Distinct Variants at *LIN28B* Influence Growth in Height from Birth to Adulthood

Elisabeth Widén,[1,*] Samuli Ripatti,[1,2] Diana L. Cousminer,[1] Ida Surakka,[1,2] Tuuli Lappalainen,[1,3,4] Marjo-Riitta Järvelin,[5] Johan G. Eriksson,[2,6,7,8] Olli Raitakari,[9,10] Veikko Salomaa,[2] Ulla Sovio,[5] Anna-Liisa Hartikainen,[11] Anneli Pouta,[2] Mark I. McCarthy,[12,13] Clive Osmond,[14] Eero Kajantie,[2] Terho Lehtimäki,[15] Jorma Viikari,[16] Mika Kähönen,[17] Chris Tyler-Smith,[18] Nelson Freimer,[19,20,21] Joel N. Hirschhorn,[22,23,24] Leena Peltonen,[1,2,18,24] and Aarno Palotie[1,18,24]

We have studied the largely unknown genetic underpinnings of height growth by using a unique resource of longitudinal childhood height data available in Finnish population cohorts. After applying GWAS mapping of potential genes influencing pubertal height growth followed by further characterization of the genetic effects on complete postnatal growth trajectories, we have identified strong association between variants near *LIN28B* and pubertal growth (rs7759938; female $p = 4.0 \times 10^{-9}$, male $p = 1.5 \times 10^{-4}$, combined $p = 5.0 \times 10^{-11}$, n = 5038). Analysis of growth during early puberty confirmed an effect on the timing of the growth spurt. Correlated SNPs have previously been implicated as influencing both adult stature and age at menarche, the same alleles associating with taller height and later age of menarche in other studies as with later pubertal growth here. Additionally, a partially correlated *LIN28B* SNP, rs314277, has been associated previously with final height. Testing both rs7759938 and rs314277 (pairwise $r^2 = 0.29$) for independent effects on postnatal growth in 8903 subjects indicated that the pubertal timing-associated marker rs7759938 affects prepubertal growth in females ($p = 7 \times 10^{-5}$) and final height in males ($p = 5 \times 10^{-4}$), whereas rs314277 has sex-specific effects on growth (p for interaction = 0.005) that were distinct from those observed at rs7759938. In conclusion, partially correlated variants at *LIN28B* tag distinctive, complex, and sex-specific height-growth-regulating effects, influencing the entire period of postnatal growth. These findings imply a critical role for *LIN28B* in the regulation of human growth.

Human growth in height is a multifaceted process including periods of accelerated and decelerated growth velocities. The postnatal growth trajectory can be conceptualized as consisting of three partially overlapping phases of growth: infant growth characterized by rapidly declining growth velocities, slowly decelerating childhood growth, and the pubertal height growth spurt.[1] Genes are estimated to account for up to 60%–80% of the within-population variation of overall height growth.[2,3] Although the individual genes still remain largely unknown, epidemiological studies propose partly overlapping genetic regulation covering multiple aspects of growth. For example, a longitudinal study of Swedish male twins suggested that a large proportion of the genes affecting postnatal growth in height are the same or are closely linked throughout the whole growth period.[3] Furthermore, a significant proportion of shared genes is thought to account for the correlation between increased prepubertal body mass index and the timing of pubertal growth and maturation.[4] Thus far, there have been no genome-wide association studies (GWAS) specifically targeting childhood height growth, whereas there have been many recent large-scale association studies successfully identifying loci influencing both body size and pubertal timing. Although there are as many as 47 verified hits influencing final stature,[2,5] 11 loci influencing adult body mass index,[6–8] and 2 loci influencing age at menarche,[9–12] these findings explain only a marginal proportion of the overall variance of each trait. Additionally, we know very little about how these loci may influence longitudinal growth.

To elucidate the genetic framework influencing height growth, we utilized the unique resource of longitudinal childhood height data available in Finnish cohorts.[13–19] We chose pubertal growth as the primary mapping target

