

Genome-wide association and longitudinal analyses reveal genetic loci linking pubertal height growth, pubertal timing and childhood adiposity

Diana L. Cousminer¹, Diane J. Berry⁴, Nicholas J. Timpson⁵, Wei Ang⁷, Elisabeth Thiering⁸, Enda M. Byrne⁹, H. Rob Taal^{10,11,12}, Ville Huikari¹³, Jonathan P. Bradfield¹⁵, Marjan Kerkhof¹⁷, Maria M. Groen-Blokhuis¹⁹, Eskil Kreiner-Møller²⁰, Marcella Marinelli^{21,22,23}, Claus Holst²⁴, Jaakko T. Leinonen¹, John R.B. Perry^{25,26,28}, Ida Surakka^{1,29}, Olli Pietiläinen^{1,31}, Johannes Kettunen^{1,29}, Verner Anttila^{1,31,35,52}, Marika Kaakinen^{13,14}, Ulla Sovio^{32,34}, Anneli Pouta¹⁴, Shikta Das³², Vasiliki Lagou^{26,27}, Chris Power⁴, Inga Prokopenko^{26,27}, David M. Evans⁵, John P. Kemp^{5,6}, Beate St Pourcain⁶, Susan Ring⁶, Aarno Palotie^{1,31,35}, Eero Kajantie^{29,49}, Clive Osmond³⁶, Terho Lehtimäki³⁷, Jorma S. Viikari³⁸, Mika Kähönen³⁹, Nicole M. Warrington⁷, Stephen J. Lye⁴⁰, Lyle J. Palmer^{40,41}, Carla M.T. Tiesler^{8,42}, Claudia Flexeder⁸, Grant W. Montgomery⁹, Sarah E. Medland⁹, Albert Hofman^{10,11}, Hakon Hakonarson^{15,16,43}, Mònica Guxens^{21,22,23}, Meike Bartels¹⁹, Veikko Salomaa²⁹, The ReproGen Consortium[†], Joanne M. Murabito⁴⁴, Jaakko Kaprio^{1,2,30}, Thorkild I.A. Sørensen^{24,45}, Ferran Ballester^{23,46,47}, Hans Bisgaard²⁰, Dorret I. Boomsma¹⁹, Gerard H. Koppelman¹⁸, Struan F.A. Grant^{15,16,43}, Vincent W.V. Jaddoe^{10,11,12}, Nicholas G. Martin⁹, Joachim Heinrich⁸, Craig E. Pennell⁷, Olli T. Raitakari³⁹, Johan G. Eriksson^{3,29,48,51}, George Davey Smith⁵, Elina Hyppönen⁴, Marjo-Riitta Järvelin^{13,14,33,53,54}, Mark I. McCarthy^{26,27,50}, Samuli Ripatti^{1,29,31} and Elisabeth Widén^{1,*} for the Early Growth Genetics (EGG) Consortium[†]

¹Institute for Molecular Medicine, Finland (FIMM), ²Department of Public Health, ³Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, Finland, ⁴Centre for Paediatric Epidemiology and Biostatistics, MRC Centre for Epidemiology of Child Health, UCL Institute of Child Health, London, UK, ⁵The Medical Research Council (MRC) Centre for Causal Analyses in Translational Epidemiology, School of Social and Community Medicine, ⁶School of Social and Community Medicine, University of Bristol, Bristol, UK ⁷School of Women's and Infants' Health, The University of Western Australia, Perth, WA, Australia, ⁸Institute of Epidemiology I, Helmholtz Zentrum München—German Research Center for Environmental Health, Neuherberg, Germany, ⁹Queensland Institute of Medical Research, Brisbane, Australia, ¹⁰The Generation R Study Group, ¹¹Department of Epidemiology, ¹²Department of Pediatrics, Erasmus Medical Center, Rotterdam, The Netherlands, ¹³Institute of Health Sciences, ¹⁴Biocenter Oulu, University of Oulu, Oulu, PO Box 5000, FI-90014, Finland, ¹⁵Center for Applied Genomics, ¹⁶Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, PA, USA, ¹⁷Department of Epidemiology, ¹⁸Department of Pediatric Pulmonology and Pediatric Allergology, GRIAC Research Institute, Beatrix Children's Hospital, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands, ¹⁹Netherlands Twin Register, Department of Biological Psychology, VU University, Amsterdam, The Netherlands, ²⁰Copenhagen Prospective Studies on Asthma in Childhood, Health Sciences, University of Copenhagen, Copenhagen University Hospital, Gentofte, Denmark, ²¹Center for Research in Environmental Epidemiology (CREAL), Barcelona, Catalonia,

*To whom correspondence should be addressed at: Institute for Molecular Medicine Finland (FIMM), Tukholmankatu 8 (PO Box 20), University of Helsinki, Helsinki 00014, Finland. Tel: +358 503812738; Fax: +358 919125737; Email: elisabeth.widen@helsinki.fi

†A full list of members is provided in the Supplementary Material.

Spain, ²²Hospital del Mar Research Institute (IMIM), Barcelona, Catalonia, Spain, ²³CIBER Epidemiologia y Salud Pública (CIBERESP), Barcelona, Catalonia, Spain, ²⁴Institute of Preventive Medicine, Copenhagen University Hospital, Copenhagen, Denmark, ²⁵Genetics of Complex Traits, Exeter medical school, University of Exeter, Exeter, UK, ²⁶Wellcome Trust Centre for Human Genetics, ²⁷Oxford Centre for Diabetes, Endocrinology and Metabolism, Radcliffe Department of Medicine, University of Oxford, Oxford, UK, ²⁸Department of Twin Research and Genetic Epidemiology, King's College London, London, UK, ²⁹Department of Chronic Disease Prevention, ³⁰Department of Mental Health & Substance Abuse Services, National Institute for Health and Welfare, Helsinki, Finland, ³¹The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK, ³²Department of Epidemiology and Biostatistics, School of Public Health, ³³Department of Epidemiology and Biostatistics, MRC Health Protection Agency (HPA) Centre for Environment and Health, School of Public Health, Imperial College, London, UK, ³⁴Department of Non-communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK, ³⁵Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA, ³⁶MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton, UK, ³⁷Department of Clinical Chemistry, Fimlab Laboratories, University Hospital and University of Tampere, Finland, ³⁸Department of Medicine, ³⁹Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku; Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Finland, ⁴⁰Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada, ⁴¹Ontario Institute for Cancer Research, Toronto, Ontario, Canada, ⁴²Division of Metabolic Diseases and Nutritional Medicine, Ludwig-Maximilians-University of Munich, Dr von Hauner Children's Hospital, Munich, Germany, ⁴³Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA, ⁴⁴Department of Medicine, Section of General Internal Medicine, Boston University School of Medicine, Boston, MA, USA, ⁴⁵The Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Copenhagen, Denmark, ⁴⁶Division of Environment and Health, Center for Public Health Research (CSISP), Valencia, Spain, ⁴⁷School of Nursing, University of Valencia, Valencia, Spain, ⁴⁸Unit of General Practice, ⁴⁹Hospital for Children and Adolescents, Helsinki University Central Hospital and University of Helsinki, Helsinki, Finland, ⁵⁰Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford OX3 7LJ, UK, ⁵¹Folkhalsan Research Centre, Helsinki, Finland, ⁵²Analytical and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA ⁵³Unit of Primary Care, Oulu University Hospital, Kajaanintie 50, PO Box 20, FI-90220 Oulu, 90029 OYS, Finland, ⁵⁴Department of Children and Young People and Families, National Institute for Health and Welfare, Aapistie 1, Box 310, FI-90101 Oulu, Finland

Received October 12, 2012; Revised and Accepted February 25, 2013

The pubertal height growth spurt is a distinctive feature of childhood growth reflecting both the central onset of puberty and local growth factors. Although little is known about the underlying genetics, growth variability during puberty correlates with adult risks for hormone-dependent cancer and adverse cardiometabolic health. The only gene so far associated with pubertal height growth, *LIN28B*, pleiotropically influences childhood growth, puberty and cancer progression, pointing to shared underlying mechanisms. To discover genetic loci influencing pubertal height and growth and to place them in context of overall growth and maturation, we performed genome-wide association meta-analyses in 18 737 European samples utilizing longitudinally collected height measurements. We found significant associations ($P < 1.67 \times 10^{-8}$) at 10 loci, including *LIN28B*. Five loci associated with pubertal timing, all impacting multiple aspects of growth. In particular, a novel variant correlated with expression of *MAPK3*, and associated both with increased prepubertal growth and earlier menarche. Another variant near *ADCY3-POMC* associated with increased body mass index, reduced pubertal growth and earlier puberty. Whereas epidemiological correlations suggest that early puberty marks a pathway from rapid prepubertal growth to reduced final height and adult obesity, our study shows that individual loci associating with pubertal growth have variable longitudinal growth patterns that may differ from epidemiological observations. Overall, this study uncovers part of the complex genetic architecture linking pubertal height growth, the timing of puberty and childhood obesity and provides new information to pinpoint processes linking these traits.

