ORIGINAL ARTICLE

A male-specific quantitative trait locus on 1p21 controlling human stature

S Sammalisto, T Hiekkalinna, E Suviolahti, K Sood, A Metzidis, P Pajukanta, H E Lilja, A Soro-Paavonen, M-R Taskinen, T Tuomi, P Almgren, M Orho-Melander, L Groop, L Peltonen, M Perola

J Med Genet 2005;42:932–939. doi: 10.1136/jmg.2005.031278

See end of article for authors' affiliations

Correspondence to:
Dr Leena Peltonen,
Department of Medical
Genetics and Molecular
Medicine, University of
Helsinki and National
Public Health Institute,
Biomedicum Helsinki PL
104, FIN-00251 Helsinki,
Finland; Leena.Peltonen@
ki.fi

Revised version received
3 March 2005
Accepted for publication
7 April 2005

Background: Many genome-wide scans aimed at complex traits have been statistically underpowered due to small sample size. Combining data from several genome-wide screens with comparable quantitative phenotype data should improve statistical power for the localisation of genomic regions contributing to these traits.

Objective: To perform a genome-wide screen for loci affecting adult stature by combined analysis of four previously performed genome-wide scans.

Methods: We developed a web based computer tool, Cartographer, for combining genetic marker maps which positions genetic markers accurately using the July 2003 release of the human genome sequence and the deCODE genetic map. Using Cartographer, we combined the primary genotype data from four genome-wide scans and performed variance components (VC) linkage analyses for human stature on the pooled dataset of 1417 individuals from 277 families and performed VC analyses for males and females separately.

Results: We found significant linkage to stature on 1p21 (multipoint LOD score 4.25) and suggestive linkages on 9p24 and 18q21 (multipoint LOD scores 2.57 and 2.39, respectively) in males-only analyses. We also found suggestive linkage to 4q35 and 22q13 (multipoint LOD scores 2.18 and 2.85, respectively) when we analysed both females and males and to 13q12 (multipoint LOD score 2.66) in females-only analyses.

Conclusions: We strengthened the evidence for linkage to previously reported quantitative trait loci (QTL) for stature and also found significant evidence of a novel male-specific QTL on 1p21. Further investigation of several interesting candidate genes in this region will help towards characterisation of this first sex-specific locus affecting human stature.

The search for genes underlying human complex traits has proven to be a more formidable task than many had anticipated. For the identification of novel genes and loci, linkage analysis using marker maps spanning the whole genome represents a hypothesis-free approach. However, success in consistent identification of chromosomal regions involved in the aetiology of complex traits has been relatively modest. Several factors contribute to this: (1) environmental factors make a significant contribution to complex phenotypes; (2) misdiagnoses and phenocopies introduce noise into the analyses, and (3) the linkage analysis has inherent limitations in detecting genes of modest to moderate effect.¹ In a review of 101 published genome-wide scans carried out to identify loci for 31 different complex phenotypes, it was concluded that the most critical parameters associated with increased success in locus finding were an increase in sample size and an attempt to genetically homogenise the study sample (that is, draw the sample from a single ethnic group).² Combining data from previously conducted genome scans in various study samples is an obvious step to increase the sample size. This has proven to be a powerful strategy in gene mapping of several complex traits.^{3–6}

There are a number of measurable quantitative phenotypes which are systematically collected in conjunction with genome scans but are merely used for ascertainment or adjustment purposes. With proper informed consent, these data should also be used for genetic analyses. We have chosen one such trait, stature (that is, adult height), for quantitative trait locus (QTL) analysis; stature is one of the most frequently collected traits in study samples and is thus

highly suitable for pooling of data across studies. Stature as a quantitative trait has also some advantages for genetic analysis: it is highly heritable with an estimated contribution of genetic components ranging from 68% to 93% of the total phenotypic variance^{7–8} and it can be easily and unambiguously measured, thus minimising the noise produced by measurement errors. Our strategy was to combine the primary genotypes produced in four genome scans performed in families ascertained for type II diabetes,^{9,10} familial combined hyperlipidaemia (FCHL),^{11,12} and low high-density lipoprotein cholesterol (HDL-C),^{13,14} and carry out a combined linkage analysis for stature in the pooled dataset of 1417 individuals. The linkage analysis of the pooled study sample provided suggestive evidence for two genetic loci on 4q35 and 22q13, controlling stature when phenotypes of both sexes were used in the analyses. Sex-specific linkage analyses implied three genetic loci on 1p21, 9p24, and 18q21 controlling male stature and one on 13q12 controlling female stature. After assessment of the empiric genome-wide significance for each of these findings, only the locus on 1p21 linked to male stature remained significant ($p < 0.01$).

METHODS

Study subjects

Genome-wide genotype information and phenotype data were available for 1417 individuals from 277 families. In

Abbreviations: FCHL, familial combined hyperlipidaemia; HDL-C, high-density lipoprotein cholesterol; HME, hereditary multiple exostosis; QTL, quantitative trait locus; VC, variance components

addition, the genotypes of 117 individuals who did not have phenotypic information were retained for the analyses. Study families were originally ascertained for type II diabetes (Botnia and Helsinki study samples), familial combined hyperlipidaemia (FCHL study sample), and low high-density lipoprotein cholesterol (Low HDL-C study sample). More detailed description of the patient selection is available elsewhere.⁹⁻¹⁴ Written informed consent was obtained from each study subject prior to participating in the study. Each separate study design was approved by local ethics committees. Personal information identifying the participants was not made available to the investigators performing the genetic analyses.