[1]Institute for Molecular Medicine Finland, University of Helsinki, 00014 Helsinki, Finland; [2]National Institute for Health and Welfare, 00271 Helsinki, Finland; [3]Department of Medical Genetics, University of Helsinki, 00014 Helsinki, Finland; [4]Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva 1211, Switzerland; [5]Department of Epidemiology and Public Health, Imperial College London, London W2 1PG, UK; [6]Department of General Practice and Primary Health Care, University of Helsinki, 00014 Helsinki, Finland; [7]Unit of General Practice, Helsinki University Central Hospital, 00029 Helsinki, Finland; [8]Folkhalsan Research Centre, Biomedicum Helsinki, University of Helsinki, 00014 Helsinki, Finland; [9]Department of Clinical Physiology, Turku University Hospital, 20520 Turku, Finland; [10]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, 20520 Turku, Finland; [11]Department of Obstetrics and Gynecology, University of Oulu, 90220 Oulu, Finland; [12]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK; [13]Oxford Centre for Diabetes, Endocrinology and Medicine, University of Oxford, Oxford OX3 7LJ, UK; [14]MRC Epidemiology Resource Centre, University of Southampton, Southampton SO16 6YD, UK; [15]Department of Clinical Chemistry, University Hospital and University of Tampere, 33521 Tampere, Finland; [16]Department of Medicine, University of Turku and Turku University Hospital, 20520 Turku, Finland; [17]Department of Clinical Physiology, University of Tampere and Tampere University Hospital, 33521 Tampere, Finland; [18]Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK; [19]Center for Neurobehavioral Genetics, University of California, Los Angeles, Los Angeles, CA 90095, USA; [20]The Jane and Terry Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, Los Angeles, CA 90095, USA; [21]Department of Psychiatry, University of California, Los Angeles, Los Angeles, CA 90095, USA; [22]Program in Genomics and Division of Endocrinology, Children's Hospital, Boston, MA 02115, USA; [23]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA; [24]Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA
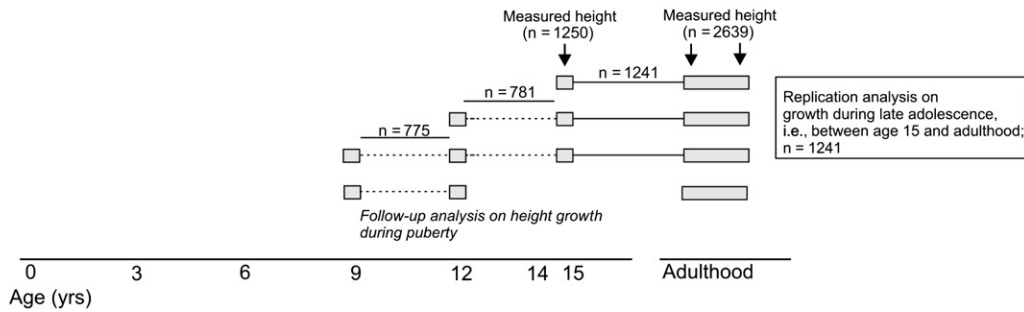*Correspondence: elisabeth.widen@helsinki.fi
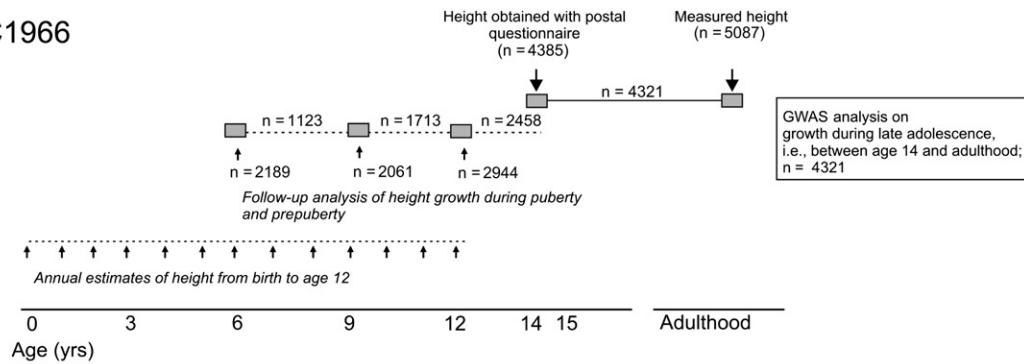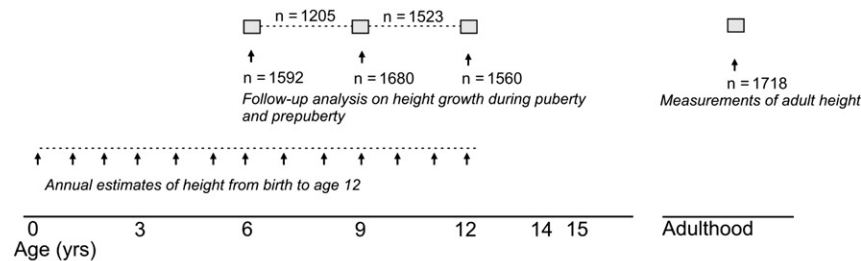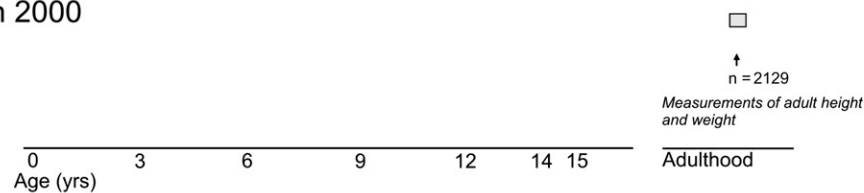
Figure 1. Study Populations and Design

Study samples:

Northern Finland Birth Cohort 1966 (NFBC1966) is a prospective cohort study conducted in the two northernmost provinces of Finland, Oulu and Lapland, representing 96% of all births in these provinces.[13,14] The offspring were followed up at 6 mos. and at 1, 14, and 31 yrs of age. Height at 14 yrs was obtained by self-report from questionnaires mailed to the adolescents. At age 31, a representative subsample of the study subjects (cohort members still living in Northern Finland or in the capital area) were invited for clinical examination. At the clinical visit, the participants' (n = 5654) height and weight were assessed, and a blood sample was obtained, from which DNA was extracted. Extended data on height and weight from 1 to 12 yrs were retrospectively obtained from growth records originally collected by the school health care service and thereafter archived at communal health clinics. In case there were two height measurements within ±1 year of a specific birthday, the height at the birthday was extrapolated from the measurements, assuming a linear relationship. For the remaining subjects, the corresponding height was evaluated by specific linear regression only if there was one measurement within ±0.5 yrs of the birthday.

Cardiovascular Risk of Young Finns Study (YF) is a prospective cohort study conducted at five university departments of medical schools in Finland (i.e., Turku, Helsinki, Kuopio, Tampere, and Oulu), with the aim of studying the levels of cardiovascular risk factors in children and adolescents in different parts of the country. The baseline study was conducted in 1980, and the study subjects were followed with

for many reasons. Importantly, this growth phase regulates final height, accounting for as much as 15%–20% of adult stature.[20,21] Furthermore, most aspects of this growth period are highly heritable.[3,22,23] Finally, the phenotypic variation is very large, with as much within-sex variation in timing as 4 yrs.[24] The design of our study is presented in Figure 1. Altogether, three different Finnish population cohorts with longitudinal height data collected at different time points during childhood and adulthood, in addition to a fourth cohort with data on adult height, were included. A precise assessment of the pubertal growth spurt requires very frequent height measurements spanning a large age range, data which typically are not readily obtained in an epidemiological setting. Therefore, we monitored this growth phase with a simple and robust measurement capturing growth during late adolescence, the increase in height between age 14 and adulthood. This approach allowed us to maximize the number of subjects available for genome-wide association mapping, thereby also increasing the statistical power. A similar measurement, i.e., the change in relative height between age 12 and adulthood in females and between age 14 and adulthood in males, has previously been shown to correlate strongly with the timing of the pubertal growth spurt.[25] Thus, our study design primarily facilitated the detection of loci influencing the timing of the pubertal growth spurt. Moreover, the availability of longitudinal data on childhood height enabled the exploration of putative shared genetic effects influencing multiple periods of postnatal height growth.

