses (bottom), and relative risk of obesity for children at a given time point in the top and bottom PRS deciles in red and blue, respectively, compared to the entire cohort (top), where obesity is defined as belonging to the top five BMI percentiles. PRS training was performed using summary statistics for birth weight (**a**) from Warrington et al.[29] and adult BMI (**b**) from Yengo et al.[15]. **c**,**d**, Zoomed views of the dashed rectangles in **a** and **b** showing the stratification by birth weight (**c**) and adult BMI (**d**) PRSs at birth and 8 years, respectively. The density of scores in this study is plotted with the different deciles coloured from left to right. Below, the relative risk of obesity for children in each decile relative to the entire cohort is plotted with the share of obese children in each decile annotated. Bottom, mean standardized BMI of children in each decile. All error bars represent 95% confidence intervals. See Supplementary Table 1 for the number of samples at each time point.

Polygenic transition across infancy and childhood. We constructed polygenic risk scores (PRS) to assess the ability of PRSs of BMI and related traits to stratify BMI and obesity during infancy and early childhood. Strong age-dependent gradients were found with opposing patterns for birth weight and BMI-related traits (Fig. 5, Extended Data Fig. 6 and Supplementary Tables 8 and 9).

For the birth weight-based PRS, the difference in standardized BMI between the 1st and 10th decile was 0.7 at birth (Fig. 5), declining considerably at 6 weeks, and subsequently stabilizing. This residual and lasting association of the birth weight PRS supports an overlap between genetic variants influencing birth weight and BMI development in infancy and childhood. Furthermore, the top risk score decile captures an elevated and consistent share of obese children, even until 7 to 8 years, where it performs similarly to scores trained on childhood BMI and obesity (Extended Data Fig. 6).

The PRS based on adult BMI displayed a shift from 3 to 8 years, where the difference in standardized BMI between the 1st and 10th decile rapidly grows (Fig. 5) and variance explained increases from 0.4% to 5.3% (Extended Data Fig. 6). In the top risk decile, 13% of children were obese at age 8 years, corresponding to a 2.6 times higher risk compared with the median at this age, and a 7.4 times higher risk compared to the bottom risk decile. The PRSs based on previous childhood BMI and obesity studies display similar patterns as those based on adult BMI studies, albeit with lower variance explained (Extended Data Fig. 6). These studies thus mainly capture the genetics of BMI after AR, where the adult architecture is already dominating. Results from both the BMI-adjusted and BMI-unadjusted T2D PRSs show an inverse correlation between BMI at birth and later T2D. However, while this effect quickly vanishes for the unadjusted T2D PRS, children in the top risk decile for BMI-adjusted T2D risk maintain lower BMI throughout infancy, possibly reflecting the key role of insulin metabolism during early growth[45] (Extended Data Fig. 6 and Supplementary Table 8).

Age-stratified polygenic risk scores improve prediction of childhood body mass index. None of the PRS models above capture the BMI development during infancy and the first years of childhood. We evaluated the improvement in performance of PRS models when training on the time-resolved GWAS results generated in this study compared to models trained on adult BMI using a set of 1,096 children in MoBa that were not included in the GWAS. Age-specific modelling vastly improved the variance explained by the PRS during infancy, especially around the AP at 6 months, where R^2 increased from 1.5% using results from adult BMI to 6.4% using age-specific results (Extended Data Fig. 6 and Supplementary Table 8). We also tested the predictive ability of the 21 variants in the transient cluster, which peaked between 6 months and 1.5 years (P value $< 1 \times 10^{-5}$), and explained between 3.0% and 4.5% of the variance during this age span. Hence, the identified variants in the transient cluster alone explain a substantial proportion of the variance in BMI around the AP. Tracking the share of children in the different risk score strata at each time point yielded interweaved trajectories illustrative of the dramatic changes in the genetics of BMI (Extended Data Fig. 6).

Discussion

The association of common genetic variation with BMI changes rapidly during infancy and early childhood, which are stages of life characterized by rapid development and drastic changes in the environment, body composition and metabolism. From the 46 independent loci that we associated with childhood BMI, 29 were not associated with birth weight or adult BMI in large meta-analyses. We proposed to group the genetic association with early BMI into four main clusters that align well with the phases of early growth (Fig. 1): the birth cluster, characterized by loci mainly acting on fetal growth; the transient and early rise clusters, which affect BMI development during the key transitions around AP and rebound; and finally, the late rise cluster of loci, which come into play later in childhood and have persistent influence on BMI into adult life. It is important to note that the assignment of variants to clusters can be misled by uncertainty in effect size estimates, especially at later ages, and uneven distribution of time points, and depends on predefined reference curves. Although the ALSPAC trajectories and summary statistics from adult BMI studies are consistent with our results, further research with larger sample sizes is needed to refine the temporal profiles of these loci and their clustering.

Most of the variants that we discovered show age-specific transient effects and thus would not be identified from GWAS in other age groups. Conversely, early rising loci display gradually stronger effects after birth lasting into the prepubertal age. These loci may be particularly important for processes preceding puberty onset, which is supported by the LD score regression profiles that show gradually increasing genetic correlation between BMI at 3 to 8 years of age and early puberty, higher stature at age 10–12 years and a shorter relative length increase after age 12 years. The age-specific association patterns demonstrate a major change in the underlying genetic architecture of childhood BMI before and after AR, where a shift in association trajectories, genetic correlations, PRS prediction power and heritability occurs. This is further underlined by the large overlap between variants identified in adult BMI and late childhood, but lower overlap with earlier childhood.

An important step in the search for more effective intervention and treatment strategies for childhood and adolescence obesity is to improve our understanding of the genetic and molecular mechanisms influencing BMI development before childhood obesity develops, typically at 5 years of age, to select predisposed children for targeted intervention. Our results point to the substantial inherited variability influencing key genes in the hypothalamic signalling pathway previously known for their role in Mendelian morbid obesity. In addition to replicating the association with variants in *LEP/LEPR*[19,20], we identify two novel variants in *LEPR* and variants near *PCSK1*, *ADCY3* and *MC4R*, all known monogenic obesity genes and central to the hypothalamic signalling pathway; all show age-dependent influences during early childhood. Thus, our findings are highly suggestive of energy intake and expenditure being central to controlling BMI during early childhood, especially before 5 years of age. Notably, many of these genes are already targets for treatment in Mendelian disease, such as leptin-replacement treatment for LEP deficiency and MC4R agonists for LEPR, PCSK1 and POMC deficiency[46–48]. As more genes implicated in monogenic obesity are found to harbour common variants associated with BMI, the notion that monogenic and polygenic obesity share underlying aetiologies is strengthened.

The identification of three novel signals within *GLP1R* offers another important link to putative treatment opportunities. The bidirectional gut–brain axis connecting the enteric with the central nervous system plays a vital role in informing the brain of peripheral energy status. However, relatively few genetic variants associated with genes that have direct or indirect roles in gastrointestinal functions have been associated with childhood obesity. First, the discovery of three novel independent associations in *GLP1R* not picked up in the much bigger meta-analyses on adult BMI is advocating for distinctly different underlying biology driving early BMI development. Second, it iterates on the importance of hypothalamic signalling and further establishes the importance of common variation in genes related to the gut–brain axis in development of early childhood BMI. Finally, increased understanding of GLP-1 signalling in early childhood BMI development is particularly important as GLP1R is a pharmaceutical target for treating adult obesity[49] and recently showed promising results for treating obesity in adolescence[50]. A study of patients treated with the GLP1R agonist liraglutide found alterations in brain activity related to highly desirable food cues and reduced activity in areas of the brain involved in

the reward system[51]. Mice injected with liraglutide show increased energy expenditure through stimulation of brown adipocyte thermogenesis acting through hypothalamic processes[52]. Animal studies have shown that GLP1Rs located in the brain mediate the effect of liraglutide on weight loss. A previous study found that knocking down *GLP1R* in the brain eliminated the effect of liraglutide, while knocking down the same receptor in the peripheral nervous system did not reduce its efficacy significantly[53]. *GLP1R* expression in adipose tissue has also been linked to increased insulin sensitivity[54].

Combined, these results point towards a key role of the central melanocortin system for appetite and energy expenditure early in life, and in particular highlights the POMC system as a putative drug target. This shows that well-powered GWAS of BMI performed in young children can identify novel genes, proteins and pathways not found in adult GWAS, with putative potential for obesity treatment. However, translating GWAS findings into function is challenging and, for most of the discovered loci, more research is needed to reveal the precise molecular and physiological mechanisms involved.

Child–mother–father trio analyses revealed that the association for two independent loci near *KLF14* is driven by the maternally transmitted allele only, suggesting that the paternal allele is silenced. Maternal imprinting for variants in *KLF14* has previously been identified for T2D, and one of our variants tags the same signal, while the other is a novel secondary imprinting effect acting on *KLF14* in early childhood. Additionally, our PRS analysis using a T2D reference study finds persistently low BMI during childhood for children in the highest decile of BMI-adjusted T2D PRSs, and it is tempting to ascribe these late effects on childhood BMI to mechanisms acting through insulin and glucose metabolism given the numerous studies associating *KLF14* with T2D. However, alleles in high LD with the infant BMI and T2D risk-increasing alleles were recently associated with lower plasma leptin levels adjusted for BMI[44], offering yet another putative link between leptin regulation and weight gain in infancy.