The contributions of different study groups are as follows: 416 individuals from 58 families (mean family size 10.78) and 668 individuals from 165 families (mean family size 7.63) from the two separate study groups ascertained for type II diabetes, 139 individuals from 29 families ascertained for FCHL (mean family size 7.09), and 194 individuals from 25 families ascertained for low HDL-C (mean family size 11.24). Body height for the majority of individuals was measured at health care centres with subjects in light clothing without shoes, but for some individuals in the FCHL and low HDL-C study groups only self reported height was available. However, it has been shown in Finnish twins that the correlation between self reported and measured height is 0.98 in males and 0.96 in females making self reported height a reliable phenotype.¹⁵ The genotype and phenotype information contribution of the study samples are summarised in table 1 and the numbers of informative relationships in the combined sample are summarised in table 2.

Genotyping and marker map construction

Genotypes were produced as described earlier.^{9 11 12 16} Original genome-wide screens combined here were genotyped with different marker sets; the Botnia and Helsinki samples were genotyped with microsatellite loci selected from public databases to provide an average marker spacing of ~10 cM, while the FCHL and Low-HDL samples were genotyped using Weber screening set 9¹⁷ and a modified version of Weber screening set 6,¹⁷ respectively.

Since the original genome scans differed in their selection of markers, the first task was to construct a merged, well ordered marker map across the four different studies. For the purpose of marker map construction, we developed a web based tool, Cartographer, which is available on our website

Table 2 Number of informative relative pairs where both members have genotype and phenotype information

	Female-Female	Male-Male	Opposite sex pairs	All pairs
Sib pairs	376	304	648	1328
Half sibs	12	5	21	38
Cousins	57	53	121	231
Parent-child	144	104	259	507
Grandparent-grandchild	9	1	18	28
Avuncular	168	147	350	665
Total	766	614	1417	2797

(<http://www.bioinfo.helsinki.fi/cartographer>). In brief, Cartographer first looks up the physical location of the markers on a locally stored copy of the UCSC database (<http://genome.ucsc.edu/>) and orders the markers based on the sequence information. The genetic location of each marker is next defined using the published deCODE genetic map locations,¹⁸ which are also stored in the UCSC database. For markers not included in the deCODE genetic map, Cartographer uses linear interpolation for obtaining estimates of the genetic locations of these markers by using the physical and genetic locations of the immediately flanking deCODE markers. If the physical location from the UCSC database and the genetic location from the deCODE genetic map for a given marker are in disagreement, Cartographer obtains an estimate of the genetic location via interpolation using the nearest flanking deCODE markers that are in agreement with the sequence information. Finally, the marker genotype information for each genome scan is combined to produce one shared marker map.

Error checking

Familial relationships were verified using the GRR program.¹⁹ The genotype and phenotype data of pairs whose reported relationship were likely to have been misspecified were removed prior to analysis. Genotypes were checked for Mendelian inconsistencies with the program PedCheck²⁰ and for Mendelian consistent genotyping errors with the error checking option implemented in MERLIN.²¹ Some 0.1% of the genotypes were determined by MERLIN to be unlikely and were removed prior to analyses using the Pedwise program distributed with the MERLIN package.

Table 1 Demographic characteristics of the study groups

Study group	Phenotypes	Genotypes	Both	Age* (years)		Stature (cm)		Skewness	Kurtosis
				Mean (SD)	Range	Mean (SD)	Range		
Botnia									
Males	216	217	211	58.7 (14.4)	23.6–84.2	174.7 (5.9)	157.0–193.0	0.18	0.42
Females	208	226	205	58.2 (14.8)	23.5–87.2	161.2 (5.9)	143.0–181.0	0.04	0.25
Helsinki									
Males	337	293	288	56.6 (12.6)	25.9–82.4	174.4 (6.7)	153.7–197.0	–0.01	0.33
Females	432	391	380	60.1 (12.8)	23.1–90.6	160.0 (6.3)	136.0–184.0	–0.03	0.70
FCHL									
Males	64	79	63	47.0 (12.7)	23.7–85.4	175.7 (5.9)	164.0–191.0	0.18	–0.07
Females	76	103	76	51.4 (13.6)	23.0–82.7	161.5 (6.2)	147.0–176.0	–0.07	–0.08
Low HDL-C									
Males	99	108	97	49.3 (14.3)	23.2–82.4	176.7 (6.2)	162.0–190.0	–0.12	–0.33
Females	99	111	97	50.0 (14.1)	23.5–90.7	164.2 (6.2)	149.5–179.0	0.21	–0.01
Combined									
Males	716	697	659	55.4 (14.0)	23.2–85.4	174.7 (6.4)	153.7–197.0	0.03	0.23
Females	815	831	758	57.6 (14.1)	23.0–90.7	161.0 (6.3)	136.0–184.0	0.01	0.46
Total	1531	1528	1417	56.6 (14.1)	23.0–90.7	167.4 (9.4)	136.0–197.0	0.07	–0.49

*Individuals <23 years of age excluded.
SD, standard deviation.