We first performed GWAS analysis of height growth during late adolescence in Northern Finland Birth Cohort 1966 (NFBC1966). The increase in height between age 14 and adulthood was used as an estimate of pubertal growth in 2073 males and 2248 females with genotypes from Illumina Infinium 370CNV Duo arrays (Figure 1).[26] Applying previously established thresholds,[26] five markers all located in a single block of linkage disequilibrium (LD) spanning 111 kb on chromosome 6q21, a block containing the 5′ end and upstream region of LIN28B (MIM 611044) (lin-28 homolog B [C. elegans]) and an uncharacterized open reading frame c6orf220, yielded evidence for a significant association signal in the combined analysis of males and females (p < 5 × 10^{-7}). The signal predominantly came from females and to a lesser extent also from males (rs7759938; beta = 0.159, standard error [SE] = 0.032, p = 7.2 × 10^{-7} in females versus beta = 0.093, SE = 0.033, p = 0.005 in males; Table 1). Testing the best-associated SNP (rs7759938) and a proxy (rs314268; beta = 0.140, SE = 0.032, p = 1.4 × 10^{-5} in females versus beta = 0.095, SE = 0.033, p = 0.004 in males) in the longitudinal replication cohort Cardiovascular Risk in Young Finns Study (YF), including 1241 individuals (Figure 1), added further support to the original observation (rs7759938, combined p in all subjects = 5.0 × 10^{-11}). The strongest associated marker explained approximately 1% of the total phenotypic variation in females and roughly 0.5% in males, estimated with the regression r^2 in the Northern Finland Birth Cohort Study with the normalized phenotype as the outcome in an additive regression model adjusting for the first two dimensions obtained from the multidimensional scaling analysis.[27] Also, the genome-wide association p values showed more apparent departure from the quantile-quantile plot in females (see Figure S1 available online). The reason for this phenomenon remains unclear. Prior to regression analysis, the phenotypic outcome was normalized and standardized, but the underlying raw phenotype differed between males and females. The average peak growth velocity occurs at approximately 12 yrs in females and 14 yrs in males. As a consequence, the GWAS outcome captures growth a bit later relative to the average peak growth velocity in girls than in boys. Thus, the difference in association signals could in part be a reflection that our measurement of growth during late adolescence targets pubertal growth with a slightly different focus and resolution in both sexes.

We hypothesized that the observed association might be a consequence of an effect on the timing of the pubertal

3 yr intervals until 1992; more recently in 2001, when blood samples for DNA extraction were drawn (n = 2620); and in 2007.[17,18] Height and weight were measured at each clinical visit. Adult stature was measured between ages 24 and 42.

Helsinki Birth Cohort Study (HBCS) includes 8760 subjects born in Helsinki between 1934 and 1944.[15,16] Between 2000 and 2002, a representative subset of 928 males and 1075 females participated in a clinical study focusing upon cardiovascular and metabolic outcomes and cognitive function. Height during childhood until age 12 was obtained from child welfare and school health care records, and height at specific birth days was extracted as described by Eriksson et al.[16] Adult stature was measured when the study subjects were between ages 59 and 70.

Adult height was available in the Health 2000 study (H2000), which was a health interview/examination survey carried out by the National Institute for Health and Welfare in Finland from fall 2000 to spring 2001, with a nationally representative sample of 10,000 individuals drawn from the Finnish population aged 18 and older. The main topics of the study were health status, major chronic conditions, functional ability and limitations, determinants of health, and use of health care.[19] From a subcohort of 6000 individuals representative of the Finnish population over age 30, roughly 1000 nondiabetic subjects meeting the International Diabetes Federation criteria for metabolic syndrome and a control cohort of 1000 subjects matched for sex, age, and residence were selected for GWAS analyses. The age range was between ages 30 and 75. In the current study, the data for two genotyped markers were used.

All cohort studies were approved by their corresponding local ethical committees, and the study subjects gave their informed consent according to the approved protocols.

Study design: The primary GWAS analysis on height growth during late adolescence (solid line) was analyzed based on height growth between age 14 and adulthood, and the replication analyses were carried out in YF via a similar estimate of pubertal height growth, i.e., the increase in height between age 15 and adulthood. Growth during early and midadolescence (broken line) was assessed as the increase in height between ages 9 and 12 in YF and HBCS or as the increase in height between ages 12 and 14 in NFBC1966 and between ages 12 and 15 in YF. Further follow-up studies of the entire growth trajectory from birth to age 12 were carried out in NFBC1966 and HBCS.

Table 1. Association Signals between Pubertal Growth during Late Adolescence and *LIN28B* Obtained in Northern Finland Birth Cohort 1966 and the Replication Cohort Cardiovascular Risk of Young Finns Study

NFBC1966 Females								YF Females			
Chr.	SNP	Position	n	Beta (SE)	p Value	A1	MAF	n	Beta (SE)	p Value	MAF
6	rs10485311	105445489	2245	0.107 (0.058)	0.06334	G	0.07	-	-	-	-
6	rs6923490	105455398	2247	0.083 (0.036)	0.02258	A	0.207	-	-	-	-
6	rs4946651	105476203	2235	0.146 (0.03)	1.481×10^{-6}	A	0.435	-	-	-	-
6	rs7759938	105485647	2242	0.159 (0.032)	7.204×10^{-7}	G	0.318	664	0.167 (0.058)	0.004311	0.33
6	rs314262	105501314	2200	0.15 (0.031)	9.103×10^{-7}	G	0.44	-	-	-	-
6	rs314280	105507530	2241	0.151 (0.03)	5.401×10^{-7}	A	0.436	-	-	-	-
6	rs12194974	105510891	2248	−0.052 (0.051)	0.31	A	0.09	-	-	-	-
6	rs314277	105514355	2214	0.126 (0.04)	0.001654	A	0.167	-	-	-	-
6	rs314268	105524671	2240	0.14 (0.032)	1.422×10^{-5}	G	0.32	666	0.135 (0.058)	0.01967	0.336

NFBC1966 Males								YF Males			
Chr.	SNP	Position	n	Beta (SE)	p Value	A1	MAF	n	Beta (SE)	p Value	MAF
6	rs10485311	105445489	2073	0.124 (0.058)	0.03156	G	0.076	-	-	-	-
6	rs6923490	105455398	2072	0.04 (0.038)	0.3027	A	0.201	-	-	-	-
6	rs4946651	105476203	2060	0.076 (0.031)	0.01611	A	0.436	-	-	-	-
6	rs7759938	105485647	2068	0.093 (0.033)	0.004647	G	0.324	569	0.144 (0.066)	0.02969	0.308
6	rs314262	105501314	2019	0.082 (0.032)	0.009394	G	0.44	-	-	-	-
6	rs314280	105507530	2066	0.072 (0.031)	0.02155	A	0.436	-	-	-	-
6	rs12194974	105510891	2072	0.001 (0.058)	0.9874	A	0.084	-	-	-	-
6	rs314277	105514355	2047	0.062 (0.041)	0.1263	A	0.174	-	-	-	-
6	rs314268	105524671	2065	0.095 (0.033)	0.003559	G	0.329	570	0.121(0.066)	0.06681	0.315