Polygenic risk prediction provides opportunities to estimate an individual-level genetic liability and may potentially be used for early identification of children with considerable risk for developing obesity. Here, we show striking differences in BMI between children in the top and bottom deciles of an adult BMI-based PRS concurrent with timing of the AR. Notably, the effect estimates in MoBa are almost identical to what was previously described for British children from ALSPAC[12], suggesting that this score is transferable between Scandinavian and British children. We also show that the PRS can identify children at considerably higher risk of being obese already from 5 years of age. As much as 13% of children in the top decile could be defined as obese at age 8 years, corresponding to a sevenfold higher risk compared to the bottom risk decile (Fig. 5). The shift in genetic architecture before age 5 years renders PRSs based on adult BMI inferior to age-resolved scores during infancy. The testing in our independent sample demonstrates that BMI in the earlier years of life is shaped by a complex interplay and transitions from both age-restricted and more long-term genetic influences that have to be considered when evaluating a child's growth pattern and the potential for targeted interventions. Although both sensitivity and specificity of current PRSs for obesity still are low[24], PRS stratification may help identify selected groups of children that benefit more from early intervention or tailored treatment.

Our study sample consists of a single cohort of northern European descent, and further research is needed to evaluate the generalization of the results to other populations. However, the larger size of the current MoBa release, the availability of parental data and the homogeneous phenotyping allowed us to perform much more detailed time-resolved analyses than typically possible in a meta-analysis involving studies performed under different protocols and data collection time points. The age-dependent association patterns identified here illustrate the importance of early age sampling, and the need for unifying data collection and measurements across cohorts to balance the putative benefit from increased sample size without introducing considerable variance in the phenotyping.

In conclusion, our results provide a fine-grained understanding of the changing genetic landscape regulating BMI from birth to 8 years. The identified loci represent clusters of association trajectories that reflect various phases of growth and highlight a fundamental role of pathways involved in appetite regulation and energy metabolism in both normal growth and rare syndromic obesity. These results demonstrate a strong genetic drive ensuring that children gather the energy necessary to sustain healthy growth.

Methods

Ethics. Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Informed consent for the use of data collected via questionnaires and clinics was obtained from study participants following the recommendations of the ALSPAC Ethics and Law Committee. Consent for biological samples was collected in accordance with the Human Tissue Act (2004). The administrative board of MoBa led by the Norwegian Institute of Public Health approved the study protocol. The establishment of MoBa and initial data collection was based on a licence from the Norwegian Data Protection Agency and the study was approval by The Regional Committee for Medical Research Ethics (no. 2012/67). The MoBa cohort is currently regulated by the Norwegian Health Registry Act.

Study population. MoBa is an open-ended cohort study that recruited pregnant women in Norway from 1999 to 2008. Approximately 114,500 children, 95,200 mothers and 75,000 fathers were enrolled in the study from 50 hospitals across Norway[27]. Anthropometric measurements of the children were carried out at hospitals at birth and during routine visits in the primary health care system by trained nurses at 6 weeks, at 3, 6 and 8 months, and at 1, 1.5, 2, 3, 5, 7 and 8 years of age. Parents later transcribed these measurements to questionnaires. In 2012, study personnel from SELECTionPREDISPOSED (an ERC AdG-supported University of Bergen project) and Better Health by Harvesting Biobanks (HARVEST) randomly selected 11,490 umbilical cord blood DNA samples from the biobank of this study for family triad genotyping, excluding samples matching any of the following criteria: (1) stillborn, (2) deceased, (3) twins, (4) non-existent data at the Norwegian Medical Birth Registry, (5) missing anthropometric measurements at birth in Medical Birth Registry, (6) pregnancies where the mother did not answer the first questionnaire (as a proxy for higher dropout rate) and (7) missing parental DNA samples. In 2016, HARVEST researchers randomly selected a second set of 8,900 triads using the same criteria. The same year NORMENT personnel selected 5,910 triads with the same selection criteria as those used for HARVEST, and extended this with 3,209 triads in 2018. Additionally, a study from 2014 genotyped 1,062 attention deficit hyperactivity disorder (ADHD) cases among the children, and, in 2015, a study genotyped 5,834 randomly selected parents.

Genotyping. Genotyping of the samples was performed in seven different batches on different Illumina platforms over 4 years. Genotyping in the SELECTionPREDISPOSED and HARVEST studies was performed using Illumina HumanCoreExome-12 (v1.1) and HumanCoreExome-24 (v1.0) arrays for 6,938 and 4,552 triads, respectively, at the Genomics Core Facility in the Norwegian University of Science and Technology, Trondheim, Norway. The second wave of genotyping in HARVEST was genotyped using Illumina's Global Screening Array (v1.0) for all 8,900 triads at the Erasmus University Medical Center in Rotterdam, the Netherlands. NORMENT genotyped 5,910 triads using InfiniumOmniExpress-24 (v1.2) in 2016 and 3,209 samples using GSA24 (v1.0) in 2018. The 1,062 ADHD cases were genotyped using InfiniumOmniExpress-24 (v1.2) in 2014 and the 5,834 randomly selected controls using HumanOmniExpress-24 (v1.0). All genotyping was conducted at deCODE genetics, Reykjavik, Iceland. The Genome Reference Consortium Human Build 37 (GRCh37) reference genome was used for all annotations.

Genotypes were called in Illumina GenomeStudio (v2011.1) for the 11,490 triads part of HARVEST and v2.0.3 for the remaining batches. Cluster positions were identified from samples with a call rate ≥ 0.98 and a GenCall score ≥ 0.15. We excluded variants with low call rates, signal intensity, quality scores and deviation from Hardy–Weinberg equilibrium based on the following quality-control (QC) parameters: call rate $< 98\%$, cluster separation < 0.4, 10% GC score < 0.3, AA T Dev > 0.025 and Hardy–Weinberg equilibrium P value $< 1 \times 10^{-6}$. Samples were excluded based on call rate $< 98\%$ and heterozygosity excess > 4 s.d. values. Study participants with non-Norwegian ancestry were excluded after merging with ancestry reference samples from the HapMap project (v3).

Pre-phasing and imputation. Before imputation, insertions and deletions were removed to make the dataset congruent with the Haplotype Reference Consortium

(HRC; v1.1) imputation panel using the HRC Imputation preparation tool by Will Rayner (v4.2.5). Allele, marker position and strand orientation were updated to match the reference panel. Pre-phasing was conducted locally using Shapeit (v2.790)[55]. Imputation was performed at the Sanger Imputation Server with positional Burrows–Wheeler transform[56] and HRC (v1.1) as the reference panel.

Phenotypes. Length/height and weight values were extracted from hospital records through the Norwegian Medical Birth Registry for measurements at birth and from the study questionnaires for the remaining time points. In addition, pregnancy duration in days calculated from ultrasound due date was obtained from the Norwegian Medical Birth Registry. Length and weight values were inspected at each age and those provided in centimetre or gram instead of metre and kilogram, respectively, were converted. Extreme outliers, typically an error in handwritten text parsing or a consequence of incorrect units, were excluded. The value x was considered as an extreme outlier if $x > m + 2 \times (\mathrm{perc}_{99} - m)$ or $x < m - 2 \times (m - \mathrm{perc}_{1})$, where m represents the median within the age group, and perc_{1} and perc_{99} represent the 1st and 99th percentiles, respectively.

Outlier detection and missing value imputation. For all children in MoBa ($n > 100{,}000$), length and weight curves were inspected for outlying values, missing values were imputed, and artefacts causing the length of children to decrease were corrected[19]. Length and weight values presenting an extreme peak or an extreme gap were removed. Missing values preceded and followed by at least two measurement points were imputed through interpolation on the growth curve. Length curves were adjusted to prevent peaks to cause length decrease[19]. These steps were conducted iteratively until no data point was changed (Extended Data Fig. 7). Finally, for all children and all time points presenting both length and weight values, the BMI was computed.

Sample selection. From the total set of growth curves, only the genotyped children passing genotype QC were retained. In addition, the following pregnancies were excluded: (1) pregnancies strictly shorter than 37 full gestational weeks (259 d); (2) plural pregnancies; (3) ADHD excess cases; (4) outliers in the principal-component analysis (PCA) of the genotypes. The set of ADHD excess cases were defined as the additional cases included by the ADHD case–control study. Outliers in the PCA represented 6% of the cohort, and were excluded to reduce the risk of systematic bias due to population stratification[57]. The resulting set of 28,681 children was used in genetic association and is referred to in the following as the full set of children. From this, we built a set of child–mother–father trios by selecting children who had both parents genotyped, with parents passing genotype QC and belonging to the central cluster in the PCA of the genotypes. If the members of two different trios were related according to an identity-by-descent analysis (PI_HAT > 0.1), one trio was randomly excluded. The resulting set of 23,538 trios is referred to in the following as the set of unrelated trios. Allele frequencies and LD scores are estimated based on the parents in the set of unrelated trios.

Phenotype standardization. For the full set of children, at each time point, the BMI was standardized using the generalized additive model for location, scale and shape (GAMLSS) v5.1-7 (https://www.gamlss.com/) in R v3.6.1 (2019-07-05)—'Action of the Toes'. Two GAMLSS models based on a log-normal distribution were fitted separately for boys and girls, using pregnancy duration as a covariate, as detailed in Supplementary Table 10. Note that the models of early BMI include a nonlinear dependency on pregnancy duration, but the nonlinear terms had to be removed after 6 months to ensure the convergence of GAMLSS. GAMLSS models were fitted solely on children from the set of unrelated trios. The models obtained were used to compute standardized BMI values for the full set of children using the 'centiles.pred' function of GAMLSS (Supplementary Table 10). All effect sizes are expressed relative to the standardized phenotypes. A child was considered obese if the standardized BMI was strictly higher than qnorm (0.95), where qnorm represents the quantile function of the standard normal distribution.