INTRODUCTION

Postnatal height growth is a heritable complex process characterized by rapid infant growth, slowly diminishing mid-childhood growth and a distinct pubertal height growth spurt. Whereas the genetics of adult stature has been evaluated in large-scale genome-wide association (GWA) analyses (1), few studies have addressed the molecular underpinnings of distinct growth phases. Moreover, specific growth patterns during childhood correlate with both altered pubertal timing and adult health risks. For example, increased height and body mass index (BMI) prior to puberty correlate with advanced pubertal onset (2–4), and early puberty associates with increased risk of adult obesity and related metabolic traits (5–7). Still, the specific mechanisms linking these traits remain elusive.

To elucidate part of the genetic architecture impacting adolescent growth, we focussed primarily on the dynamic and highly variable pubertal growth spurt that reflects both the activation of central puberty and local growth factors (8,9) while accounting for up to 15–20% of adult stature (10). Narrow-sense heritability estimates place the genetic contribution to variation in pubertal growth between 60 and 90% (11–13), and twin studies suggest a substantial proportion of shared genetic variance with other phases of childhood growth (13). Specifically, we aimed to (i) identify genetic variants associated with the onset, total magnitude and tail end of the pubertal growth spurt and (ii) investigate these variants' longitudinal effects on overall childhood growth and the timing of puberty (study design outlined in Fig. 1).

Due to large variation in the timing and rate of the pubertal growth spurt (shown schematically in Fig. 2), an accurate model typically requires frequent height measurements spanning a large age range, often difficult to obtain. Furthermore, girls enter puberty and, thus, begin their growth spurt, an average of 2 years earlier than boys. Taking these challenges into consideration, we aimed to characterize loci influencing growth during puberty by leveraging heterogeneous height measurements taken at varied ages throughout childhood across participating cohorts to maximize statistical power. Therefore, we modelled the pubertal height growth spurt for GWA using three partially correlated simple measures (Supplementary Material, Table S1; Fig. 2) that also partly reflect the timing of puberty (14). In Analysis I, we targeted the take-off phase of the growth spurt [height standard deviation score (SDS) at 10 years in girls and 12 years in boys] by reasoning that increased height relative to the population mean in early puberty reflects either overall genetic height potential or entrance into the pubertal growth spurt. Because a large proportion of adult stature is achieved prior to the onset of puberty, we expected a significant part of the detected variants to associate with overall height growth potential, whereas a minority would have specific pubertal timing effects. In Analysis II, we assessed the overall contribution of growth across puberty to adult height (height change SDS between 8 years and adult) that reflects the total magnitude of growth during the pubertal growth spurt. Finally, in Analysis III, we approximated the timing of peak height velocity by looking at the height change SDS between age 14 years

and adult because early maturing individuals grow less during late adolescence than late-maturing individuals, who still have much of their remaining growth to achieve after age 14. Similar simple height measurements across puberty have previously proven robust in the GWA setting for detecting common genetic variation influencing both height growth and pubertal timing (15).

RESULTS

Discovery and follow-up meta-analyses reveal 10 genome-wide significant loci associated with pubertal growth

Nine cohorts contributed partly overlapping population-based samples (Supplementary Material, Table S2) with childhood height measurements and approximately 2.5 million directly genotyped or imputed SNPs to three discovery GWA analyses (Supplementary Material, Table S3), in which we meta-analysed data from males and females both separately and combined for the three models. We observed significant deviation from the expected distribution of *P*-values for all three combined-gender analyses (I, II and III), males and females separately for Analysis I and females only for Analysis II (Supplementary Material, Fig. S1A).

All three models resulted in genome-wide significant loci, although we had most power (Supplementary Material, Table S4) to detect loci for Analysis I (height SDS at age 10 years in girls and 12 years in boys) (Table 1). In total, nine loci contained markers that reached *P*-values below the genome-wide significance threshold corrected for testing three primary phenotypes ($P < 1.67 \times 10^{-8}$, after genomic control). Of these, only rs7759938 nearby *LIN28B* was previously known to influence pubertal growth (15,16).

Due to the requirement of Analyses II and III to have both childhood and adult height measurements for the same individuals, there were no additional samples available for follow-up of suggestive signals for these analyses. Thus, we only performed follow-up for Analysis I. An additional 6 cohorts comprising up to 9710 samples were available for follow-up of the 22 suggestive signals ($P < 1 \times 10^{-5}$) for Analysis I (Supplementary Material, Table S5). Joint analysis of discovery and follow-up stages for Analysis I robustly confirmed a single novel variant, rs4788196 ($P = 9.49 \times 10^{-11}$, $n = 18\,737$; Table 1), thus bringing the number of loci reaching the genome-wide significance threshold to 10, of which 7 were associated with Analysis I.

Expression quantitative trait loci (eQTL) analysis links rs4788196 (G) to decreased expression of nearby gene *MAPK3* and pathway analyses highlight the TGF-beta signalling pathway and pathways in cancer

To link the identified association signals with putative biological processes, we tested all significantly associated gene regions for association with leukocyte gene expression levels and performed gene pathway analyses. Expression quantitative trait loci (eQTL) analysis in whole blood (17) linked rs960273 with the gene *GNA12*, as previously reported (1), as well as highlighting a previously unknown

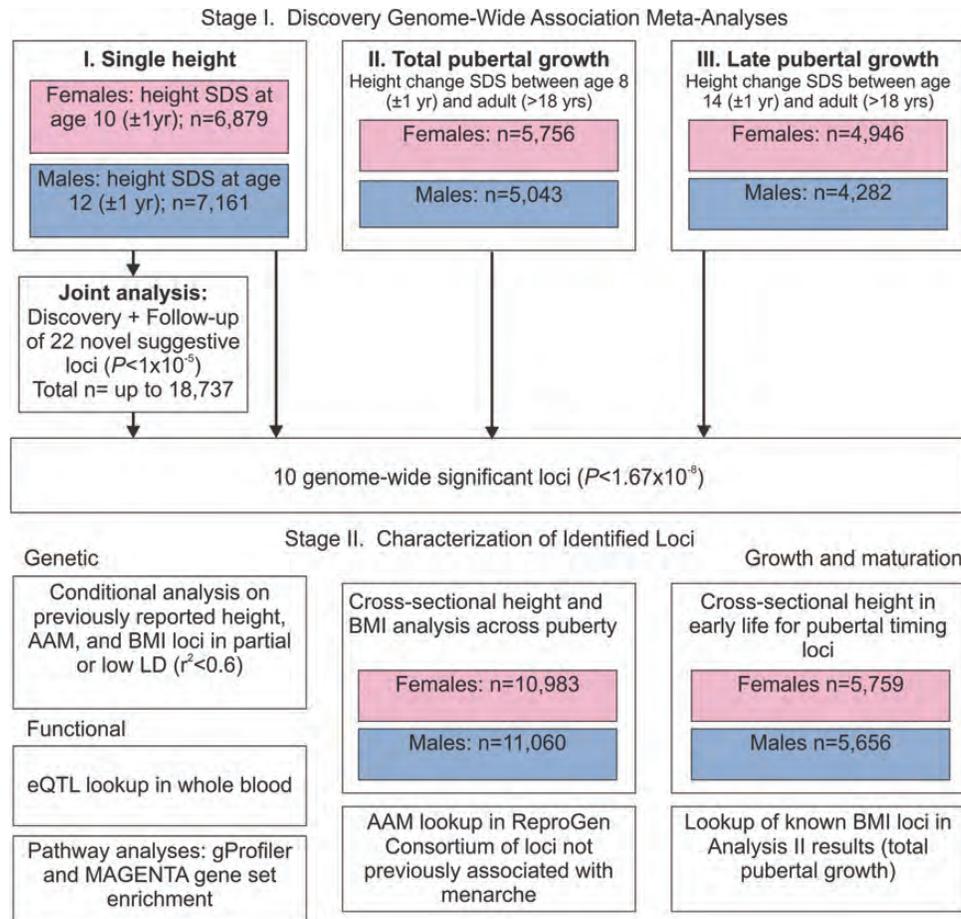


Figure 1. Study design. We performed a two-stage study to detect and characterize loci influencing the pubertal phase of childhood growth. Stage 1 consisted of locus mapping using a GWA approach on three related simple measurements of the pubertal growth spurt and joint analysis of discovery and follow-up studies of novel variants for height SDS at age 10 years in girls and 12 years in boys (Analysis I). Validated loci were then characterized using a variety of methods in Stage 2, including genetic characterization (conditional analysis on previously reported nearby SNPs in low or partial pairwise linkage disequilibrium with pubertal growth signals), functional characterization (eQTL analysis of pubertal growth SNPs on the expression levels of nearby genes and pathway analysis on biological pathways using g:Profiler and MAGENTA), characterization of the growth and maturational effects of the identified loci (on height and BMI across puberty, on the timing of menarche for those signals not previously associated with AAM and on early height from 1 to 4 years for loci that influence the timing of menarche and pubertal growth) and the association between BMI-increasing alleles and total pubertal growth.

role for the extracellular signal-regulated kinase 1 (*MAPK3*, also known as *ERK1*) in prepubertal height growth (Supplementary Material, Table S6). More specifically, the adolescent height-increasing allele (G) at rs4788196 on 16p11.2 (Analysis I) correlated with decreased expression of *MAPK3*, consistent with previous studies linking deactivation of the gene with increased bone growth in mice (18). We subsequently performed pathway analyses using the g:Profiler Gene Group Functional Profiler tool [g:GOS (19); Supplementary Material, Table S7A] and MAGENTA Gene Set Enrichment Analyses [GSEA (20); Supplementary Material, Table S7B] that commonly highlighted the TGF-beta signaling pathway and pathways in cancer for loci identified in Analysis I. Whereas g:Profiler identified the MAPK-pathway, the GSEA showed enrichment of lower than expected *P*-values for genes belonging to the *TOBI* pathway, although the individual implicated gene regions were only suggestively associated in Analysis I.