Pubertal growth was estimated as height increase after age 14 (in NFBC1966) or age 15 (in YF). The data were analyzed by linear regression implemented in PLINK[27] covering the autosomes, assuming additive inheritance and including individual-specific scores of the first two dimensions of the multidimensional scaling identical-by-state (IBS) analysis as covariates. The phenotype distribution for the increase in height between ages 14 and 31 was normalized by logarithm transformation, and sex-specific Z scores computed from the normalized phenotype were used as input in the association analysis. The genomic inflation factor λ was 1.04 in both sexes. The association results from males and females were combined into fixed-effect meta-analysis with reciprocal weighting on the square of standard errors of the effect size estimates via the MetABEL package for the R software. Positions are based on the NCBI B36 assembly. Abbreviations are as follows: NFBC1966, Northern Finland Birth Cohort 1966; YF, Cardiovascular Risk in Young Finns Study; SE, standard error; MAF, minor allele frequency.

growth spurt. Therefore, we conducted follow-up analyses to characterize the growth effect at marker locus rs7759938 in more detail. Heights at ages 9, 12, and 15 were available in YF, prompting analysis of the height increase between ages 9 and 12 and between ages 12 and 14/15 to estimate pubertal growth during early and midpuberty separately in the subset of NFBC1966 study subjects with archived growth charts available. Finally, we estimated growth during early adolescence based on archived growth charts in a third Finnish cohort, the Helsinki Birth Cohort Study (HBCS) (Figure 1). The results are presented in Table S1. Considering growth between ages 9 and 12 as a proxy for the onset of the pubertal growth spurt, the observed correlation between the G (ancestral) allele at rs7759938 and decreased growth during early puberty both in males and females ($p = 1.7 \times 10^{-4}$) further supports an association between the G allele and later timing of the pubertal growth spurt.

Interestingly, the *LIN28B* region has previously been associated with the timing of menarche in females.[9–12]

Four independent studies reported significant association either with rs7759938, yielding the strongest evidence for association in our study, or with markers in tight correlation, the same alleles associating with later age of menarche in other studies as with later timing of pubertal growth here (Table S2; Figure S3). Marker rs7759938 was also associated with age of menarche in the Finnish cohorts (n = 4379, beta = 0.124, SE = 0.023, p = 8.3 × 10^{-8}). Even though both the pubertal growth spurt and menarche are secondary manifestations of pubertal maturation, the growth spurt represents an early marker of pubertal development in girls, whereas menarche is a late event. To clarify the possible causal direction of the associations, we performed multiple regression analysis including both rs7759938 and height growth during early puberty (between ages 9 and 12) as a proxy for timing of the pubertal growth spurt in the regression model. This conditioned analysis resulted in a major reduction of the association signal with age of menarche (beta = 0.046, SE = 0.030, p = 0.13, n = 1932), suggesting that the effects

Table 2. Linear Regression Analysis of Final Height in Three Finnish Population Cohorts Evaluating the Effects of rs7759938 and rs314277

| | Regression of Single Markers | | Simultaneous Regression of Both Markers | |
	rs7759938	rs314277	rs7759938	rs314277
NFBC1966 males	0.087 (0.030)	0.036 (0.037)	0.121 (0.041)	−0.063 (0.050)
HBCS males	0.061 (0.057)	0.034 (0.077)	0.072 (0.072)	−0.025 (0.097)
H2000 males	0.066 (0.048)	0.000 (0.062)	0.102 (0.060)	−0.077 (0.077)
All males (beta [SE], p)	0.078 (0.023), p = 8 × 10^{-4}	0.028 (0.029), p = 0.35	0.107 (0.031), p = 5 × 10^{-4}	−0.061 (0.039), p = 0.12
n	4211	4184	4176	4176
NFBC1966 females	0.062 (0.030)	0.089 (0.037)	0.028 (0.039)	0.0721 (0.048)
HBCS females	0.090 (0.051)	0.159 (0.065)	0.021 (0.067)	0.137 (0.086)
H2000 females	0.021 (0.047)	0.039 (0.059)	0.004 (0.059)	0.036 (0.076)
All females (beta [SE], p)	0.058 (0.023), p = 0.01	0.091 (0.028), p = 0.001	0.021 (0.029), p = 0.48	0.075 (0.037), p = 0.04
n	4692	4676	4645	4645
Females and males (beta [SE], p)	0.074 (0.017), p = 3 × 10^{-5}	0.062 (0.020), p = 0.003	0.062 (0.021), p = 0.003	0.011 (0.027), p = 0.69

The study populations included in the analysis were Northern Finland Birth Cohort 1966 (NFBC1966), Helsinki Birth Cohort Study (HBCS), and Health 2000 (H2000). The regression analysis of single markers at *LIN28B* is shown on the left, and the multiple regression analysis including both markers at *LIN28B* simultaneously in the regression model is shown on the right. The effect alleles are G at rs7759938 and A at rs314277. The regression analyses were run with cohort- and sex-specific Z scores as the independent variable. In H2000, adult height was adjusted for age, and the individual scores were obtained from the first two dimensions of multidimensional scaling IBS analysis of genotypes generated with Illumina 610 Bead Chips. SE, standard error.

on both the timing of the growth spurt and the timing of menarche might not be independent. Rather, the timing of pubertal growth and age of menarche may be mediated through a common underlying mechanism.

In addition to association with pubertal timing, the *LIN28B* region also coincides with a previously verified height locus.[28,29] Contrary to the published association findings with age of menarche, association to final height has been reported at two distinct marker loci that are only partially correlated with each other (pairwise r^2 = 0.26), at rs314277 with a p value of 1.1 × 10^{-8} in roughly 25,000 individuals[28] and at rs314268 with a p value of 7.7 × 10^{-7} in 49,000 individuals.[29] Marker rs314268 appears to overlap with the pubertal timing effect, showing strong correlation with the pubertal timing-associated SNP rs7759938 (r^2 = 0.94; Figure S3). Thus, the effect on final height could be mediated through pubertal timing because individuals maturing later, as a result of a delayed growth spurt and an extended overall period of height growth, may grow taller than their earlier-maturing peers.[30,31] Also, the previously published association finding at rs314277 could reflect the same underlying functional mutation, in spite of the modest pairwise correlation between rs314277 and the pubertal timing-associated marker rs7759938 (r^2 = 0.29). Alternatively, rs314277 might tag a second independent effect influencing final height.