Genetic association. The association between the genotypes and the standardized phenotypes using linear mixed models was conducted using BOLT-LMM (v2.3.4)[58] in the full set of children using genotyping batch, sex, pregnancy duration and ten principal components as covariates. LD scores were taken from samples of European ancestry in the 1000 Genomes Project[59], and the genetic map files embedded with BOLT-LMM. The genetic relationship matrix was calculated using a set of high-quality markers having minor allele frequency > 0.05 and INFO score > 0.98. A genetic variant was deemed genome-wide significant at a P value $< 5 \times 10^{-8}$ at any given time point. At all loci reaching genome-wide significance, approximate COJO multiple-SNP analyses were conducted using COJO in GCTA (1.93.2b)[28]. Throughout all analyses, the age at peak association refers to the age of lowest P value in the association with BMI, and the effect allele refers to the BMI-increasing allele at age at peak association.

Effect size estimates for the top hits in ALSPAC. Age, weight and height of children were obtained from the ALSPAC cohort[1,30], which corresponds to 15,454 pregnancies, resulting in 15,589 fetuses. Of these, 14,901 were alive at 1 year of age. We jointly used both self-reported values and measurements from the Children in Focus group as obtained from the ALSPAC cohort. Data are available on the ALSPAC website through a fully searchable data dictionary and variable search tool ('Data availability').

Only children listed in the set of unrelated children as provided by the cohort were used. BMI values were computed unless already provided. Values were binned at birth, around 4 and 8 months, around 1, 1.5, 2 and 2.5 years, and around every year from 3 to 18 years of age. When multiple values for the same child were present in the same bin, the one closest to its individual BMI curve was retained. BMI values were standardized using GAMLSS as done for MoBa. Genotypes were extracted using PLINK 1.9 and a linear association between genotypes and standardized BMI was conducted in R.

Obesity gene enrichment analysis. The gene enrichment analysis around the 46 top hits was conducted using the union of two panels of genes implicated in monogenic and severe early-onset obesity: Blueprint Genetics Monogenic Obesity Panel (test code KI1701; https://blueprintgenetics.com/tests/panels/endocrinology/monogenic-obesity-panel/), consisting of 36 genes, and Genomics England severe early-onset obesity panel v2.2, consisting of 32 genes (https://panelapp.genomicsengland.co.uk/panels/130/). The union of the two resulted in 42 genes used in analysis. A list containing gene locations for hg19 was obtained from PLINK 1.9 resources (https://www.cog-genomics.org/plink/1.9/resources/), which contained 25,303 unique genes used in the analysis. A 500-kb window was used to identify genes in the vicinity of the top hits. The significance for the enrichment of monogenic genes compared to random sampling was estimated using the distribution function of the hypergeometric distribution via the function 'phyper' from the R package stats.

Comparison with adult body mass index in MoBa. Pre-pregnancy BMI values were computed using self-reported height and weight for the parents who were genotyped and passed QC, excluding outliers in the PCA of the genotypes (27,088 mothers and 26,239 fathers), yielding 26,062 and 22,719 values for mothers and fathers, respectively. As detailed in Supplementary Table 10, BMI values were standardized using GAMLSS for mothers and fathers separately, using their birth year as the covariate, as a proxy for age. Like for children, GAMLSS models were fitted solely on parents from the set of unrelated trios and used to compute standardized values for all parents, including related parents.

The association between parent BMI and genotypes was computed for mothers and fathers separately, using BOLT-LMM (v2.3.4)[58] as done for the children. The covariates used were the genotyping batch, birth year and ten principal components.

Clustering of association profiles. For each of the 46 independent genome-wide significant variants, alleles were aligned so that the association with standardized BMI was positive at the age of peak association. Effect sizes for all time points were then combined into an association profile for this variant, that is, a vector $\beta = (\beta_{\mathrm{birth}}, \beta_{6\mathrm{w}}, ..., \beta_{8\mathrm{y}})$. Reference profiles showing effect sizes over time corresponding to an association at birth waning afterwards and an increasing association after 1 year of age towards adulthood were built using equations (1) and (2), respectively.

$$x_1(\mathrm{age}) = 10^{-3\frac{\mathrm{age}}{365.25}} \tag{1}$$

$$x_2(age) = 0 \text{ if age } < 365.25, \left(\frac{age - 365.25}{7 \times 365.25}\right)^2 \text{ else} \tag{2}$$

Where x_1 and x_2 represent the reference profiles and 'age' is the age at a given time point in days. These reference profiles are predefined constructs and their parameterization can influence the clustering. They were not tuned towards specific outcomes to avoid overfitting. The association profiles of each variant were then projected onto these reference profiles, by fitting a linear model according to equation (3):

$$\beta \sim x_1 + x_2 + 1 \tag{3}$$

The resulting projection is shown in Fig. 1b. The profiles of equations (1) and (2) correspond to the curves on the West and South cardinal directions of Fig. 1b, while the profiles in all other cardinal and intercardinal directions correspond to linear combinations of these two, yielding eight reference profiles: early fall and late rise (SE), early fall (E), early and late fall (NE), late fall (N), early rise and late fall (NW), early rise (W), early rise and late rise (SW) and late rise (S).

Each variant was plotted on Fig. 1b using the sum of the absolute values of the effect size over time as the radial coordinate, hence avoiding dependency on the reference profiles for this coordinate, and the relative association with x_1 and x_2 to define the angular coordinate, as described in equations (4) and (5), respectively.

$$\rho = \Sigma|\beta| \tag{4}$$

$$\theta = -atan2(\beta_{x_2}, \beta_{x_1}) + \theta_0 \tag{5}$$

Where ρ represents the radial coordinate, θ the angular coordinate, β_{x_1} and β_{x_2} the association between the genetic association profile and the reference profiles x_1 and x_2, respectively, and θ_0 a constant.

Each association profile was plotted after normalization to the association level at age at peak in Fig. 1c using the angular coordinate θ as baseline on the ordinate.

A cardinal or intercardinal cluster was defined for each of the eight reference profiles corresponding to the cardinal and intercardinal directions in Fig. 1b. Every cardinal and intercardinal cluster was assigned a first element chosen to be the variant with the angular coordinate θ closest to the direction (that is, most correlated to that profile). The other variants were then assigned to a cluster based on their angular nearest neighbour, yielding the clustering displayed by the dendrogram of Fig. 1d. Finally, as illustrated in Fig. 1e, the cardinal and intercardinal clusters were grouped into four main clusters: (birth) SE + E + NE; (transient) N + NW; (early rise) W; and (late rise) SW + S.

Mapping to pathways. The lead SNPs of the 46 independent loci were submitted to the Ensembl Variant Effect Predictor[60]. All proteins coded by genes reported with a consequence other than downstream_gene_variant, upstream_gene_variant or intergenic_variant were retained as potentially affected by a given variant. If no such gene was found, the protein coded by the closest gene within 500 kb was retained. Proteins were matched to Reactome[61] using PathwayMatcher[62]. Then, for each of the four main clusters, we built the smallest set of top-level pathways that explained the protein set returned by the Variant Effect Predictor analysis, and counted the number of variants in this cluster affecting a protein in one of these top-level pathways (Fig. 1f). Results for each SNP are reported in Supplementary Table 3.

Mapping to other traits. For each SNP, other associated traits were extracted using PhenoScanner[63,64]. PhenoScanner was queried using 'EUR' and an R^2 threshold of 0.8 for proxies and 5×10^{-8} as the P-value threshold. Synonymous terms were grouped, and, for each of the four main clusters, the number of variants mapping to a given trait relative to the number of variants in the cluster was plotted in Fig. 1g. Results for each SNP are reported in Supplementary Table 3.

Comparison with birth weight and adult body mass index. Summary statistics on birth weight and adult BMI were obtained from Warrington et al.[29] and Yengo et al.[15], respectively. Variants were matched by rsID. For the variants with no match, proxies were sought using LDproxy (https://ldlink.nci.nih.gov/) using a window of 500 kb, CEU as the reference population and an R^2 threshold of 0.2, and alleles were aligned. From the 46 top hits, variants were considered novel if there were no nearby proxy SNPs in high LD ($R^2>0.6$) with the lead SNP in data from Warrington et al.[29] and Yengo et al.[15] that had *a* P value lower than 5×10^{-8}. For comparisons, for each of the 46 top hits, the variant in statistics from Warrington et al.[29] and Yengo et al.[15] with the lowest P value with an LD R^2 value higher or equal to 0.2 was extracted. Summary statistics for all variants in the three datasets are available in Supplementary Table 11.

Subsequently, for all variants associated with own birth weight in the statistics from Warrington et al.[29], and all variants associated with adult BMI in data from Yengo et al.[15], the association profile in MoBa was extracted and the angular coordinate of Fig. 1b was computed by projecting onto the reference profiles as before. The angular density of each study was subsequently computed using sliding windows over θ, normalized to the number of variants in each study, and plotted in Fig. 1h.

Child–mother–father trio and haplotype analysis. At all time points, for all 46 independent genome-wide significant variants, the association with the children's genome was conditioned on the genomes of the parents in the set of unrelated trios using the linear model described in equation (6).

$$\mathrm{BMI} \sim \mathrm{child} + \mathrm{mother} + \mathrm{father} + 1 \quad (6)$$

Where 'BMI' refers to the standardized BMI of the child at a given time point, and 'child', 'mother' and 'father' refer to the number of tested alleles (hard-call genotypes) for this variant in the child, mother, and father genomes, respectively.

Taking advantage of the phasing of the children's genotypes, we could infer the parent of origin of the genotyped alleles as done by Chen et al.[65]. This results in an alternative model that allows studying the association per haplotype in the set of unrelated trios, as detailed in equation (7).

$$\mathrm{BMI} \sim \mathrm{MnT} + \mathrm{MT} + \mathrm{FnT} + \mathrm{FT} + 1 \quad (7)$$

Where MnT and MT refer to the number of tested alleles non-transmitted and transmitted by the mother to the child, respectively. Similarly, FnT and FT refer to the number of tested alleles non-transmitted and transmitted by the father to the child, respectively.