The novel locus near *MAPK3* associates transiently with height growth in childhood and earlier menarche

Although there are no published studies implicating *MAPK3* in human height growth, rare recurrent CNVs near *MAPK3* on chromosome 16p11.2 have been shown to associate with early onset obesity (21,22). Nonetheless, the adolescent height effect that we observed did not appear to be CNV-mediated (Supplementary Material, Table S8). To characterize the *MAPK3* variant in more depth, we evaluated the longitudinal height and BMI effects of rs4788196. We plotted the effect size (beta) against six age bins across puberty from 8 years to adult (Fig. 3A) and investigated early height yearly from ages 1 to 4 (Supplementary Material, Table S9). These analyses revealed a transient effect on height growth for the G allele from age 4 in both males and females that was diluted by adulthood, with no apparent effect on BMI (Supplementary Material, Fig. S2). Finally, because rapid

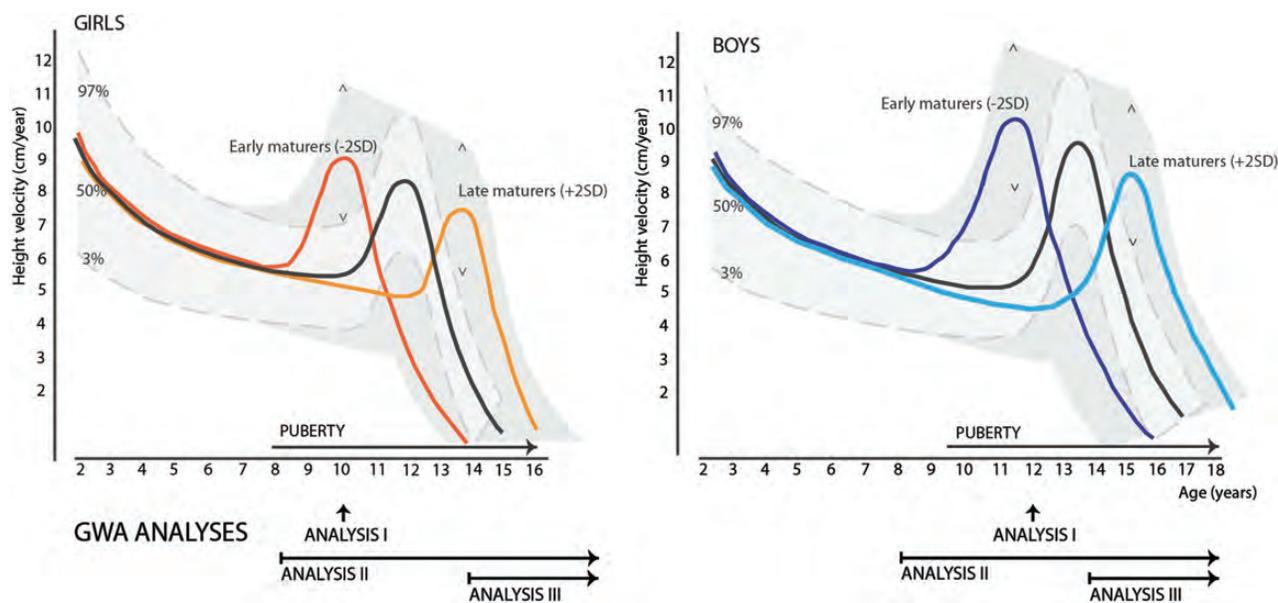


Figure 2. Schematic picture of postnatal height and the three partly correlated GWA phenotypes describing pubertal growth. Childhood and pubertal growth rates at the 3rd, 50th and 97th percentile are shown for girls in the left panel and boys in the right panel. Growth rates during puberty vary as a consequence of variable timing of the growth spurt. The black growth curves illustrate a growth pattern representing the mean timing of the pubertal growth spurt, whereas two SDs early (-2 SD) and late ($+2$ SD) timing of the pubertal growth spurt are shown in dark versus light orange in girls and dark versus light blue in boys. The three genome-wide analysis strategies are illustrated in the bottom of each panel. Analysis I aims at capturing the take-off phase of the pubertal growth spurt and includes a single height measurement relative to the population mean (height SDS) at age 10 years in girls and 12 years in boys. Analyses II measures the change in relative height between age 8 years and adulthood and targets the total magnitude of pubertal growth. Analysis III, measuring the relative height change between age 14 years and adult, targets the late end of the pubertal spurt in height growth that varies depending on the timing of peak height velocity of the growth spurt.

growth during childhood and early adolescence may correlate with early timing of sexual maturation (8), we also tested rs4788196 for association with age at menarche (AAM) (23) and found that the height-increasing allele associated with earlier AAM ($P = 1.42 \times 10^{-4}$), surpassing the significance threshold of 0.007, that corresponds to a Bonferroni-correction accounting for follow-up of seven loci not previously associated with AAM; Supplementary Material, Table S10).

Five of the discovered pubertal growth loci are also associated with pubertal timing, eight are adult stature loci and one is a BMI locus

Even though the *MAPK3*-locus associated with both prepubertal height growth and the timing of puberty, it showed little evidence for association with adult anthropometric traits (Table 2). Nonetheless, epidemiological data support phenotypic correlations between earlier pubertal timing, increased adult obesity and decreased final height. To better understand how the loci we detected contribute to a genetic link underlying these traits, we performed a systematic analysis of all leading SNPs significantly associated with pubertal height and growth and noted substantial evidence for overlap (Fig. 4). Assessing the pubertal timing effect, both based on published GWA of AAM (23) and by *in silico* meta-analysis of leading signals not previously implicated in the timing of menarche, as described above, showed that all three discovery analysis approaches detected pubertal timing loci (near *MAPK3*, *PXMP3*, *VGLL3*, *ADCY3-POMC* and *LIN28B*). Moreover,

eight signals overlapped with adult stature loci (1), of which *ADCY3-POMC* has further been implicated in childhood (24) and adult obesity (25). However, two of our signals within 1 Mb of reported height loci showed partial linkage disequilibrium ($r^2 < 0.6$) with previously published SNPs. Conditioning for the previously reported marker revealed evidence for a further independent association nearby *CABLES1* (rs6507528; $P = 0.00011$) (Supplementary Material, Table S11). This variant may represent allelic heterogeneity or partially tag the same causative variant (26).

Longitudinal analyses across puberty show that the identified pubertal growth loci represent both overall growth potential and pubertal timing

To further evaluate the leading signals, we compared their height effects longitudinally across puberty, revealing multiple distinct growth trajectories. This approach divided the loci associated with various measures of pubertal height and growth into two groups based on association with pubertal timing. One group of loci (near *ZBTB38*, *EFEMP1*, *CABLES1*, *ADAMTSL3* and *GNAI2*), not associated with pubertal timing, all impacted height SDS across multiple growth phases, strongly and steadily from prepuberty to adulthood (Fig. 3C). Thus, these loci likely reflect overall growth potential, rather than puberty-specific effects. In contrast, the five pubertal timing-associated variants displayed diverse effects on the timing and tempo of growth, both before and during puberty (Table 2; Fig. 3A and B).