To test for the presence of two separate height-regulating effects at *LIN28B*, we analyzed final stature in three Finnish cohorts, both by considering the effect of the pubertal timing-associated markers rs7759938 and rs314277 separately and by including both markers simultaneously in the regression model. Consistent with the previous studies on adult height, we found the markers to be associated

with adult stature when analyzed individually (n = 8903 at rs7759938, n = 8860 at rs314277; Table 2). Separate evaluation of females and males showed that rs314277 appeared to contribute to final height more significantly in females, whereas the effect at rs7759938 was, if anything, stronger in males. An analysis of rs314277 conditional on rs7759938 showed opposite effects on height in males and females, also suggestive of a sex-specific effect of this SNP (Table 2). To formally test for interaction between genotypes and sex, we performed multiple regression analysis, including both of the SNPs, sex, and all of their possible interactions in the model (Table 3), finding evidence for sex-genotype interaction at rs314277 (p = 0.005). The data thus support two independent effects at *LIN28B* influencing final height, one tagged by the pubertal timing-associated markers and one tagged by rs314277.

Because the different phases of postnatal growth likely share a significant proportion of the regulating genes,[3] we wanted to evaluate whether the specific height growth effects at *LIN28B* may also influence height growth prior to puberty by analyzing the effect of both *LIN28B* markers simultaneously on the complete postnatal growth trajectory at 1 yr intervals in two independent cohorts (NFBC1966 and HBCS). We first visualized the effects by plotting the marker-specific betas and standard errors obtained from the multiple regression analyses of height throughout childhood, including both markers in the regression model, as a line chart (Figure 2). The analyses showed markedly different patterns of association with height in both sexes.

Consistent with our previous analysis of growth during early and late adolescence, the G allele at the pubertal timing-associated marker rs7759938 was associated with

Table 3. Multiple Regression Analysis of Adult Stature in Three Finnish Cohorts Evaluating Marker-Marker and Marker-Sex Interactions

Independent Variable	Beta	SE	p Value
rs7759938	0.10523	0.02986	0.000427
rs314277	−0.16806	0.07567	0.026377
SEX	−0.65858	0.02893	$<2 \times 10^{-16}$
rs7759938 × rs314277 interaction	0.04167	0.04989	0.403553
rs7759938 × SEX interaction	−0.07596	0.04143	0.066754
rs314277 × SEX interaction	0.29371	0.10412	0.004802
rs7759938 × rs314277 × SEX interaction	−0.07654	0.06932	0.269571

The regression analysis was run with cohort-specific Z scores as the dependent variable in Northern Finland Birth Cohort 1966, Helsinki Birth Cohort Study, and Health 2000; n = 8821. Markers rs7759938 and rs314277 and sex, in addition to marker and sex interaction, were included in the regression model. In H2000, adult height was adjusted for age, and the individual scores were obtained from the first dimensions of multidimensional scaling IBS analysis of genotypes generated with Illumina 610 Bead Chips prior to the multiple regression analysis. The AA, AG, and GG genotypes at marker locus rs7759938 were coded as 0, 1, and 2; the CC, CA, and AA genotypes at rs314277 were coded as 0, 1, and 2; and male and female sex were coded as 0 and 1.

Figure 2. Linear Regression Analysis of Postnatal Height Evaluating the Independent Effects of rs7759938 and rs314277
The results of linear regression of standardized birth length, standardized height at 1 yr intervals between age 1 and 12, standardized height at age 14, and adult stature in Northern Finland Birth Cohort 1966 and Helsinki Birth Cohort Study at marker locus rs7759938 are shown in the upper panel and at rs314277 are shown in the lower panel. The sex-specific effect sizes and standard errors were calculated, including both rs7759938 and rs314277 in the regression model. For clarity, only one-sided standard errors are shown. The nominal p values of the sex specific effects are indicated in the figure as follows: #p < 0.1–0.05, *p < 0.05–0.01, **p < 0.01. The effect alleles are G and A at rs7759938 and rs314277, respectively. n at ages 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, and 14 and at adulthood equals 2315, 2199, 1925, 1590, 1356, 1335, 1702, 2254, 1855, 1744, 1767, 2153, 2127, 2042, and 3136 in males and 2638, 2537, 2176, 1852, 1630, 1577, 2010, 2 514, 2054, 1935, 2001, 2417, 2292, 2208, and 3555 in females. Height data was available for both cohorts up until age 12 and as an adult. Height at age 14 was available only in NFBC1966. Birth length in both cohorts has been adjusted for gestational age.

a reduction of the effect on relative height after age 9, followed by a significant increase in the effect after ages 12 and 14 (Figure 2), whereas no such effect was seen at rs314277. Analyzing growth after age 14, including both rs7759938 and rs314277 in the regression model, confirmed that only rs7759938 influenced the timing of the pubertal growth spurt in both sexes (females and males combined rs7759938 beta = 0.130, SE = 0.030, p = 1.8×10^{-5} versus rs314277 beta = −0.007, SE = 0.037, p = 0.85). However, the independent effect of the pubertal timing-associated marker appeared to not be restricted to puberty alone. In females, rs7759938 was associated with a distinct reduction of effect size already during prepuberty, between ages 6 and 9. In males, the effect sizes remained stable up until age 10. Multiple regression analysis of the change in relative height between ages 6 and 9 confirmed a significant female-specific effect (rs7759938 beta = −0.214, SE = 0.054, p = 6.65×10^{-5} and rs314277 beta = −0.023, SE = 0.135, p = 0.87; rs7759938 × rs314277 interaction beta = 0.145, SE = 0.093, p = 0.122; Table S3). At the other height-associated locus, rs314277, the A (derived) allele, previously associated with increased adult stature, consistently showed negative effect sizes throughout childhood and adulthood in males, whereas the female analysis showed an accumulating increase in effect sizes initiated already during infancy and reaching suggestive p value levels of 0.01 at ages 9 and 10 (Figure 2). Analyzing both males and females simultaneously also indicated suggestive evidence for genotype-sex interaction at rs314277, with p values of 0.03 and 0.01 at ages 9 and 10, respectively (Figure 2).
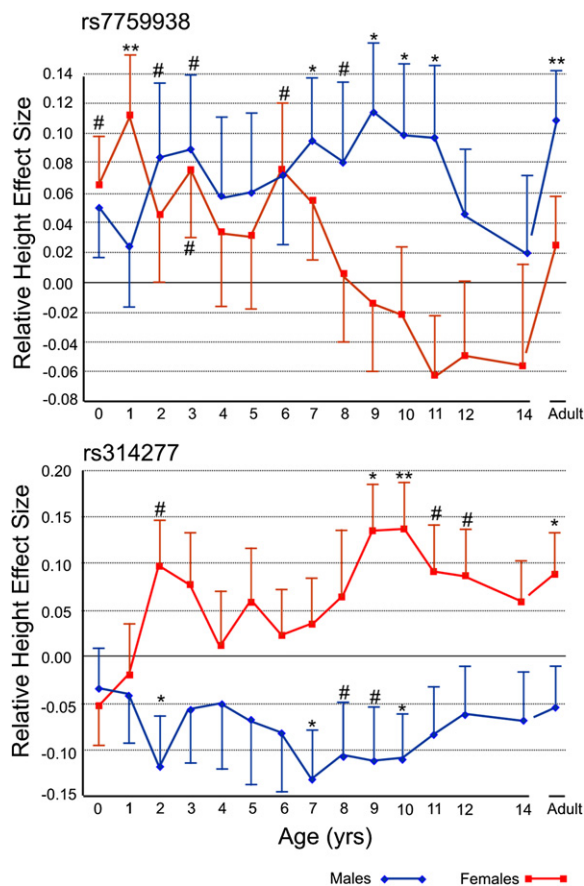
To further characterize the specific effects on height growth, we also conducted multiple regression analyses of the total magnitude of height growth during infancy, childhood, and puberty. The analyses did not reveal any statistically significant associations (Table S4), even though patterns consistent with the cross-sectional analyses of height could be observed. For example, rs314277, which associated significantly with both childhood and adult female stature, showed a positive correlation with