For a given variant, the share of Mendelian errors in the set of unrelated trios was estimated using trios presenting a homozygous parent. Then, a Mendelian error results in a value of −1 or +2 in the non-transmitted allele count. The share of Mendelian error was estimated by comparing the number of such erroneous genotypes to the number of trios with a homozygous parent expected from the tested allele frequency. When the estimated share of Mendelian errors was over 50%, the alleles of the children were swapped.

For the chromosome X, no filtering was done based on ploidy; when only one chromosome was found the allele was assumed to be inherited from the mother. Note that the chromosome X was not phased, yielding a high share of Mendelian errors, approximately 50%, indicative of a random assignment of children alleles. Haplotype analysis was therefore not possible for the variant on chromosome X, while trio analysis is unaffected by this.

For both models, the same covariates were used as for the genetic association analysis using BOLT-LMM, that is, genotyping batch, sex, gestational age and ten principal components, and both phenotypes and genotypes were adjusted for covariates in the same way as BOLT-LMM does. Haplotype and trio analyses were conducted using TrioGen v0.5.0 (https://github.com/mvaudel/TrioGen/) in the OpenJDK Runtime Environment (Zulu 8.20.0.5-linux64; build 1.8.0_121-b15). Summary statistics for all variants are available in Supplementary Table 7.

LD score regression. LD score regression was performed with LD Hub v1.9.0 using LDSC v1.0.027 with all markers remaining after filtering on the provided SNP list as recommended by the LD Hub authors. A total of 1,215,001 markers remained after filtering. All available phenotypes were selected for correlation analyses. Results for all variables along with heritability and QC reports are available in Supplementary Table 4.

Polygenic risk scores. PRSs were calculated using PRSice-2 v2.3.0 (https://www.prsice.info/). For scores based on study results from previous meta-analyses, the results were obtained from EGG (http://egg-consortium.org/) for birth weight, childhood BMI and childhood obesity, GIANT for adult BMI (https://portals.broadinstitute.org/collaboration/giant/) and DIAGRAM (https://diagram-consortium.org/) for T2D. PRSs were calculated separately for all time points for each phenotype using ten principal components, sex, gestational age and genotyping batch as covariates. Samples without a valid BMI measurement for a specific age were excluded from the analysis at that age, but would be included in analyses of other ages should BMI measurement be available. Among the samples reaching analysis at any age; none had missing genotype data because only markers available in HRC 1.1 were used in the analyses and none of the samples had missing covariates. The target dataset provided to PRSice included all markers available after imputation as hard-called genotypes, but were filtered to only include variants present in the respective reference data used in the respective analysis (supplying beta weights for each variant). Variants were excluded by LD pruning using the target dataset and default settings for PRSice (250-kb clump window, R^2 threshold of 0.1, no P-value threshold). In the resulting set of samples and markers, multiple PRS models were generated and fitted by gradually incrementing the inclusion P value by 5×10^{-5}. Finally, the assessed PRS models were ranked by P value of model fit. The PRS model with the best fit at each age was used in downstream stratification analyses. From the full set of children, one in each pair of samples with PI_HAT > 0.1 was removed at random, leaving 25,113 samples for the PRS analyses. Time-resolved scores used age-specific summary results from the primary analyses as the base with the independent set of 1,062 samples from MoBa as the target. Here, ten principal components, sex and gestational age were used as covariates. Defaults were used for all other parameters. A PRS report as formalized by Wand et al.[66] is available in Supplementary Table 12.

Figures. All figures were generated in R v3.6.1 (2019-07-05)–Action of the Toes (https://www.r-project.org/). In addition to the base packages, the following packages were used: tidyr (v1.1.0), janitor (v2.0.1), conflicted (v1.0.4), glue (v1.4.0), stringr (v1.4.0), dplyr (v1.0.0), scico (v1.1.0), RColorBrewer (v1.1-2), ggplot2 (v3.3.2), ggrepel (v0.8.2), grid (v3.6.1), gtable (v0.2.0), patchwork (v1.1.1), PhenoScanner (v1.0) and ggfx (v0.0.0.900).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The full GWAS summary statistics for all time points are available at https://www.fhi.no/en/studies/moba/for-forskere-artikler/gwas-data-from-moba/. Access to genotypes and phenotypes from MoBa is subject to controlled access by the Norwegian Institute of Public Health in accordance with national and international regulations. Conditions of access including contact details for requests can be found at the Norwegian Institute of Public Health website (https://www.fhi.no/en/studies/moba/).

HRC or 1000G Imputation preparation and checking: https://www.well.ox.ac.uk/~wrayner/tools/.

Sanger imputation service: https://imputation.sanger.ac.uk/.

LD score repository: https://alkesgroup.broadinstitute.org/LDSCORE/.

Genotype-Tissue Expression: https://www.gtexportal.org/.

Birth weight reference data[29]: http://egg-consortium.org/BW5/Fetal_BW_European_meta.NG2019.txt.gz.

Adult BMI reference data[15]: http://portals.broadinstitute.org/collaboration/giant/images/c/c8/Meta-analysis_Locke_et_al%2BUKBiobank_2018_UPDATED.txt.gz.

T2D[67]: https://www.diagram-consortium.org/downloads.html.
- T2D GWAS meta-analysis–unadjusted for BMI[67].
- T2D GWAS meta-analysis–adjusted for BMI[67].

Childhood obesity[18]: http://egg-consortium.org/Childhood_Obesity_2019/CHILDHOOD_OBESITY.TRANS_ANCESTRAL.RESULTS.txt.gz.
Childhood BMI[16]: http://egg-consortium.org/Childhood_BMI/EGG_BMI_HapMap_DISCOVERY.txt.gz.
ALSPAC data dictionary and variable search tool: http://www.bristol.ac.uk/alspac/researchers/our-data/.

Received: 26 May 2021; Accepted: 9 February 2022;
Published online: 21 March 2022