Table 1. Pubertal growth loci reaching genome-wide significance ($P < 1.67 \times 10^{-8}$)

SNP	Nearby gene(s)	Previously associated related trait ^a	Chromosome	Position (bp) ^b	Effect allele/ other allele ^b	Effect allele frequency	Discovery Relative height change β (SE) ^c	P^d	P_{het}	$P_{\text{sex-het}}^c$	Follow-up Relative height change β (SE) ^c	P	Joint Discovery + Follow-up Relative height change β (SE) ^c	P	n
Analysis I: single height															
<i>Height SDS at age 10 in females and age 12 in males combined</i>															
rs6764769	ZBTB38, RASA2	H	3	142 582 970	G/A	0.45	0.08 (0.012)	4.6×10^{-10}	0.52	0.35					13 960
rs7846385	PXMP3, PKIA	H, AAM	8	78 322 734	C/T	0.26	0.09 (0.014)	5.27×10^{-10}	0.13	0.13					13 942
rs1346789	EFEMP1	H	2	55 945 556	C/T	0.22	-0.08 (0.015)	1.15×10^{-8}	0.41	0.85					13 960
rs6507528	CABLES1	H	18	19 025 578	G/A	0.55	-0.09 (0.016)	1.31×10^{-8}	0.44	0.2					13 160
rs1365198	ADAMTSL3	H	15	82 189 907	G/T	0.76	0.08 (0.014)	1.5×10^{-8}	0.66	0.41					13 946
rs4788196	MAPK3	Novel	16	29 874 935	G/A	0.44	0.06 (0.012)	6.38×10^{-7}	0.51	0.43	0.08 (0.021)	4.25×10^{-5}	0.07 (0.011)	9.49×10^{-11}	18 737
<i>Height SDS at age 12 in males only</i>															
rs960273	GNA12	H	7	2 824 402	C/T	0.32	-0.1 (0.018)	2.51×10^{-8}	0.96	5.17×10^{-4}					6986
Analysis II: pubertal growth															
<i>Height change SDS (8 adults) in females and males combined</i>															
rs1172294	ADCY3, DNAJC27, POMC	H, BMI	2	25 022 704	G/A	0.45	-0.08 (0.014)	1.02×10^{-8}	0.87	0.92					10 799
<i>Height change SDS (8 adults) in females only</i>															
rs7628864	VGLL3	AAM	3	86 933 308	G/A	0.38	-0.11 (0.019)	3.17×10^{-9}	0.76	6.83×10^{-6}					5756
Analysis III: late pubertal growth															
<i>Height change SDS (14 adults) in females and males combined</i>															
rs7759938	LIN28B	H, AAM	6	105 485 647	C/T	0.32	0.11 (0.016)	3.87×10^{-9}	0.23	0.23					8863

^aPreviously associated related traits are adult stature (H), AAM or BMI.

^bMarker position reported according to Build 36 and allele coding based on the positive strand.

^cEffect sizes are change in height or growth SDS score.

^d P -value adjusted for genomic control.

^e P -value assessed by t -test for sexual heterogeneity.

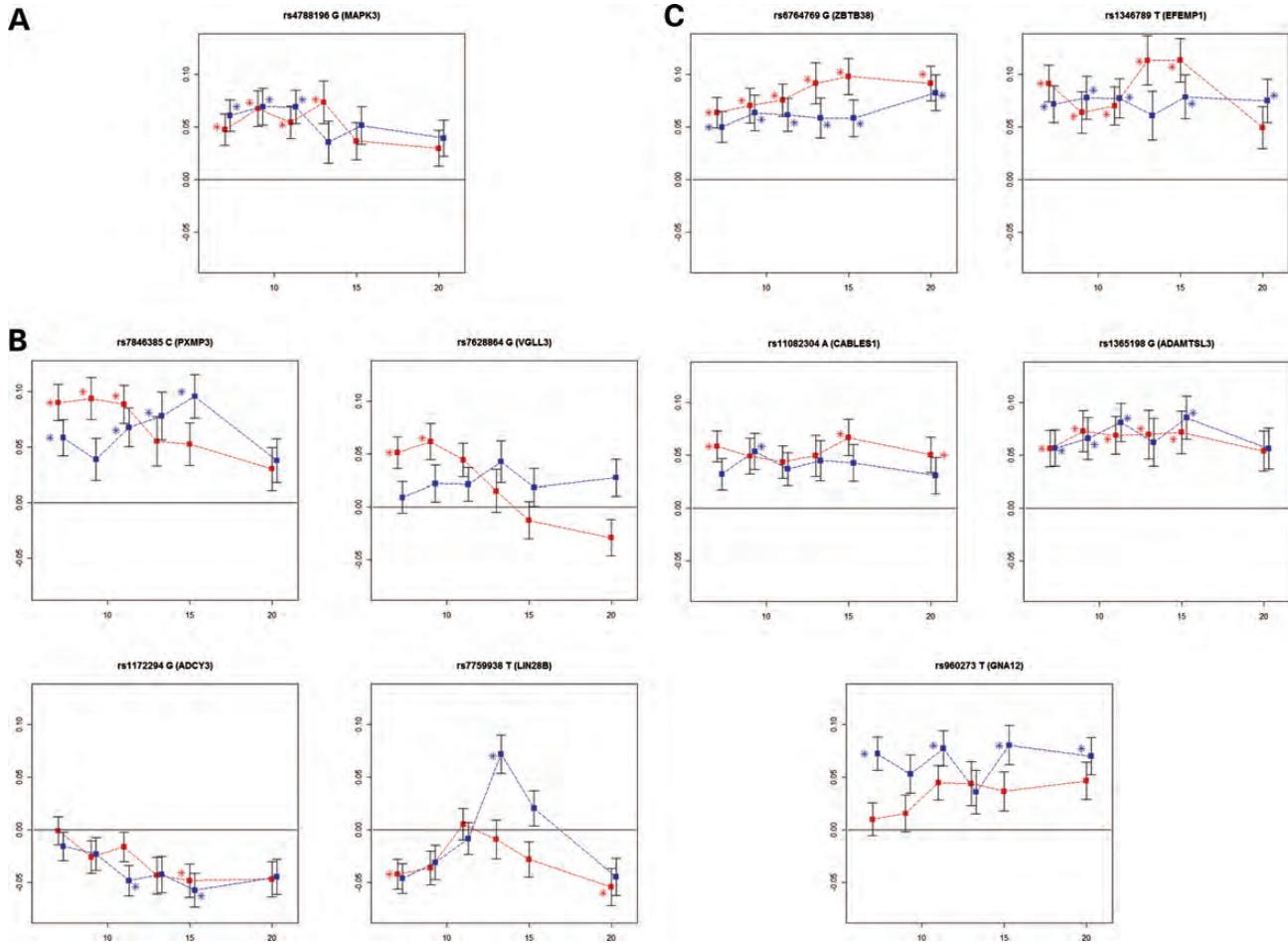


Figure 3. Height across multiple growth periods for 10 pubertal growth loci. The effect size of linear regression of height SDS against genotype at six age bins from prepuberty to adulthood was plotted longitudinally across adolescence for males (blue) and females (red). (A) *MAPK3* (B) Pubertal growth loci not associated with the timing of menarche are shown for the height-increasing allele. All these variants are also associated with adult stature. (C) Pubertal growth loci associated with pubertal timing, shown for the menarche-advancing allele. These loci show divergent association with prepubertal growth. The x-axis represents age and the y-axis is the effect size (beta) of the association between the indicated allele and relative height at each age. * $P < 0.002$.

The locus associating with increased childhood BMI also associates with decreased pubertal growth but not with prepubertal height

Of the pinpointed loci associating with pubertal growth and timing, only one had previously been associated with BMI (Supplementary Material, Fig. S2). The strong correlation between childhood obesity and prepubertal height (2–4) and between prepubertal height and AAM (27) predicts that the BMI-associated marker, rs1172294 (*ADCY3-POMC*), would associate with increased prepubertal stature. However, the variant showed no association with stature before puberty (Fig. 3B). Nonetheless, the BMI-increasing allele (G) was associated with earlier menarche ($P = 8.64 \times 10^{-6}$) and a decline in pubertal growth in both males and females, as expected. Consistently, other variants previously associated with childhood obesity (24) showed a parallel between elevated BMI and diminished growth across puberty (Supplementary Material, Table S12). We also found that rs3817334

(*MTCH2*), previously associated with adult (25) but not childhood BMI, also associated with the same decrease in overall pubertal growth.