growth during all three periods of growth in females. Thus, both our cross-sectional analyses of childhood height and our follow-up analysis of height growth suggest distinct effects on height growth prior to puberty. Whereas rs7759938 is associated with more rapid changes in growth, taking effect during a shorter period of time than the age ranges covering the complete childhood or pubertal growth periods, rs314277 appears to tag a more slowly accumulating effect influencing height predominantly in females.

In light of our results, *LIN28B* emerges as an important overall regulator of postnatal height growth. Our primary genome-wide analysis identified significant association between the gene region and the timing of pubertal growth. However, our follow-up analyses show that the genetic effects are not restricted to puberty, even though both the current and previous studies[9–12] show a strong influence on pubertal timing. Furthermore, even if final height and pubertal growth are well correlated with each other, we found that *LIN28B* significantly influences height growth already during early childhood prior to the initiation of pubertal development. Moreover, the association with final height was not affected by controlling for age at menarche in the multiple regression model, neither at rs7759938 (n = 4645, beta = 0.02, SE = 0.031) nor at rs314277 (beta = 0.081, SE = 0.039). Even though all of our study populations are Finnish, we have reason to believe that our findings are also applicable to other populations of European descent, because both markers subjected to follow-up studies have similar allelic frequencies in the Finnish cohorts as in other European populations. Overall, our study highlights the complex genetic regulation underlying a multifactorial trait, which in this case could be successfully dissected by sex-specific analyses considering two partially correlated genetic effects. The strong sex interaction observed in the current study might in part explain why the previous large-scale association studies of final height detected only barely significant or suggestive association when analyzing both sexes together.[28,29] Sex-specific mechanisms certainly can be expected to mediate a noticeable influence on both height growth and final stature. Therefore, we anticipate that future sex-stratified association analyses of these traits will likely add to our understanding of the underlying regulatory network.

A major strength of the current study is the rather unique longitudinal height data available in multiple well-characterized cohorts. This setup enabled the inspection of multiple growth phases within the same individuals, even though the specific time points of the individual height measurements varied. All cohorts were Finnish, and they represented different geographical parts of the country. Testing for population stratification within the GWAS cohort showed no evidence for major substructure. Also, the two markers analyzed showed similar allelic frequencies across the country. Nevertheless, with respect to year of birth, there was a substantial range, from the 1920s to the

1960s, a circumstance which might impact both the pattern of height growth and the estimate of final height. For example, the environmental conditions influencing height growth can be expected to vary significantly between both the earlier and the latter part of the 20th century, and also between the southern industrialized part and the northern rural parts of the country. Furthermore, estimating final height in elderly study subjects, e.g., above 60 yrs of age, might be hampered by, for instance, changes in posture or compression of the discs between the vertebrae. Despite these factors, expected to increase random error, the genetic effects influencing height growth and final stature were remarkably similar between cohorts. In particular, the sex-specific effect influencing final height at rs314277 was almost identical in all cohorts, as were the effects on childhood growth (Table 2; Figure S4). The similarity of the association results among cohorts, in addition to the similarity of the association patterns between childhood and adulthood, make it unlikely that the observed genetic associations are spurious. Given the varying environmental factors, the consistency of the association results further suggests that the genetic effects are rather pervasive even though they explain only a minor proportion of the overall phenotypic variation.

The markers subjected to follow-up analysis both had genome-wide significant (p < 5 × 10^{-8}) prior evidence for exerting an effect on height growth. Nevertheless, multiple follow-up analyses were performed. Many of these tests were not independent because of the strong correlation between measurements of childhood height at consecutive years. After applying a conservative Bonferroni-corrected threshold, p = 0.005, correcting for ten independent sex-specific tests, the main findings of consistent evidence across cohorts for two independent growth effects with sex-specific action and for the influence on height growth prior to puberty remain significant. However, in the exploratory analyses we also report results with p < 0.05, although we realize that further evidence is needed to evaluate which of these suggestive signals are real. For example, the nominal evidence for association with birth length at rs7759938 in females and males (beta = 0.058, SE = 0.025, p = 0.02) may support an effect on fetal growth, an effect well in accordance with the expression pattern of *LIN28B*,[32] although we do not have sufficient power to further validate this putative effect in the current study.