References

1. Fraser, A. et al. Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int. J. Epidemiol.* **42**, 97–110 (2013).
2. Rolland-Cachera, M. F., Deheeger, M., Maillot, M. & Bellisle, F. Early adiposity rebound: causes and consequences for obesity in children and adults. *Int. J. Obes.* **30**, S11–S17 (2006).
3. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in body mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2,416 population-based measurement studies in 128.9 million children, adolescents, and adults. *Lancet* **390**, 2627–2642 (2017).
4. World Health Organization. Consideration of the evidence on childhood obesity for the Commission on Ending Childhood Obesity: report of the ad hoc working group on science and evidence for ending childhood obesity (WHO, Geneva, 2016) http://apps.who.int/iris/bitstream/10665/206549/1/9789241565332_eng.pdf?ua=1
5. Singh, A. S., Mulder, C., Twisk, J. W. R., van Mechelen, W. & Chinapaw, M. J. M. Tracking of childhood overweight into adulthood: a systematic review of the literature. *Obes. Rev.* **9**, 474–488 (2008).
6. Woo, J. G. et al. Prediction of adult class II/III obesity from childhood BMI: the i3C consortium. *Int. J. Obes.* **44**, 1164–1172 (2020).
7. Geserick, M. et al. Acceleration of BMI in early childhood and risk of sustained obesity. *N. Engl. J. Med.* **379**, 1303–1312 (2018).
8. MacLean, P. S., Higgins, J. A., Giles, E. D., Sherk, V. D. & Jackman, M. R. The role for adipose tissue in weight regain after weight loss. *Obes. Rev.* **16**, 45–54 (2015).
9. Silventoinen, K. et al. Genetic and environmental effects on body mass index from infancy to the onset of adulthood: an individual-based pooled analysis of 45 twin cohorts participating in the COllaborative project of Development of Anthropometrical measures in Twins (CODATwins) study. *Am. J. Clin. Nutr.* **104**, 371–379 (2016).
10. Silventoinen, K. et al. Differences in genetic and environmental variation in adult BMI by sex, age, time period, and region: an individual-based pooled analysis of 40 twin cohorts. *Am. J. Clin. Nutr.* **106**, 457–466 (2017).
11. Kilpeläinen, T. O. et al. Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. *PLoS Med.* **8**, e1001116 (2011).
12. Khera, A. V. et al. Polygenic prediction of weight and obesity trajectories from birth to adulthood. *Cell* **177**, 587–596 (2019).
13. Yang, J. et al. Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat. Genet.* **47**, 1114–1120 (2015).
14. Yang, J. et al. Genome partitioning of genetic variation for complex traits using common SNPs. *Nat. Genet.* **43**, 519–525 (2011).
15. Yengo, L. et al. Meta-analysis of genome-wide association studies for height and body mass index in ~700,000 individuals of European ancestry. *Hum. Mol. Genet.* **27**, 3641–3649 (2018).
16. Felix, J. F. et al. Genome-wide association analysis identifies three new susceptibility loci for childhood body mass index. *Hum. Mol. Genet.* **25**, 389–403 (2016).
17. Vogelezang, S. et al. Novel loci for childhood body mass index and shared heritability with adult cardiometabolic traits. *PLoS Genet.* **16**, e1008718 (2020).
18. Bradfield, J. P. et al. A trans-ancestral meta-analysis of genome-wide association studies reveals loci associated with childhood obesity. *Hum. Mol. Genet.* **28**, 3327–3338 (2019).
19. Helgeland, Ø. et al. Genome-wide association study reveals dynamic role of genetic variation in infant and early childhood growth. *Nat. Commun.* **10**, 4448 (2019).
20. Alves, A. C. et al. GWAS on longitudinal growth traits reveals different genetic factors influencing infant, child, and adult BMI. *Sci. Adv.* **5**, eaaw3095 (2019).
21. Farooqi, I. S. et al. Effects of recombinant leptin therapy in a child with congenital leptin deficiency. *N. Engl. J. Med.* **341**, 879–884 (1999).
22. Licinio, J. et al. Phenotypic effects of leptin replacement on morbid obesity, diabetes mellitus, hypogonadism, and behavior in leptin-deficient adults. *Proc. Natl Acad. Sci. USA* **101**, 4531–4536 (2004).
23. Turcot, V. et al. Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity. *Nat. Genet.* **50**, 26–41 (2018).
24. Loos, R. J. F. & Yeo, G. S. H. The genetics of obesity: from discovery to biology. *Nat. Rev. Genet.* https://doi.org/10.1038/s41576-021-00414-z (2021).
25. Flannick, J., Johansson, S. & Njølstad, P. R. Common and rare forms of diabetes mellitus: towards a continuum of diabetes subtypes. *Nat. Rev. Endocrinol.* **12**, 394–406 (2016).
26. Marenne, G. et al. Exome sequencing identifies genes and gene sets contributing to severe childhood obesity, linking *PHIP* variants to repressed POMC transcription. *Cell Metab.* **31**, 1107–1119 (2020).
27. Magnus, P. et al. Cohort Profile Update: The Norwegian Mother and Child Cohort Study (MoBa). *Int. J. Epidemiol.* **45**, 382–388 (2016).
28. Yang, J. et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375 (2012).
29. Warrington, N. M. et al. Maternal and fetal genetic effects on birth weight and their relevance to cardio-metabolic risk factors. *Nat. Genet.* https://doi.org/10.1038/s41588-019-0403-1 (2019).
30. Boyd, A. et al. Cohort Profile: the 'children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *Int. J. Epidemiol.* **42**, 111–127 (2013).
31. Sun, Q. et al. Genome-wide association study identifies polymorphisms in *LEPR* as determinants of plasma soluble leptin receptor levels. *Hum. Mol. Genet.* **19**, 1846–1855 (2010).
32. Saeed, S. et al. Loss-of-function mutations in *ADCY3* cause monogenic severe obesity. *Nat. Genet.* **50**, 175–179 (2018).
33. Stergiakouli, E. et al. Genome-wide association study of height-adjusted BMI in childhood identifies functional variant in *ADCY3*. *Obesity* **22**, 2252–2259 (2014).
34. Krashes, M. J., Lowell, B. B. & Garfield, A. S. Melanocortin-4 receptor-regulated energy homeostasis. *Nat. Neurosci.* **19**, 206–219 (2016).
35. Jackson, R. S. et al. Obesity and impaired prohormone processing associated with mutations in the human prohormone convertase 1 gene. *Nat. Genet.* **16**, 303–306 (1997).
36. Martín, M. G. et al. Congenital proprotein convertase 1/3 deficiency causes malabsorptive diarrhea and other endocrinopathies in a pediatric cohort. *Gastroenterology* **145**, 138–148 (2013).
37. Ramos-Molina, B., Martin, M. G. & Lindberg, I. *PCSK1* variants and human obesity. *Prog. Mol. Biol. Transl. Sci.* **140**, 47–74 (2016).
38. Sun, B. B. et al. Genomic atlas of the human plasma proteome. *Nature* **558**, 73–79 (2018).
39. Alvarez, E. et al. The expression of GLP-1 receptor mRNA and protein allows the effect of GLP-1 on glucose metabolism in the human hypothalamus and brainstem. *J. Neurochem.* **92**, 798–806 (2005).
40. Yau, A. M. W. et al. A pilot study investigating the influence of glucagon-like peptide-1 receptor single-nucleotide polymorphisms on gastric emptying rate in Caucasian men. *Front. Physiol.* **9**, 1331 (2018).
41. Small, K. S. et al. Regulatory variants at *KLF14* influence type 2 diabetes risk via a female-specific effect on adipocyte size and body composition. *Nat. Genet.* **50**, 572–580 (2018).
42. Kong, A. et al. Parental origin of sequence variants associated with complex diseases. *Nature* **462**, 868–874 (2009).
43. Yang, Q. & Civelek, M. Transcription factor KLF14 and metabolic syndrome. *Front. Cardiovasc. Med.* **7**, 91 (2020).
44. Yaghootkar, H. et al. Genetic studies of leptin concentrations implicate leptin in the regulation of early adiposity. *Diabetes* https://doi.org/10.2337/db20-0070 (2020).
45. Murray, P. G. & Clayton, P. E. Endocrine control of growth. *Am. J. Med. Genet. C Semin. Med. Genet.* **163C**, 76–85 (2013).
46. Yeo, G. S. H. et al. The melanocortin pathway and energy homeostasis: from discovery to obesity therapy. *Mol. Metab.* **48**, 101206 (2021).
47. Clément, K. et al. Efficacy and safety of setmelanotide, an MC4R agonist, in individuals with severe obesity due to LEPR or POMC deficiency: single-arm, open-label, multicentre, phase 3 trials. *Lancet. Diabetes Endocrinol.* **8**, 960–970 (2020).
48. Clément, K. et al. MC4R agonism promotes durable weight loss in patients with leptin receptor deficiency. *Nat. Med.* **24**, 551–555 (2018).
49. González-García, I., Milbank, E., Diéguez, C., López, M. & Contreras, C. Glucagon, GLP-1 and thermogenesis. *Int. J. Mol. Sci.* **20**, 3445 (2019).
50. Kelly, A. S. et al. A randomized, controlled trial of liraglutide for adolescents with obesity. *N. Engl. J. Med.* **382**, 2117–2128 (2020).
51. Farr, O. M. et al. GLP-1 receptors exist in the parietal cortex, hypothalamus and medulla of human brains and the GLP-1 analogue liraglutide alters brain activity related to highly desirable food cues in individuals with diabetes: a crossover, randomised, placebo-controlled trial. *Diabetologia* **59**, 954–965 (2016).
52. Beiroa, D. et al. GLP-1 agonism stimulates brown adipose tissue thermogenesis and browning through hypothalamic AMPK. *Diabetes* **63**, 3346–3358 (2014).

53. Sisley, S. et al. Neuronal GLP1R mediates liraglutide's anorectic but not glucose-lowering effect. *J. Clin. Invest.* **124**, 2456–2463 (2014).
54. Vendrell, J. et al. Study of the potential association of adipose tissue GLP-1 receptor with obesity and insulin resistance. *Endocrinology* **152**, 4072–4079 (2011).
55. Delaneau, O., Zagury, J.-F. & Marchini, J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat. Methods* **10**, 5–6 (2013).
56. Durbin, R. Efficient haplotype matching and storage using the positional Burrows–Wheeler transform. *Bioinformatics* **30**, 1266–1272 (2014).
57. Marees, A. T. et al. A tutorial on conducting genome-wide association studies: quality control and statistical analysis. *Int. J. Methods Psychiatr. Res.* **27**, e1608 (2018).
58. Loh, P.-R. et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat. Genet.* **47**, 284–290 (2015).
59. 1000 Genomes Project Consortium et al. An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
60. McLaren, W. et al. The Ensembl variant effect predictor. *Genome Biol.* **17**, 122 (2016).
61. Jassal, B. et al. The reactome pathway knowledgebase. *Nucleic Acids Res.* **48**, D498–D503 (2020).
62. Sánchez, L. F. H. et al. PathwayMatcher: proteoform-centric network construction enables fine-granularity multiomics pathway mapping. *Gigascience* **8**, giz088 (2019).
63. Staley, J. R. et al. PhenoScanner: a database of human genotype–phenotype associations. *Bioinformatics* **32**, 3207–3209 (2016).
64. Kamat, M. A. et al. PhenoScanner V2: an expanded tool for searching human genotype–phenotype associations. *Bioinformatics* **35**, 4851–4853 (2019).
65. Chen, J. et al. Dissecting maternal and fetal genetic effects underlying the associations between maternal phenotypes, birth outcomes, and adult phenotypes: a Mendelian-randomization and haplotype-based genetic score analysis in 10,734 mother–infant pairs. *PLoS Med.* **25**, 17 (2020).
66. Wand, H. et al. Improving reporting standards for polygenic scores in risk prediction studies. *Nature* **591**, 211–219 (2021).
67. Mahajan, A. et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat. Genet.* **50**, 1505–1513 (2018).

Acknowledgements

This work was supported by grants (to S.J.) from Helse Vest's Open Research (nos. 912250 and F-12144), the Novo Nordisk Foundation (NNF19OC0057445) and the Research Council of Norway (no. 315599)); and (to P.R.N.) from the European Research Council (AdG SELECTionPREDISPOSED no. 293574), the Bergen Research Foundation (Utilizing the Mother and Child Cohort and the Medical Birth Registry for Better Health), Stiftelsen Kristian Gerhard Jebsen (Translational Medical Center), the University of Bergen, the Research Council of Norway (FRIPRO no. 240413), the Western Norway Regional Health Authority (Strategic Fund Personalized Medicine for Children and Adults), the Novo Nordisk Foundation (no. 54741) and the Norwegian Diabetes Association. This work was partly supported by the Research Council of Norway through its Centres of Excellence funding scheme (nos. 262700 and 223273), Better Health by Harvesting Biobanks (no. 229624) and The Swedish Research Council, Stockholm, Sweden (2015-02559), the Research Council of Norway (FRIMEDBIO nos. 547711 and 273291) and March of Dimes (no. 21-FY16-121). MoBa is supported by the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research, National Institutes of Health (NIH)/NIEHS (contract no. N01-ES-75558), NIH/NINDS (grant nos. UO1 NS 047537-01 and UO1 NS 047537-06A1).

We are grateful to all the families in Norway who are taking part in the ongoing MoBa cohort study.

We are extremely grateful to all the families who took part in the ALSPAC cohort study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council (MRC) and Wellcome (grant ref. 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and S.J and M.V. serve as guarantors for the contents of this paper. A comprehensive list of grant funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf); this research was specifically funded by Wellcome Trust and the MRC (core; 076467/Z/05/Z). ALSPAC GWAS data were generated by sample logistics and genotyping facilities at Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe.