DISCUSSION

Taken together, the simple approach used in this study to model the pubertal growth spurt for GWAS in more than 18 000 study subjects of European descent identified 10 significantly associated pubertal growth loci. Utilizing unique longitudinal childhood measurements, we described the distinct height growth and BMI effects of these variants across puberty and noted significant longitudinal associations at each locus. More specifically, half of the identified loci also associated with pubertal timing and provided evidence linking a robust novel growth and menarche locus with *MAPK3* expression levels. Despite prior association between several genes in the MAPK-pathway and skeletal growth

Table 2. Association of pubertal growth variants with height growth across puberty

SNP (allele) nearby gene	Prepubertal height Height SDS at 7–8 years			Height at early puberty Analysis I height SDS at 10 years in females and 12 years in males			Total pubertal growth Analysis II height SDS change 8 years in adult			Late adolescent growth Analysis III height SDS change 14 years in adult			Adult height ^b Adult stature SDS			GIANT height ^c <i>P</i> , <i>N</i>
	Beta	SE	<i>P</i>	Beta	SE	<i>P</i> ^a	Beta	SE	<i>P</i> ^a	Beta	SE	<i>P</i> ^a	Beta	SE	<i>P</i>	
rs7846385 (C) <i>PXMP3</i>																
Females	0.090	0.016	3.73 × 10⁻⁸	0.096	0.018	8.71 × 10⁻⁸	-0.057	0.021	0.006	-0.043	0.023	0.068	0.03	0.019	0.114	
Males	0.058	0.016	0.031	0.067	0.019	5.46 × 10⁻⁴	-0.009	0.009	0.696	-0.068	0.026	0.008	0.038	0.02	0.055	
Combined	0.074	0.012	1.48 × 10⁻¹⁰	0.082	0.013	3.04 × 10⁻¹⁰	-0.035	0.015	0.022915	-0.054	0.017	0.002	0.034	0.014	0.014	2.988 × 10⁻¹⁰ , 133 772
rs7628864 (G) <i>VGLL3</i>																
Females	0.051	0.015	0.001	0.044	0.016	0.007	-0.112	0.019	3.17 × 10⁻⁹	-0.059	0.021	0.006	-0.029	0.017	0.087	
Males	0.009	0.015	0.549	0.023	0.018	0.201	0.011	0.020	0.578	-0.001	0.022	0.969	0.028	0.017	0.113	
Combined	0.030	0.011	0.004	0.035	0.012	0.004	-0.054	0.014	9.41 × 10⁻⁵	-0.031	0.015	0.043	0.00	0.012	0.967	0.6118, 133 795
rs7759938 (T) <i>LIN28B</i>																
Females	-0.042	0.014	0.003	0.004	0.017	0.792	-0.027	0.020	0.169	-0.112	0.022	2.89 × 10⁻⁷	-0.054	0.018	0.002	
Males	-0.046	0.014	0.001	-0.011	0.018	0.547	-0.020	0.021	0.327	-0.074	0.023	0.002	-0.045	0.018	0.011	
Combined	-0.044	0.010	8.80 × 10⁻⁶	-0.003	0.012	0.804	-0.024	0.014	0.097	-0.094	0.016	3.87 × 10⁻⁹	-0.05	0.013	1.0 × 10⁻⁴	8.691 × 10⁻¹⁸ , 133 774
rs4788196 (G) <i>MAPK3</i>																
Females	0.047	0.015	0.001	0.059	0.017	8.72 × 10⁻⁴	-0.0003	0.018	0.958	-0.025	0.021	0.231	0.03	0.017	0.081	
Males	0.061	0.015	4.51 × 10⁻⁵	0.066	0.017	1.49 × 10⁻⁴	0.003	0.02	0.867	-0.046	0.023	0.04	0.039	0.017	0.023	
Combined	0.054	0.011	3.56 × 10⁻⁷	0.063	0.012	6.38 × 10⁻⁷	0.001	0.014	0.921	-0.035	0.015	0.023	0.035	0.012	0.004	0.003336, 133 818
rs1172294 (G) <i>ADCY3</i>																
Females	-0.001	0.013	0.945	-0.006	0.017	0.718	-0.078	0.019	3.16 × 10⁻⁵	-0.046	0.021	0.028	-0.047	0.017	0.005	
Males	-0.016	0.014	0.246	-0.045	0.017	0.01	-0.08	0.02	5.97 × 10⁻⁵	-0.01	0.023	0.673	-0.045	0.017	0.008	
Combined	-0.008	0.01	0.404	-0.026	0.012	0.04	-0.079	0.014	1.02 × 10⁻⁸	-0.02	0.015	0.184	-0.046	0.012	1.3 × 10⁻⁴	1.7 × 10⁻¹³ , 133 780

^a*P*-values taken from GWA discovery analyses, genomic-control corrected.

^bMeta-analysis of association results for adult stature taken from the cohorts participating in the longitudinal analyses.

^cAssociation to adult stature taken from publically available results of the GIANT Consortium adult height meta-analysis (http://www.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files.) Beta estimates were not available for public download.

All effect sizes are given for the menarche-advancing allele. In bold are associations reaching a *P*-value threshold of 0.002 (Bonferroni-corrected threshold accounting for follow-up analysis of five markers in five analyses).

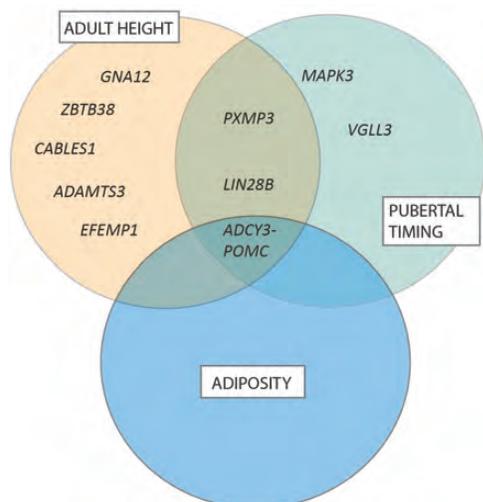


Figure 4. Loci associated with pubertal height and their overlap with partially correlated phenotypes. The 10 genome-wide significant loci were assessed for their association with the correlated traits adult stature, pubertal timing (AAM) and adiposity (BMI) by examining previously published literature. Additionally, for loci not previously associated with pubertal timing, we queried the leading SNPs in GWA data of AAM by the ReproGen Consortium. All the pubertal growth loci showed pleiotropic associations with one or more related phenotype.

syndromes (28–31), *MAPK3* has not been associated with human height before. Interestingly, gonadotropin-releasing hormone, crucial for regulating the onset of puberty, activates *MAPK3* (32), providing a putative biological link between rs4788196 and pubertal timing.

This study has several strengths. It is based on a large dataset of rather unique longitudinal height data from multiple well-characterized study cohorts. The two-stage approach applied in this study enabled wide-ranging characterization of growth and maturation phenotypes associated with the 10 leading association signals. Nonetheless, one limitation is that the height data available for analysis varied among the cohorts. Furthermore, the height measurements available in most of the cohorts were not frequent enough to allow detailed modelling of the pubertal growth spurt. To overcome the lack of very frequent height measurements and the variability of height assessments between cohorts, we chose to adopt an analysis strategy aiming to maximize the number of study subjects. By utilizing three simple and robust height growth estimates to model the pubertal growth spurt, in addition to applying rigorous statistical significance thresholds, we were successfully able to identify and characterize novel loci significantly associated with pubertal height growth. As expected, a proportion of these loci also associated with pubertal timing, as assessed by age of menarche, and with adult height.

In fact, our data affirm a complex genetic architecture underlying growth, pubertal timing and adiposity. In particular, specific genetic effects may contradict epidemiological correlations. Epidemiological studies have observed a developmental pattern linking taller prepubertal stature to earlier puberty, accelerated skeletal maturation and short adult stature due to early cessation of growth (2,3). Although the majority of loci we assessed showed the expected parallel association between early menarche and decreased overall

pubertal height growth, their prepubertal height effects varied. Three variants (near *MAPK3*, *PXMP3* and *VGLL3*) followed the expected pattern, linking taller prepubertal stature with earlier AAM, whereas the early puberty-associated allele (T) at rs7759938 (*LIN28B*) correlated with shorter prepubertal childhood height, as reported previously (15).

The relationship between puberty and adult stature is similarly complex, whereas epidemiological studies show a correlation between early puberty and reduced adult height (3), and a genetic association study found that early puberty alleles may associate with either increased or decreased adult stature (23). An example is rs7846385 (*PXMP3*), for which, contradictory to the predicted pattern, the early menarche allele associates with increased adult height. Our results show that tall adult height is achieved because the early-menarche allele (C) also influences tall childhood height and a limited reduction in total pubertal growth. These data, thus, agree with a recent candidate gene study suggesting that loci associated with adult height may have a stronger influence on prepubertal growth than during the pubertal growth spurt (16). However, utilizing a genome-wide approach with greater sample sizes, our study identifies loci previously missed that specifically target pubertal growth, and we find that they are associated with diverse and unique longitudinal growth patterns. We also find that not all loci influencing pubertal growth also impact adult stature.

Additionally, epidemiological studies link increased childhood adiposity with advanced puberty and increased prepubertal height. Although all childhood BMI-increasing alleles assessed in this study also showed an association with decreased overall pubertal growth, at the *ADCY3-POMC* locus, the same allele associated with both earlier puberty and increased childhood BMI, but not with prepubertal stature. The correlation between obesity and pubertal growth may be consequential of hormonal changes associated with childhood adiposity. However, because the same association pattern was also present at a locus uniquely associated with adult BMI (*MTCH2*), an underlying shared genetic effect remains likely.

Given the complexity of the relationships between these developmental traits, tracking unique gene effects across multiple growth periods may help to elucidate specific pathways linking childhood events to adult outcomes, as illustrated here with height growth, pubertal timing and adult stature. While epidemiological studies have described correlations between distinct childhood growth events and adult health, genetically defined association patterns may pinpoint molecular processes linking these traits. Characterization of these pathways may thus provide new insight towards a better understanding of the relationships between early growth patterns, pubertal timing and adult disease risk.