Three published genome-wide selection scans, which used haplotype-based[33,34] and allele frequency spectrum statistics,[35] report signals suggestive of selection in the vicinity of *LIN28B*. Therefore, given a link between the timing of puberty and reproduction, the locus could be a candidate target for natural selection. In the presence of positive selection, one haplotype pattern is rapidly driven to high frequency, and typically the selected allele is therefore surrounded by a much longer haplotype stretch than the other alleles. At *LIN28B*, the derived allele A of rs7759938, associating with early pubertal timing, is

both common and present on a long haplotype extending toward the telomere over at least 100 kb (Figure S5). Visualization of the haplotype variation by median-joining network analysis[36] of the HapMap3 data also showed a network pattern consistent with positive selection of haplotypes carrying the derived early pubertal timing-associated allele A at rs7759938 (Figure S6). In contrast, the formal tests of a selective sweep (the iHS and F_{ST} statistics) indicated that the selection signal lies about 250 kb from the SNPs with the strongest association to pubertal timing (Figure S7). This might imply that the target of selection is distinct from the effect influencing pubertal timing. However, determining the precise target of selection remains a challenge because the localization of an iHS peak is influenced by several factors, including recombination rates and stochastic events. Also, our coalescent simulations of positive selection of a single variant in the middle of a genomic segment of 800 kb showed that the peak of the iHS landscape might be located as much as 200 kb away from the best selected variant, even assuming uniform recombination rates (data not shown; see Lappalainen et al.[34] for details of the simulations). Certainly pubertal timing, being closely linked to reproduction, is an intuitive candidate target for recent positive selection. Recent studies have indeed suggested that the short stature pygmy phenotype in part would be a by-product of selection for early reproduction.[37,38] Nevertheless, further studies are needed to fully characterize the role of selection in the evolutionary history of the *LIN28B* region.

In conclusion, we provide compelling evidence that the genomic region harboring *LIN28B*, recently associated with adult stature and timing of menarche,[9–12,28,29] regulates multiple aspects of human postnatal growth, with a particularly strong influence at puberty. Finding more than one independent genetic variant in the region tagging distinctive and sex-specific effects suggests that *LIN28B* might be an important switch regulating critical aspects of human growth and development. One of the markers tagging the effects is located in the 5′ upstream region of *LIN28B*, whereas the other one is intronic. Yet the causative variants are yet unknown, as are the molecular mechanisms mediating the effects. *LIN28B* is a homolog of a key regulator of developmental timing in *C. elegans*,[39] so the effect on pubertal timing could be mediated through a conserved pathway influencing a developmental clock mechanism. Recently, *LIN28B* was shown to inhibit miRNA *let-7* biosynthesis, resulting in upregulation of *let-7* targets.[32,39] Several *let-7* targets have previously been associated with final height.[28] Thus, *LIN28B* might influence height growth in part through altered expression of *let-7*-regulated genes, although functional studies are needed to validate this proposed mode of action.

Supplemental Data

Supplemental Data include eight figures and four tables and can be found with this article online at http://www.ajhg.org.

Web Resources

The URLs for data presented herein are as follows:

Online Mendelian Inheritance in Man, http://www.ncbi.nlm.nih.gov/Omim/
MetABEL Software 0.0-3, http://mga.bionet.nsc.ru/~yurii/ABEL/
The R Project for Statistical Computing, http://www.r-project.org/
Haploview, http://www.broadinstitute.org/mpg/haploview
International HapMap Project, http://hapmap.ncbi.nlm.nih.gov/
Hapmap 3 Data Release, http://www.sanger.ac.uk/humgen/hapmap3/
SNP Annotation and Proxy Search (SNAP), http://www.broadinstitute.org/mpg/snap/

References

1. Okasha, M., Gunnell, D., Holly, J., and Davey Smith, G. (2002). Childhood growth and adult cancer. Best Pract. Res. Clin. Endocrinol. Metab. *16*, 225–241.
2. Weedon, M.N., and Frayling, T.M. (2008). Reaching new heights: Insights into the genetics of human stature. Trends Genet. *24*, 595–603.
3. Silventoinen, K., Pietiläinen, K.H., Tynelius, P., Sørensen, 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.
4. Kaprio, J., Rimpelä, A., Winter, T., Viken, R.J., Rimpelä, M., and Rose, R.J. (1995). Common genetic influences on BMI and age at menarche. Hum. Biol. *67*, 739–753.
5. Lettre, G. (2009). Genetic regulation of adult stature. Curr. Opin. Pediatr. *21*, 515–522.
6. Cotsapas, C., Speliotes, E.K., Hatoum, I.J., Greenawalt, D.M., Dobrin, R., Lum, P.Y., Suver, C., Chudin, E., Kemp, D., Reitman, M., et al. GIANT Consortium. (2009). Common body mass index-associated variants confer risk of extreme obesity. Hum. Mol. Genet. *18*, 3502–3507.

7. Willer, C.J., Speliotes, E.K., Loos, R.J.F., Li, S., Lindgren, C.M., Heid, I.M., Berndt, S.I., Elliott, A.L., Jackson, A.U., Lamina, C., et al. Wellcome Trust Case Control Consortium; Genetic Investigation of ANthropometric Traits Consortium. (2009). Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat. Genet. *41*, 25–34.

8. Thorleifsson, G., Walters, G.B., Gudbjartsson, D.F., Steinthorsdottir, V., Sulem, P., Helgadottir, A., Styrkarsdottir, U., Gretarsdottir, S., Thorlacius, S., Jonsdottir, I., et al. (2009). Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. Nat. Genet. *41*, 18–24.

9. Ong, K.K., Elks, C.E., Li, S., Zhao, J.H., Luan, J., Andersen, L.B., Bingham, S.A., Brage, S., Davey Smith, G., Ekelund, U., et al. (2009). Genetic variation in LIN28B is associated with the timing of puberty. Nat. Genet. *41*, 729–733.

10. Sulem, P., Gudbjartsson, D.F., Rafnar, T., Holm, H., Olafsdottir, E.J., Olafsdottir, G.H., Jonsson, T., Alexandersen, P., Feenstra, B., Boyd, H.A., et al. (2009). Genome-wide association study identifies sequence variants on 6q21 associated with age at menarche. Nat. Genet. *41*, 734–738.

11. He, C., Kraft, P., Chen, C., Buring, J.E., Paré, G., Hankinson, S.E., Chanock, S.J., Ridker, P.M., Hunter, D.J., and Chasman, D.I. (2009). Genome-wide association studies identify loci associated with age at menarche and age at natural menopause. Nat. Genet. *41*, 724–728.

12. Perry, J.R.B., Stolk, L., Franceschini, N., Lunetta, K.L., Zhai, G., McArdle, P.F., Smith, A.V., Aspelund, T., Bandinelli, S., Boerwinkle, E., et al. (2009). Meta-analysis of genome-wide association data identifies two loci influencing age at menarche. Nat. Genet. *41*, 648–650.