All analyses were performed using digital laboratories in HUNT Cloud at the Norwegian University of Science and Technology, Trondheim, Norway. We are grateful for outstanding support from the HUNT Cloud community.

Author contributions

Ø.H., M.V., P.R.N. and S.J. designed the study. Ø.H. and M.V. analysed the data. Ø.H., M.V. and S.J. interpreted the data. J.J., J.B., G.P.K., T.R.K., P.M., C.S. and O.A.A. contributed to sample acquisition and genotyping. J.J. and J.B. assisted with genotype QC. P.S.-N., C.F., I.L.K., B.B.J., B.J. and P.R.N. critically revised the manuscript for important intellectual content. Ø.H., M.V. and S.J. wrote the manuscript. All authors participated in preparing the manuscript by reading and commenting on drafts before submission. P.R.N. and S.J. acquired the funding.

Competing interests

O.A.A. is a consultant to HealthLytix. The other authors declare no competing interests.

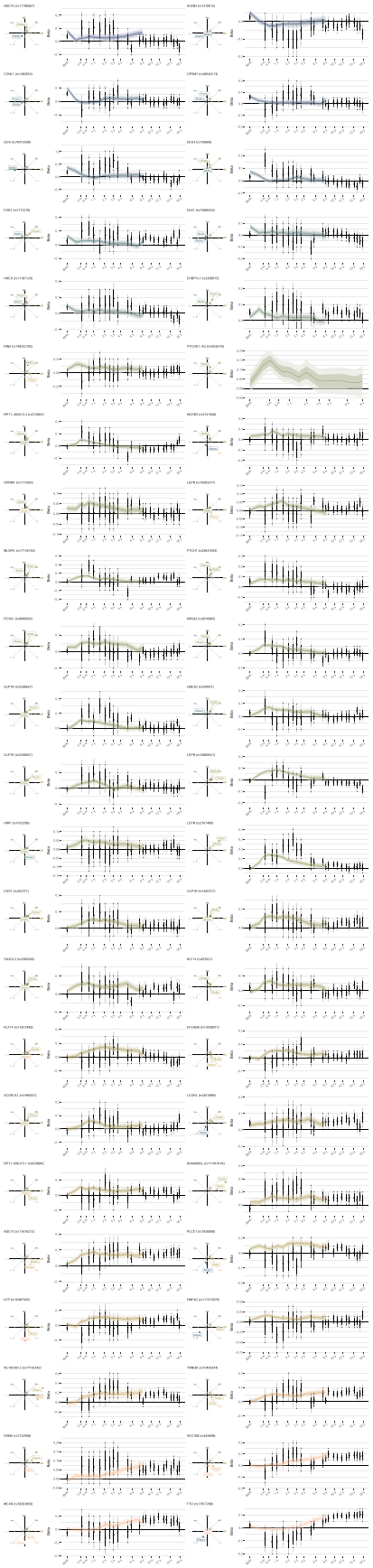
Additional information

Extended data is available for this paper at https://doi.org/10.1038/s42255-022-00549-1.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42255-022-00549-1.

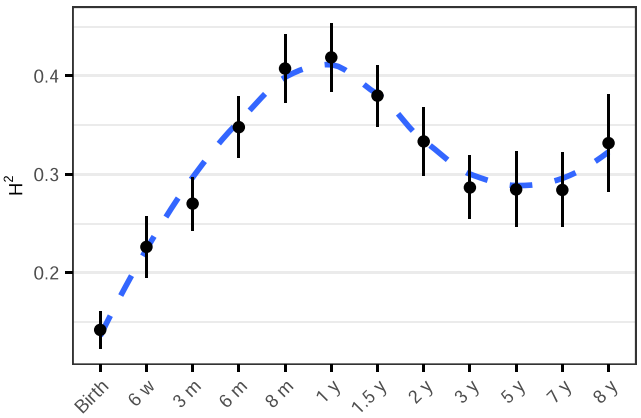
Correspondence and requests for materials should be addressed to Pål R. Njølstad or Stefan Johansson.

Peer review information *Nature Metabolism* thanks Timothy Frayling and the other, anonymous, reviewers for their contribution to the peer review of this work. Primary Handling Editor: Isabella Samuelson, in collaboration with the *Nature Metabolism* team.

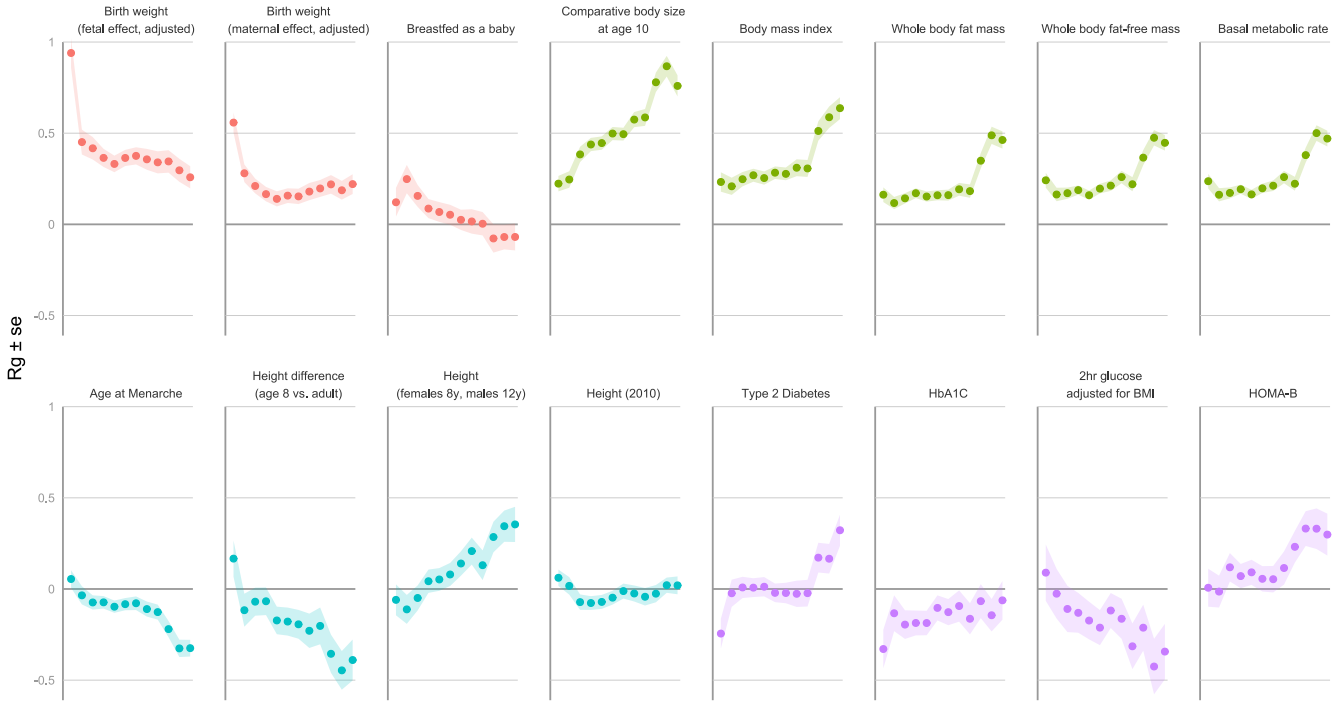


Extended Data Fig. 1 | See next page for caption.

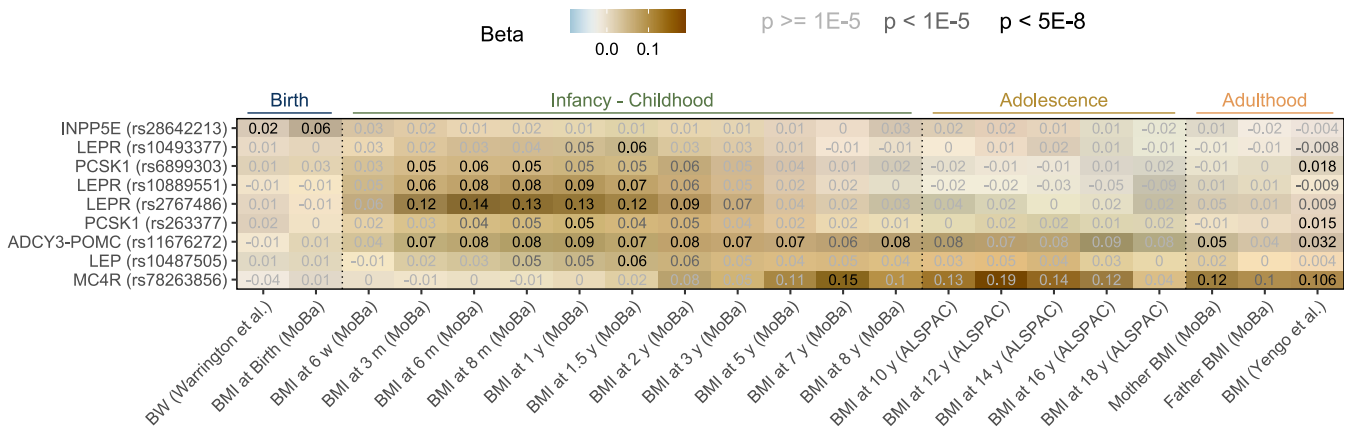
Extended Data Fig. 1 | Overlay with association profiles obtained from ALSPAC. Effect size estimates for all 46 hits obtained in the MoBa and ALSPAC cohorts. The quadrant plots to the left display the shape of the effect size estimate over time as obtained in Fig. 1B, for both cohorts, between birth and eight years of age. The effect size estimates are plotted at each age to the right using line and ribbons for MoBa and point and error bars for ALSPAC. Note that to maintain readability of earlier time points, the scale of the x axis is not linear. Thick and thin error bars/ribbons represent one standard error estimate on each side of the effect size estimates and 95% confidence intervals, respectively. See Supplementary Table 1 for the number of samples at each time point.