MATERIALS AND METHODS

Phenotypes and study subjects

Discovery study subjects were included from cohorts participating in the Early Growth Genetics Consortium (43), namely the Avon Longitudinal Study of Parents and Children (ALSPAC), 1958 British Birth Cohort (BC58-T1DGC and

BC58- WTCCC), Cardiovascular Risk in Young Finns Study (YFS), Helsinki Birth Cohort Study (HBCS), Lifestyle-Immune System-Allergy Plus Environment and Genetics Study (LISApus), Northern Finland Birth Cohort 1966 (NFBC1966), Queensland Institute of Medical Research and Western Australia Pregnancy Study (RAINE). Cohort-specific details for all analyses can be found in Supplementary Material, Table S3. The data annotation, exchange and storage have been facilitated by the SIMBioMS platform (33).

Three primary phenotypes were analysed that were defined as follows.

Analysis I: single height: girls with height measurements available at age 10 (± 1 year) and boys with height measured at age 12 (± 1 year) were included. Sex-specific SDSs for each individual were calculated within each study by dividing the difference between the individual's measured height and the within-population height mean by the population standard deviation (SD). A total of 14 040 samples (7161 males and 6879 females) from 9 contributing cohorts were included.

Analysis II: total pubertal growth: individuals with a childhood height measurement at age 8 (± 1 year) and at adulthood (≥ 18 years of age) were included. Height difference was calculated between the two measurements, and sex-specific SDSs of this difference were calculated within each study as described above. Six cohorts with up to 10 799 samples (5043 males and 5756 females) contributed to the analysis.

Analysis III: late pubertal growth: subjects with a height measurement in adolescence at age 14 (± 1 year) and at adulthood (≥ 18 years of age) were included. Height difference was calculated between the two measurements, and logarithm-transformed sex-specific Z-scores were calculated within each study. Log transformation was performed prior to SDS calculation. Five cohorts with up to 9228 subjects (4282 males and 4946 females) were included in this analysis.

Genotyping and quality control

Genome-wide genotypes were obtained using high-density SNP arrays on Illumina and Affymetrix platforms. Before imputation, SNPs with minor allele frequency of $< 1\%$, call rate $< 95\%$ or Hardy-Weinberg equilibrium $P < 1 \times 10^{-6}$ were excluded. Samples were also excluded if they contained duplicates, excess heterozygosity, non-European ancestry or ambiguous gender. Imputation was performed using IMPUTE (34) or MACH (35) for roughly 2.5 million SNPs against HapMap Phase II (release 21 of 22). Imputed SNPs were filtered prior to meta-analysis to exclude poorly imputed SNPs (IMPUTE filter PROPER INFO < 0.4 , MACH filter $r^2 < 0.3$).

GWA analyses

Within each cohort, association analyses were performed by linear regression using an additive model across genotyped and imputed SNPs (dosages), for males and females separately. For all analyses, age at adolescent measurement was included as a covariate where available, and the first two principal components were adjusted within each study sample if necessary. For the association tests, PLINK (36), ProbABEL (38), SNPtest (34) or MACH2QTL (35) was used.

Meta-analyses

A fixed effects inverse-variance meta-analysis model was used to test the effect of each variant on height, total pubertal growth or late pubertal growth separately for males and females. Sex-specific results from each study were also meta-analysed for each phenotype in three combined-gender analyses. The R package MetABEL (37) (v.2.11.1) was used to perform all meta-analyses. MetABEL corrects each individual result for its respective genomic inflation factor (λ) according to the genomic control method for population stratification. Subsequently, an additional genomic control correction was applied using the overall genomic inflation factor calculated for each of the nine meta-analysed results. The threshold for genome-wide significance was set at a conservative Bonferroni-corrected threshold of $P < 1.67 \times 10^{-8}$, accounting for testing three primary phenotypes. A further significance threshold of 0.002 (accounting for examining 10 loci in males and females) was applied to all follow-up analyses unless otherwise stated.

Conditional analyses

To determine whether our signals represent independent effects on growth during puberty from previously reported related phenotypes, we performed linear regression using an additive model on the primary pubertal growth phenotypes, adjusting each of the six markers (imputed genotype dose) for the previously reported marker nearby our signal. As in the primary analysis, age at adolescent measurement (where available) and optional adjustment for population substructure were included as covariates (Supplementary Material, Table S3).

Follow-up analyses of suggestive association signals

Genetic markers yielding association P -values of 1×10^{-5} to 1.67×10^{-8} in Analysis I and not previously associated with related traits adult stature, AAM or BMI were selected for follow-up genotyping ($n = 22$). Additional cohorts participated in the follow-up analyses, including *in-silico* analyses by ALSPAC (follow-up sample), Children's Hospital of Philadelphia (CHOP), Finnish Twin Cohort Study, Genome-Wide Population-Based Association Study of Extremely Overweight Young Adults and Lifestyle-Immune System-Allergy study & German Infant Study on the influence of Nutrition Intervention plus environment and genetics (GINIplus; follow-up sample). Association results for a marker showing borderline significance, rs281379, were also provided by Netherlands Twin Registry (NTR). *De novo* genotyping was done for selected markers (success rate $> 98\%$) from Northern Finland Birth Cohort 85-86 (NFBC8586) with TaqMan Pre-Designed SNP Genotyping Assays on LightCycler 480 Real-Time PCR System (Roche) according to the manufacturer's instructions at the Finnish Genome Center (Helsinki, Finland).

Statistical analysis in replication samples was performed similarly as in the discovery analyses with PLINK (36), ProbABEL (38) or SNPtest (34), using linear regression models for each of the 22 markers under an additive model, with age at adolescent measurement and correction for population

substructure as optional covariates. Genomic control-corrected discovery results were meta-analysed together with the individual linear regression results from contributing cohorts for each SNP, using the MetABEL (37) package of R (v.2.11.1).

CNV analysis of 16p11.2

CNVs were genotyped using signal intensity distributions and B-allele frequency of the genotyping probes with PennCNV software (39) and adjusted for genomic waves according to genomic GC content, as previously described (40). The CNV scan was completed for 2310 individuals in YFS and 4931 in NFBC1966 (41).

Expression quantitative trait loci

We queried significant SNPs from the Dietary, Lifestyle, and Genetic determinants of Obesity and Metabolic syndrome study, an extension of FINRISK 2007. The eQTL methods are previously published (17). Briefly, whole blood was extracted from 518 unrelated individuals and genotyped on the Illumina 610-Quad SNP array. In parallel, mRNA expression was quantified with Illumina HumanHT-12 Expression BeadChips. Linear regression was used to test association between transcript expression levels and each SNP.

Pathway analyses

We entered the nearest gene to all signals at $P < 1 \times 10^{-4}$ (one per locus) into the g:Profiler Gene Group Functional Profiler tool (g:GOST) (19), a webtool that, briefly, queries databases of biological pathways for enrichment of user-entered genes. For Analysis I, we also entered *MAPK3* because the gene was implicated by eQTL evidence to be functionally relevant ($n = 93$, corresponding to 0.0003% of all discovery signals for Analysis I). We also ran GSEA using MAGENTA (20), a program that calculates a P -value for each gene in the genome based on GWA results and then searches biological databases for pathways showing an enrichment of genes with lower than expected P -values. Analysis II and III data are not reported here due to the lack of significant findings.

Association analyses of age of menarche

To assess the relevance of the pubertal growth-associated variants for pubertal timing, leading signals not previously implicated in the timing of menarche were queried from *in silico* meta-analysis data of 87 802 women published by the ReproGen Consortium (23).

Cross-sectional height and BMI analyses

Height or BMI measurements from childhood to adulthood were divided into six age bins: (i) prepuberty (6.5–8.5 years old), (ii) early puberty (8.6–10.5 years old), (iii) mid-puberty for females (10.6–12.5 years old), (iv) mid-puberty for males (12.6–14.5 years old), (v) late puberty (14.6–17.5 years old) and finally (vi) adult (>17.6 years old). In each cohort, each marker of interest (imputed genotype dosage) was tested for

association with sex-specific height or BMI SDS for all age bins available, using linear regression assuming an additive model and adjusting for exact age at measurement (to the nearest month), along with optional correction for population stratification. A single measurement was included per study subject per bin, with the age closest to the mean used when more than one measurement was available. Altogether 23 SNPs were analysed for height and BMI across pubertal growth (only significantly associated markers are reported here). Summary statistics were meta-analysed like the primary analyses in each age bin, separately for males and females, for both height and BMI distinctly. Effect sizes were plotted versus age.