13. Taponen, S., Martikainen, H., Järvelin, M.-R., Sovio, U., Laitinen, J., Pouta, A., Hartikainen, A.-L., McCarthy, M.I., Franks, S., Paldanius, M., Ruokonen, A., and Northern Finland Birth Cohort 1966 Study. (2004). Metabolic cardiovascular disease risk factors in women with self-reported symptoms of oligomenorrhea and/or hirsutism: Northern Finland Birth Cohort 1966 Study. J. Clin. Endocrinol. Metab. *89*, 2114–2118.

14. Laitinen, J., Power, C., and Järvelin, M.-R. (2001). Family social class, maternal body mass index, childhood body mass index, and age at menarche as predictors of adult obesity. Am. J. Clin. Nutr. *74*, 287–294.

15. Ylihärsilä, H., Kajantie, E., Osmond, C., Forsén, T., Barker, D.J., and Eriksson, J.G. (2008). Body mass index during childhood and adult body composition in men and women aged 56-70 y. Am. J. Clin. Nutr. *87*, 1769–1775.

16. Eriksson, J.G., Forsén, T., Tuomilehto, J., Osmond, C., and Barker, D.J. (2001). Early growth and coronary heart disease in later life: Longitudinal study. BMJ *322*, 949–953.

17. Raitakari, O.T., Juonala, M., Rönnemaa, T., Keltikangas-Järvinen, L., Räsänen, L., Pietikäinen, M., Hutri-Kähönen, N., Taittonen, L., Jokinen, E., Marniemi, J., et al. (2008). Cohort profile: The cardiovascular risk in Young Finns study. Int. J. Epidemiol. *37*, 1220–1226.

18. Kivimäki, M., Lawlor, D.A., Smith, G.D., Elovainio, M., Jokela, M., Keltikangas-Järvinen, L., Vahtera, J., Taittonen, L., Juonala, M., Viikari, J.S.A., and Raitakari, O.T. (2008). Association of age at menarche with cardiovascular risk factors, vascular structure, and function in adulthood: The Cardio-vascular Risk in Young Finns study. Am. J. Clin. Nutr. *87*, 1876–1882.

19. Heistaro S., ed. (2008). Methodology Report: Health 2000 Survey (Helsinki, Finland: National Public Health Institute).

20. Luo, Z.C., and Karlberg, J. (2000). Critical growth phases for adult shortness. Am. J. Epidemiol. *152*, 125–131.

21. Perry, R.J., Farquharson, C., and Ahmed, S.F. (2008). The role of sex steroids in controlling pubertal growth. Clin. Endocrinol. (Oxf.) *68*, 4–15.

22. Fischbein, S. (1977). Intra-pair similarity in physical growth of monozygotic and of dizygotic twins during puberty. Ann. Hum. Biol. *4*, 417–430.

23. 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.

24. Palmert, M.R., and Boepple, P.A. (2001). Variation in the timing of puberty: Clinical spectrum and genetic investigation. J. Clin. Endocrinol. Metab. *86*, 2364–2368.

25. 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.

26. Sabatti, C., Service, S.K., Hartikainen, A.-L., Pouta, A., Ripatti, S., Brodsky, J., Jones, C.G., Zaitlen, N.A., Varilo, T., Kaakinen, M., et al. (2009). Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. Nat. Genet. *41*, 35–46.

27. 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., and Sham, P.C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575.

28. Lettre, G., Jackson, A.U., Gieger, C., Schumacher, F.R., Berndt, S.I., Sanna, S., Eyheramendy, S., Voight, B.F., Butler, J.L., Guiducci, C., et al. Diabetes Genetics Initiative; FUSION; KORA; Prostate, Lung Colorectal and Ovarian Cancer Screening Trial; Nurses' Health StudySardiNIA. (2008). Identification of ten loci associated with height highlights new biological pathways in human growth. Nat. Genet. *40*, 584–591.

29. Gudbjartsson, D.F., Walters, G.B., Thorleifsson, G., Stefansson, H., Halldorsson, B.V., Zusmanovich, P., Sulem, P., Thorlacius, S., Gylfason, A., Steinberg, S., et al. (2008). Many sequence variants affecting diversity of adult human height. Nat. Genet. *40*, 609–615.

30. 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.

31. Sandhu, J., Ben-Shlomo, Y., Cole, T.J., Holly, J., and Davey Smith, G. (2006). The impact of childhood body mass index on timing of puberty, adult stature and obesity: A follow-up study based on adolescent anthropometry recorded at Christ's Hospital (1936–1964). Int. J. Obes. *30*, 14–22.

32. Moss, E.G., and Tang, L. (2003). Conservation of the heterochronic regulator Lin-28, its developmental expression and microRNA complementary sites. Dev. Biol. *258*, 432–442.

33. Pickrell, J.K., Coop, G., Novembre, J., Kudaravalli, S., Li, J.Z., Absher, D., Srinivasan, B.S., Barsh, G.S., Myers, R.M., Feldman, M.W., and Pritchard, J.K. (2009). Signals of recent positive selection in a worldwide sample of human populations. Genome Res. *19*, 826–837.

34. Lappalainen, T., Salmela, E., Andersen, P.M., Dahlman-Wright, K., Sistonen, P., Savontaus, M.-L., Schreiber, S.,

Lahermo, P., and Kere, J. (2010). Genomic landscape of positive natural selection in Northern European populations. Eur. J. Hum. Genet. *18*, 471–478.

35. Carlson, C.S., Thomas, D.J., Eberle, M.A., Swanson, J.E., Livingston, R.J., Rieder, M.J., and Nickerson, D.A. (2005). Genomic regions exhibiting positive selection identified from dense genotype data. Genome Res. *15*, 1553–1565.

36. Bandelt, H.J., Forster, A., and Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. *16*, 37–48.

37. Migliano, A.B., Vinicius, L., and Lahr, M.M. (2007). Life history trade-offs explain the evolution of human pygmies. Proc. Natl. Acad. Sci. USA *104*, 20216–20219.

38. Perry, G.H., and Dominy, N.J. (2009). Evolution of the human pygmy phenotype. Trends Ecol. Evol. *24*, 218–225.

39. Viswanathan, S.R., Powers, J.T., Einhorn, W., Hoshida, Y., Ng, T.L., Toffanin, S., O'Sullivan, M., Lu, J., Phillips, L.A., Lockhart, V.L., et al. (2009). Lin28 promotes transformation and is associated with advanced human malignancies. Nat. Genet. *41*, 843–848.