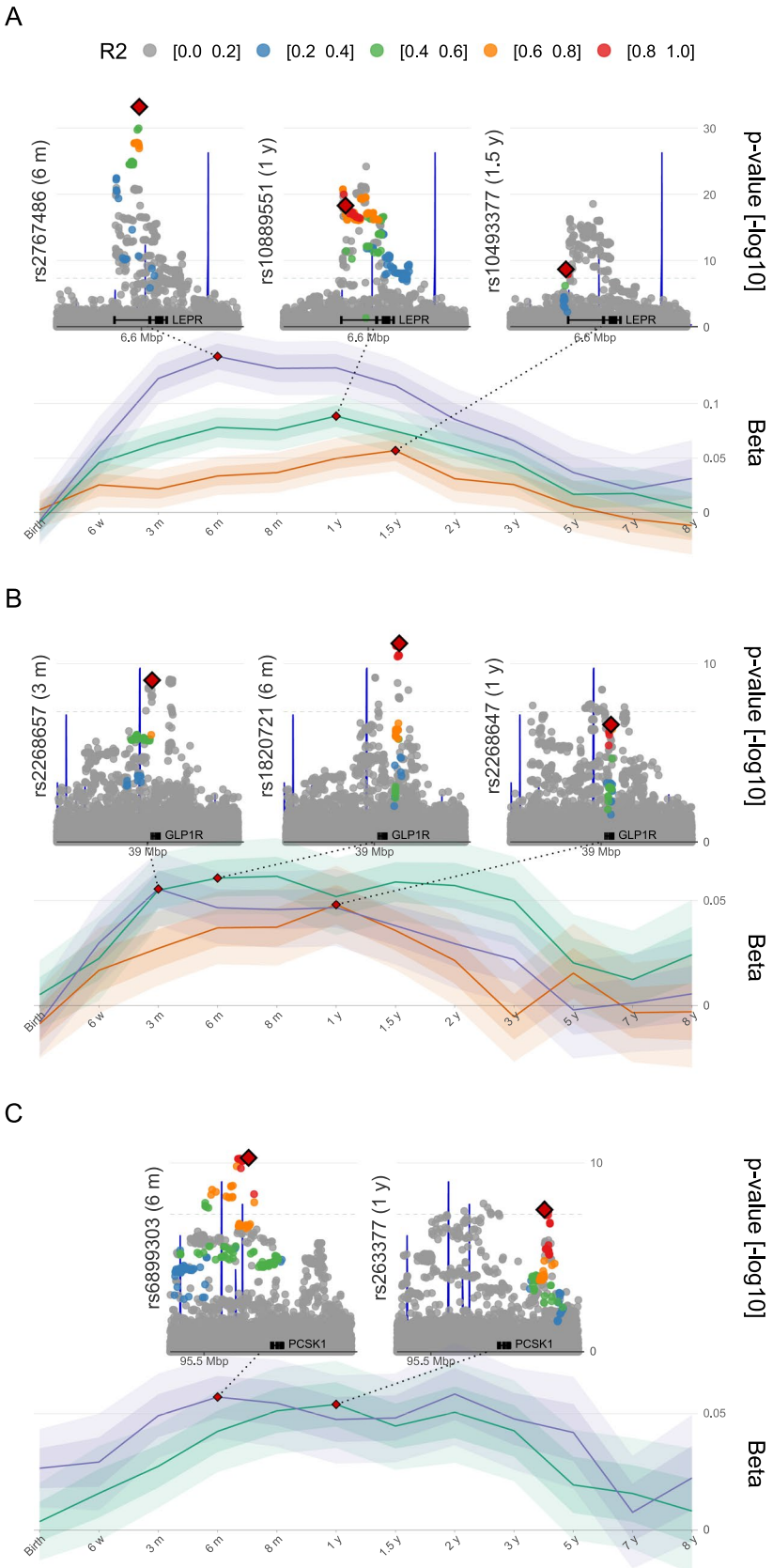
Extended Data Fig. 2 | SNP-based heritability. H^2 estimates from LD score regression for BMI plotted at each time point (black) along with locally estimated scatterplot smoothing (LOESS) local regression (in blue). Error bars represent ±SEM. See Supplementary Table 1 for the number of samples at each time point.



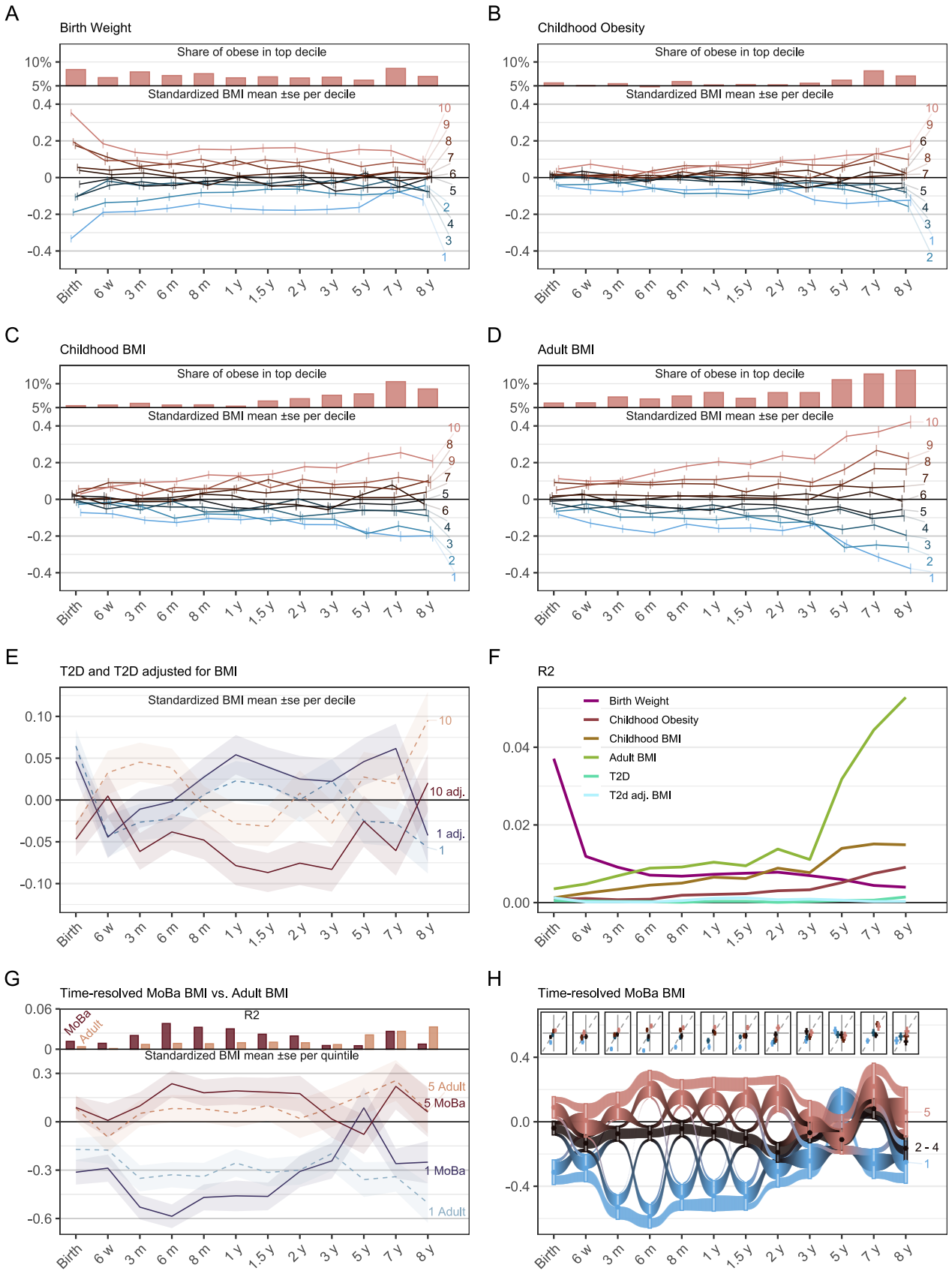
Extended Data Fig. 3 | LD-score regression. Genetic correlation estimate, r_g, of selected traits with early growth BMI at birth, 6 weeks, 3, 6, 8 months, and 1, 1.5, 2, 3, 5, 7, and 8 years of age. Ribbons represent one standard error estimate on each side of the r_g estimate. See methods for details and Supplementary Table 7 for correlation with other traits. See Supplementary Table 1 for the number of samples at each time point.



Extended Data Fig. 4 | Comparison with previous studies on birth weight and adult BMI for the variants near monogenic obesity genes. Heatmap of the effect size for the none top hits near monogenic diabetes genes (see Supplementary Table 2) from birth to adulthood obtained similarly as for Fig. 2A. Note that, in contrast to Fig. 2A, the name of the nearest monogenic obesity gene is used on the y axis, and not the locus name. See Supplementary Table 1 for the number of samples at each time point.