Early growth analyses

Cohorts with height measurements available at 1, 2, 3 or 4 years were included, namely the CHOP, Copenhagen Study on Asthma in Childhood, Generation R Study (Generation R), HBCS, Infancia y Medio Ambiente (Environment and Childhood) Project (INMA), LISApplus&GINIplus, NTR, Northern Finland Birth Cohort 1966 (NFBC1966), Prevention and Incidence of Asthma and Mite Allergy birth cohort study and Western Australian Pregnancy study (RAINE). Length was measured at 12 months (range 6–18 months) and height at 24 (range 18–30), 36 (range 30–42) and 48 (range 42–54) months. If multiple measurements per individual were available, those closest to 12, 24, 36 or 48 months were used. Sex- and age-adjusted SDSs were calculated using Growth Analyser 3.0 (Dutch Growth Research Foundation, Rotterdam, The Netherlands) in each study separately (42). The sex-specific association between each marker genotype and length or height SDS was assessed using linear regression, assuming an additive model. Imputed genotypes were used, where directly assayed genotypes were unavailable. We meta-analysed the within-cohort sex-stratified linear regression results using the inverse-variance method. A fixed-effects model was assumed, using RMeta in R (v.2.7.0).

URLS

SIMBioMS, <http://www.simbioms.org/>; R, <http://www.r-project.org/>; PLINK, <http://pngu.mgh.harvard.edu/~purcell/plink/>; IMPUTE, <http://mathgen.stats.ox.ac.uk/impute/impute.html>; MACH, <http://www.sph.umich.edu/csg/abecasis/MACH/index.html>; The International HapMap Project, <http://hapmap.ncbi.nlm.nih.gov/>; GenABEL, <http://www.genabel.org/>; SNP test, https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html; g:Profiler, <http://biit.cs.ut.ee/gprofiler/>; MAGENTA, <http://www.broadinstitute.org/mpg/magenta/>; Growth Analyser, <http://www.growthanalyser.org>.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

Detailed acknowledgements by study can be found in the supplementary note.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by the Academy of Finland (grant numbers 120315, 134839, 129297, 218029, 209072, 129255, 126925, 121584, 124282, 129378, 117787, 41071, 100499, 205585, 118555, 104781, 120315, 129418, 141054); Academy of Finland Center of Excellence in Complex Disease Genetics (grant numbers 213506, 129680); Australian National Health and Medical Research Council (241944, 339462, 389927, 389875, 389891, 389892, 389938, 442915, 442981, 496739, 552485, 552498); Australian Research Council (grant numbers A7960034, A79906588, A79801419, DP0770096, DP0212016, DP0343921); British Heart Foundation; Canadian Institutes of Health Research (grant number MOP-82893); Conselleria de Sanitat Generalitat Valenciana, and Fundació Roger Torné; Dutch Asthma Fonds; Emil Aaltonen Foundation; Erasmus Medical Center, Rotterdam, the Erasmus University Rotterdam and the Netherlands Organization for Health Research and Development (grant number 21000074); European Commission (ENGAGE project and grant agreement HEALTH-F4-2007-201413; EURO-BLCS, Framework 5 award QLGI-CT-2000-01643); European Research Council (grant number ERC-230374); Finnish Cultural Foundation; Finnish Foundation of Cardiovascular Research; Fundació La Marató de TV3, Generalitat de Catalunya (grant number CIRIT 1999SGR 00241); Genomics of Developmental Trajectories in Twins (grant number IRC2MH089 995-01); German Bundesministerium fuer Forschung und Technology (grant numbers 01 AK 803 A-H, 01 IG 07015 G); Infants Research Foundation; Institute Development Award from the Children's Hospital of Philadelphia; Instituto de Salud Carlos III (grant numbers CB06/02/0041, G03/176, FIS PI041436, PI081151, PI041705, PS09/00432, FIS-FEDER 03/1615, 04/1509, 04/1112, 04/1931, 05/1079, 05/1052, 06/1213, 07/0314, 09/02647); Juho Vainio Foundation; Ludwig-Maximilians-University's innovative research priority project MC-Health; Medical Research Council (grant numbers G0601653, G0000934, G0500539, G0600705, G0601653); National Institutes of Health (grant numbers R01 HD056465, 5R01MH63706:02); National Institute for Alcohol Abuse and Alcoholism (grant numbers AA-12502, AA-09203, AA-08315); NIH Heart, Lung and Blood Institute (grant number 5R01HL087679-02) through the STAMPEED program (grant number IRL1MH083268-01); National Health and Medical Research Council of Australia (grant numbers 403981, 003209); Netherlands Organization for Scientific research (grant number NWO/SPI 56-464-14192); Spanish Ministry of Science and Innovation (grant number SAF2008-00357); Stichting Astmabestrijding and Ministry of the Environment; Social Insurance Institution of Finland; Telethon Institute for Child Health Research and Women; Twin-family database for behavior genetics and genomics studies (grant number NWO 480-04-004); Type 1 Diabetes

Genetics Consortium (grant number U01 DK062418); Paavo Nurmi Foundation; Raine Medical Research Foundation; Research Development Award from the Cotswold Foundation; UK Medical Research Council (grant number 74882); University of Bristol; University of Western Australia; University Hospital Oulu; University of Oulu (grant number 75617); UWA Faculty of Medicine, Dentistry and Health Sciences; Wellcome Trust (grant numbers 076467, 068545/Z/02, 084762MA); and ZonMw: the Netherlands Organisation for Health Research and Development. Personal funding was provided by the Helsinki Biomedical Graduate Program to D.L.C.; Netherlands Organization for Health Research and Development (grant numbers ZonMw 90700303, 916.10159) to V.W.V.J.; Dutch Kidney Foundation (grant number C08.2251) to H.R.T.; MRC Centre for Causal Analyses in Translational Epidemiology (grant number RD1634) to G.D.S., D.M.E. and N.J.T.; European Community's Seventh Framework Programme (grant number FP7/2007-2013) and ENGAGE project, grant agreement HEALTH-F4-2007-201413 to I.P. and V.L.; Tampere and Turku University Hospital Medical Funds (grant numbers 9M048, 9N035) to T.L.; Tampere Tuberculosis Foundation to T.L.; National Health and Medical Research Council (grant number 613608) to E.M.B. and Fellowship Scheme to G.W.M.; and Wellcome Trust Sir Henry Wellcome Postdoctoral Research Fellow (grant number 092447/Z/10/Z) to J.R.B.P. J.P.K. is funded by a Wellcome Trust 4-year PhD studentship in molecular, genetic, and life course epidemiology (WT083431MA).

REFERENCES

- Lango Allen, H., Estrada, K., Lettre, G., Berndt, S.I., Weedon, M.N., Rivadeneira, F., Willer, C.J., Jackson, A.U., Vedantam, S., Raychaudhuri, S. *et al.* (2010) Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*, **467**, 832–838.
- Cooper, C., Kuh, D., Egger, P., Wadsworth, M. and Barker, D. (1996) Childhood growth and age at menarche. *Br. J. Obstet. Gynaecol.*, **103**, 814–817.
- He, Q. and Karlberg, J. (2001) BMI in childhood and its association with height gain, timing of puberty, and final height. *Pediatr. Res.*, **49**, 244–251.
- Aksglaede, L., Juul, A., Olsen, L.W. and Sorensen, T.I.A. (2009) Age at puberty and the emerging obesity epidemic. *PLoS One*, **4**, e8450.
- Silventoinen, K., Baker, J.L. and Sorensen, T.I.A. (2012) Growth in height in childhood and risk of coronary heart disease in adult men and women. *PLoS One*, **7**, e30476.
- Widen, E., Silventoinen, K., Sovio, U., Ripatti, S., Cousminer, D.L., Hartikainen, A.L., Laitinen, J., Pouta, A., Kaprio, J., Jarvelin, M.R. *et al.* (2012) Pubertal timing and growth influences cardiometabolic risk factors in adult males and females. *Diabetes Care*, **35**, 850–856.
- Lakshman, R., Forouhi, N., Luben, R., Bingham, S., Khaw, K., Wareham, N. and Ong, K.K. (2008) Association between age at menarche and risk of diabetes in adults: Results from the EPIC-norfolk cohort study. *Diabetologia*, **51**, 781–786.
- Tanner, J.M., Whitehouse, R.H., Marubini, E. and Resele, L.F. (1976) The adolescent growth spurt of boys and girls of the Harpenden growth study. *Ann. Hum. Biol.*, **3**, 109–126.
- Perry, R.J., Farquharson, C. and Ahmed, S.F. (2008) The role of sex steroids in controlling pubertal growth. *Clin. Endocrinol.*, **68**, 4–15.
- Luo, Z.C. and Karlberg, J. (2000) Critical growth phases for adult shortness. *Am. J. Epidemiol.*, **152**, 125–131.
- Sharma, J.C. (1983) The genetic contribution to pubertal growth and development studied by longitudinal growth data on twins. *Ann. Hum. Biol.*, **10**, 163–171.
- Morris, D.H., Jones, M.E., Schoemaker, M.J., Ashworth, A. and Swerdlow, A.J. (2011) Familial concordance for age at menarche:

- analyses from the Breakthrough Generations Study. *Paediatr. Perinat. Epidemiol.*, **25**, 306–311.
13. Silventoinen, K., Pietiläinen, K.H., Tynelius, P., Sorensen, T.I., Kaprio, J. and Rasmussen, F. (2008) Genetic regulation of growth from birth to 18 years of age: The Swedish Young Male Twins Study. *Am. J. Hum. Biol.*, **20**, 292–298.
 14. Wehkalampi, K., Silventoinen, K., Kaprio, J., Dick, D.M., Rose, R.J., Pulkkinen, L. and Dunkel, L. (2008) Genetic and environmental influences on pubertal timing assessed by height growth. *Am. J. Hum. Biol.*, **20**, 417–423.
 15. Widén, E., Ripatti, S., Cousminer, D.L., Surakka, I., Lappalainen, T., Järvelin, M.R., Eriksson, J.G., Raitakari, O., Salomaa, V., Sovio, U. *et al.* (2010) Distinct variants at LIN28B influence growth in height from birth to adulthood. *Am. J. Hum. Genet.*, **86**, 773–782.
 16. Vrieze, S.I., McGue, M., Miller, M.B., LeGrand, L.N., Schork, N.J. and Iacono, W.G. (2011) An assessment of the individual and collective effects of variants on height using twins and a developmentally informative study design. *PLoS Genet.*, **7**, e1002413.
 17. Inouye, M., Kettunen, J., Soininen, P., Silander, K., Ripatti, S., Kumpula, L.S., Hämäläinen, E., Jousilahti, P., Kangas, A.J., Männistö, S. *et al.* (2010) Metabonomic, transcriptomic, and genomic variation of a population cohort. *Mol. Syst. Biol.*, **6**, 441.
 18. Sebastian, A., Matsushita, T., Kawanami, A., Mackem, S., Landreth, G.E. and Murakami, S. (2011) Genetic inactivation of ERK1 and ERK2 in chondrocytes promotes bone growth and enlarges the spinal canal. *J. Orthop. Res.*, **29**, 375–379.
 19. Reimand, J., Kull, M., Peterson, H., Hansen, J. and Vilo, J. (2007) g:Profiler—a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Res.*, **35**, W193–W200.
 20. Segre, A.V., DIAGRAM Consortium, MAGIC investigators, Groop, L., Mootha, V.K., Daly, M.J. and Altshuler, D. (2010) Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLoS Genet.*, **6**, e1001058.
 21. Walters, R.G., Jacquemont, S., Valsesia, A., de Smith, A.J., Martinet, D., Andersson, J., Falchi, M., Chen, F., Andrieux, J., Lobbens, S. *et al.* (2010) A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. *Nature*, **463**, 671–675.
 22. Jacquemont, S., Reymond, A., Zufferey, F., Harewood, L., Walters, R.G., Kutalik, Z., Martinet, D., Shen, Y., Valsesia, A., Beckmann, N.D. *et al.* (2011) Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. *Nature*, **478**, 97–102.
 23. Elks, C.E., Perry, J.R., Sulem, P., Chasman, D.I., Franceschini, N., He, C., Lunetta, K.L., Visser, J.A., Byrne, E.M., Cousminer, D.L. *et al.* (2010) Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. *Nat. Genet.*, **42**, 1077–1085.
 24. Bradfield, J.P., Taal, H.R., Timpson, N.J., Scherag, A., Leccoer, C., Warrington, N.M., Hyppönen, E., Holst, C., Valcarcel, B., Thiering, E. *et al.* (2012) A genome-wide association meta-analysis identifies new childhood obesity loci. *Nat. Genet.*, **44**, 526–531.
 25. Speliotes, E.K., Willer, C.J., Berndt, S.I., Monda, K.L., Thorleifsson, G., Jackson, A.U., Lango Allen, H., Lindgren, C.M., Luan, J., Mägi, R. *et al.* (2010) Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.*, **42**, 937–948.
 26. Wood, A.R., Hernandez, D.G., Nalls, M.A., Yaghoobkar, H., Gibbs, J.R., Harries, L.W., Chong, S., Moore, M., Weedon, M.N., Guralnik, J.M. *et al.* (2011) Allelic heterogeneity and more detailed analyses of known loci explain additional phenotypic variation and reveal complex patterns of association. *Hum. Mol. Genet.*, **20**, 4082–4092.
 27. Okasha, M., McCarron, P., McEwen, J. and Smith, G.D. (2001) Age at menarche: secular trends and association with adult anthropometric measures. *Ann. Hum. Biol.*, **28**, 68–78.
 28. Rousseau, F., Bonaventure, J., Legeai-Mallet, L., Pelet, A., Rozet, J.M., Maroteaux, P., Le Merrer, M. and Munnich, A. (1994) Mutations in the gene encoding fibroblast growth factor receptor-3 in achondroplasia. *Nature*, **371**, 252–254.
 29. Shiang, R., Thompson, L.M., Zhu, Y.Z., Church, D.M., Fielder, T.J., Bocian, M., Winokur, S.T. and Wasmuth, J.J. (1994) Mutations in the transmembrane domain of FGFR3 cause the most common genetic form of dwarfism, achondroplasia. *Cell*, **78**, 335–342.
 30. Bellus, G.A., McIntosh, I., Smith, E.A., Aylsworth, A.S., Kaitila, I., Horton, W.A., Greenhaw, G.A., Hecht, J.T. and Francomano, C.A. (1995) A recurrent mutation in the tyrosine kinase domain of fibroblast growth factor receptor 3 causes hypochondroplasia. *Nat. Genet.*, **10**, 357–359.
 31. Aoki, Y., Niihori, T., Kawame, H., Kurosawa, K., Ohashi, H., Tanaka, Y., Filocamo, M., Kato, K., Suzuki, Y., Kure, S. *et al.* (2005) Germline mutations in HRAS proto-oncogene cause Costello syndrome. *Nat. Genet.*, **37**, 1038–1040.
 32. Liu, F., Austin, D.A., Mellon, P.L., Olefsky, J.M. and Webster, N.J. (2002) GnRH activates ERK1/2 leading to the induction of c-fos and LHbeta protein expression in LbetaT2 cells. *Mol. Endocrinol.*, **16**, 419–434.
 33. Krestyaninova, M., Zarins, A., Viksna, J., Kurbatova, N., Rucevskis, P., Neogi, S.G., Gostev, M., Perheentupa, T., Knuutila, J., Barrett, A. *et al.* (2009) A system for information management in BioMedical studies—SIMBioMS. *Bioinformatics*, **25**, 2768–2769.
 34. Marchini, J., Howie, B., Myers, S., McVean, G. and Donnelly, P. (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.*, **39**, 906–913.
 35. Li, Y., Willer, C.J., Ding, J., Scheet, P. and Abecasis, G.R. (2010) MaCH: Using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet. Epidemiol.*, **34**, 816–834.
 36. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.
 37. Aulchenko, Y.S., Ripke, S., Isaacs, A. and van Duijn, C.M. (2007) GenABEL: an R library for genome-wide association analysis. *Bioinformatics*, **23**, 1294–1296.
 38. Aulchenko, Y.S., Struchalin, M.V. and van Duijn, C.M. (2010) ProbABEL package for genome-wide association analysis of imputed data. *BMC Bioinformatics*, **11**, 134.
 39. Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S.F., Hakonarson, H. and Bucan, M. (2007) PennCNV: an integrated hidden markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.*, **17**, 1665–1674.
 40. Diskin, S.J., Li, M., Hou, C., Yang, S., Glessner, J., Hakonarson, H., Bucan, M., Maris, J.M. and Wang, K. (2008) Adjustment of genomic waves in signal intensities from whole-genome SNP genotyping platforms. *Nucleic Acids Res.*, **36**, e126.
 41. Pietiläinen, O.P., Rehnstrom, K., Jakkula, E., Service, S.K., Congdon, E., Tilgmann, C., Hartikainen, A.L., Taanila, A., Heikura, U., Paunio, T. *et al.* (2011) Phenotype mining in CNV carriers from a population cohort. *Hum. Mol. Genet.*, **20**, 2686–2695.
 42. Fredriks, A.M., van Buuren, S., Burgmeijer, R.J., Meulmeester, J.F., Beuker, R.J., Brugman, E., Roede, M.J., Verloove-Vanhorick, S.P. and Wit, J.M. (2000) Continuing positive secular growth change in the Netherlands 1955–1997. *Pediatr. Res.*, **47**, 316–323.
 43. Freathy, R.M., Mook-Kanamori, D.O., Sovio, U., Prokopenko, I., Timpson, N.J., Berry, D.J., Warrington, N.M., Widen, E., Hottenga, J.J., Kaakinen, M. *et al.* (2010) Variants in ADCY5 and near CCNL1 are associated with fetal growth and birth weight. *Nat. Genet.*, **42**, 430–435.