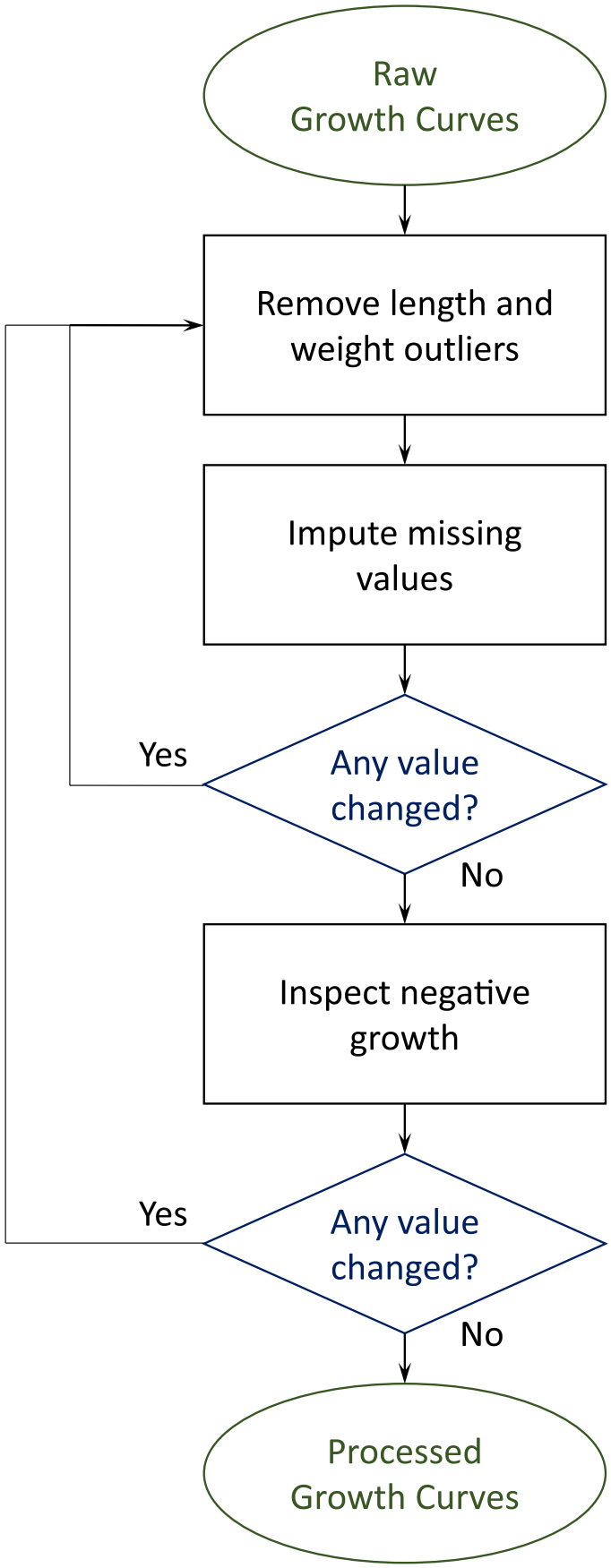


Extended Data Fig. 5 | Loci with multiple independent associations signals. Effect size estimates with child BMI from birth to eight years of age for the lead SNPs of the signals near (A) LEPR, (B) GLP1R, and (C) PCSK1. Dark and light Ribbons represent one standard error estimate on each side of the effect size estimate and 95% confidence intervals, respectively. For each SNP, regional plots are displayed at the age at peak association, highlighting the lead SNPs with red diamonds and SNPs coloured according to the LD R^2, with the exons of the gene according to Ensembl at the bottom and recombination rates in blue. See Supplementary Table 1 for the number of samples at each time point, the methods for the statistical analysis.



Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Polygenic risk score (PRS) analyses. A-D) Mean standardized BMI of children in this study at each time point after stratification in PRS deciles using PRS trained using summary statistics from meta-analyses (bottom), and share of obese children at a given time point in the top PRS decile (top), where obesity is defined as belonging to the top 5 BMI percentile. PRS training was performed using summary statistics for (A) birth weight from Warrington et al[29]. (B) childhood obesity from Bradfield et al[18]. (C) childhood BMI from Felix et al[16]. (D) adult BMI from Yengo et al[15]. (E) Mean standardized BMI for the children in this study falling in the top and bottom deciles of type 2 Diabetes (T2D) risk scores at each time point. PRS for T2D and T2D adjusted for BMI, represented in dashed and solid lines, respectively, were trained using summary statistics from Mahajan et al[67]. (F) R^2 estimated at each time point when training the PRS for Birth weight, childhood obesity, childhood BMI, adult BMI, and T2D in Fig. 5A-E. (G) Mean standardized BMI of children in MoBa that were kept out of the discovery sample falling in the top and bottom quintiles of time-resolved early growth PRSs trained using summary statistics of this study at each time point (solid lines) and of the adult BMI PRS of Fig. 1D (dashed line) (Bottom), along with the respective R^2 estimated when training the PRSs (Top). (H) Mean standardized BMI of children in MoBa that were kept out of the discovery sample falling in the bottom, intermediate, and top quintiles of time-resolved early growth PRSs trained using summary statistics of this study at each time point, in blue, black, and red, respectively. At each time point, rectangles represent one standard error estimate on each side of the mean estimate. Transitions between time points represent the share of children moving from one quintile category to the other. For each time point, mean BMI estimates for these children after stratification in quintiles are plotted for time-resolved early growth PRSs against the adult BMI PRS of Fig. 1D in inserts. All error bars/ribbons represent one standard error estimate on each side of the mean estimate. See Supplementary Table 1 for the number of samples at each time point.



Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Growth curves processing. Length and Weight curves were inspected for outliers and missing values were imputed. This process was repeated until no value was changed. Then length values were inspected for negative growth and adjusted. The entire process was repeated until no value was changed.

nature research

Corresponding author(s): Professor Stefan Johansson
Professor Pål Rasmus Njølstad

Last updated by author(s): Feb 5, 2022

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used.
Data analysis	Genotype calling, Illumina GenomeStudio v.2011.1 and v.2.0.3, commercial. pre-phasing, HRC Imputation preparation tool by Will Rayner version 4.2.5, open source. Phasing, Shapeit v2.790, open source. Phenotype processing, R v. 3.6.1, open source. Genetic association, BOLT-LMM v2.3.4, open source. Trio and haplotype analysis, TrioGen v. 0.5.0, open source. LD score regression, LD Hub v.1.9.0, open source. PRS, PRSice-2 v. 2.3.0, open source. Statistical analysis and figures, R v. 3.6.1, open source, tidyr version 1.1.0, janitor version 2.0.1, conflicted version 1.0.4, glue version 1.4.0, stringr version 1.4.0, dplyr version 1.0.0, scico version 1.1.0, RColorBrewer version 1.1-2, ggplot2 version 3.3.2, ggrepel version 0.8.2, grid version 3.6.1, gtable version 0.2.0, patchwork version 1.1.1, phenoscanner version 1.0, ggfx version 0.0.0.900.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The full GWAS summary statistics for all time points are available for download at www.fhi.no/en/studies/moba/for-forskere-artikler/gwas-data-from-moba. Access to genotypes and phenotypes from the Norwegian Mother, Father and Child Cohort Study (MoBa) is subject to controlled access by the Norwegian Institute of Public Health in accordance with national and international regulations. Conditions of access including contact details for requests can be found at the Norwegian

Institute of Public Health website (fhi.no/en/studies/moba).
HRC or 1000G Imputation preparation and checking: well.ox.ac.uk/~wrayner/tools
Sanger Imputation Service, imputation.sanger.ac.uk
LD Score repository, data.broadinstitute.org/alkesgroup/LDSCORE
GTEx, the Genotype-Tissue Expression portal, gtexportal.org
Birth weight reference data (Warrington et al 201929), http://egg-consortium.org/BW5/Fetal_BW_European_meta.NG2019.txt.gz
Adult BMI reference data (Yengo et al 201815),
http://portals.broadinstitute.org/collaboration/giant/images/c/c8/Meta-analysis_Locke_et_al%2BUKBiobank_2018_UPDATED.txt.gz
Type 2 diabetes (Mahajan et al 201848), https://www.diagram-consortium.org/downloads.html
T2D GWAS meta-analysis - Unadjusted for BMI
Published in in Mahajan et al (2018b)
T2D GWAS meta-analysis - Adjusted for BMI
Published in in Mahajan et al (2018b)
Childhood obesity (Bradfield et al 201918),
http://egg-consortium.org/Childhood_Obesity_2019/CHILDHOOD_OBESITY.TRANS_ANCESTRAL.RESULTS.txt.gz
Childhood BMI (Felix et al 201516),
http://egg-consortium.org/Childhood_BMI/EGG_BMI_HapMap_DISCOVERY.txt.gz
The ALSPAC data dictionary and variable search tool, www.bristol.ac.uk/alspac/researchers/our-data

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	All available samples from MoBa were considered for the study.
Data exclusions	Samples failing QC were excluded, as well as 1) pregnancies strictly shorter than 37 full gestational weeks (259 days); 2) plural pregnancies; 3) ADHD excess cases.
Replication	No replication is possible due to the uniqueness of the cohort.
Randomization	No experimental groups were considered.
Blinding	No experimental groups were considered.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	Children and parents participating in the Norwegian Mother, Father and Child Cohort Study. Biological material and questionnaire data have been collected since the 17th week of pregnancy. The cohort recruitment did not involve discrimination based on sex of the children (49% of females), genetics, past and current health, aiming at an unbiased

	sampling of the Norwegian population. Self-selection biases are similar as to other cohort studies, albeit less prominent than cohorts outside Scandinavia (eg UKBB), and might reduce the prevalence of families with lower socioeconomic status.
Recruitment	The Norwegian Mother, Father and Child Cohort Study (MoBa) is an open-ended cohort study that recruited pregnant women in Norway from 1999 to 2008. Families were invited to participate at the ultrasound offered around week 17 - 19 of pregnancy. Details and invitation letters are available at the Norwegian Institute of Public Health. www.fhi.no/en/studies/moba www.fhi.no/globalassets/dokumenterfiler/trykksaker/invitation-to-participants---moba.pdf
Ethics oversight	The administrative board of the Norwegian Mother, Father and Child Cohort Study led by the Norwegian Institute of Public Health. The Norwegian Data Protection Agency and approval from The Regional Committee for Medical Research Ethics. The Regional Committee for Medical Research Ethics.

Note that full information on the approval of the study protocol must also be provided in the manuscript.