Statistical analysis

Preliminary analysis of the stature distribution was performed with SPSS 11.5 (SPSS, Chicago, IL). In the combined study group stature followed a standard normal distribution. The trait distribution for each study group as well as the combined study group are summarised in table 1. The sex and age of each study subject at the time of phenotype determination were used as covariates in the variance components model. QTL analyses were performed using the variance components method implemented in the program MERLIN. The empirical significance of the obtained linkage results was determined using the simulate option in MERLIN as described in the Results section.

RESULTS

The mean family size in the combined sample was 8.56, reflecting the large average size of the pedigrees and the high number of potentially informative meiotic events in the pooled study sample. The phenotypic characteristics of each study sample as well as the combined data set are summarised in table 1. The phenotypes of the study subjects less than 23 years of age ($n = 57$) were excluded from the statistical analyses since these individuals may still be growing. However, all the available genotypes were included in the analyses to maximise linkage phase information. Outliers in the combined sample (defined as deviating four standard deviations (SD) or more from the sex-specific mean) were excluded from the analyses. The heritability of stature in the combined study group was ≈ 0.84 , which closely follows previous estimates.^{7, 8}

Although the number of the markers genotyped in four genome scans was an impressive 1127, only 159 markers (14%) were genotyped in the combined sample. A total of 686 markers ($\sim 62\%$) were genotyped in only one study sample, 142 ($\sim 13\%$) in two, and 120 ($\sim 11\%$) in three study samples. After precisely ordering all the markers genotyped in various genome scans on the genetic map using the Cartographer program developed here, 20 out of 1127 markers genotyped in the scans ($\approx 1.8\%$ of all markers) could not be accurately positioned due to missing or ambiguous sequence information and they were excluded from the analysis. The average marker heterozygosity was 0.74 (SD 0.02) and the average information content of the markers in the combined data set was 0.47 (SD 0.08).

Since stature is affected by both sex and age, these parameters were used as covariates in the variance components model. Also, to control for potential heterogeneity between the different study samples, we used the study sample identifier (Botnia, Helsinki, FCHL, or Low HDL-C) as a covariate in the variance-component model. We also examined if affection status for type 2 diabetes, FCHL, or familial Low-HDL affects adult stature by adding this information to the variance components model; this, however, did not affect the results significantly (data not shown) and the results reported here were obtained using age, sex, and study sample identifier as covariates in the model. Multipoint LOD scores over 2.0 were observed with markers located on 4q and 22q where the highest LOD scores were obtained with markers D4S426 (LOD score 2.18) and D22S282 (LOD score 2.85) in the combined sample. In the two-point variance component analyses, the LOD scores for the markers were 1.15 for D4S426 and 1.49 for D22S282.

We also analysed stature for both genders separately by ignoring the phenotype data of all individuals from one gender while retaining genotypes from both genders. In these sex-specific multipoint analyses, three loci showed evidence for linkage in males: 1p21 with marker D1S1631 (LOD score 4.25), 9p24 with marker D9S2169 (LOD score 2.57), and 18q21 with marker D18S60 (LOD score 2.39). In females

these loci showed no evidence of linkage: the respective multipoint LOD scores were 0.00 for D1S1631, 0.00 for D9S2169, and 0.00 for D18S60. The LOD scores for the markers in the two-point variance component analyses in males were 4.81 for D1S1631, 3.17 for D9S2169, and 1.74 for D18S60. The two-point LOD scores for these markers in females were 0.11 for D1S1631, 0.00 for D9S2169, and 0.00 for D18S60. One locus on 13q12 at marker D13S221 (multipoint LOD score 2.66, two-point LOD score 1.02) showed suggestive evidence for linkage in females. No evidence for linkage was found at this marker (multipoint and two-point LOD scores 0.00) in males.

To assess the significance of the obtained linkage results, we simulated 100 genome scans using the simulate option of MERLIN. This option generates simulated chromosomes for the founders conditional on the allele frequencies at each marker and segregates these in the pedigrees described in the actual data set while retaining the genetic map as well as the missing data patterns and phenotypes for each individual. The same analyses that were conducted with the actual data (multi- and two-point VC analyses on both sexes separately and jointly) were repeated for each of the simulated genomes. Since any evidence of linkage found in the simulated genomes is due to chance, only these simulations allowed us to evaluate the false-positive rate. These simulations were automated using AUTOGSCAN.²² Since not all of the individual multipoint LOD scores are independent, we determined the empirical genome-wide significance of a given LOD score as the fraction of simulated genome scans in which this LOD score was reached or exceeded. The most significant LOD scores and their empirical genome-wide significance are listed in table 3. The only linkage that remained significant was the linkage to male stature on 1p21 ($p < 0.01$). However, it must be noted that since we were restricted to performing only 100 simulated genomes due to computational restraints, the confidence intervals for the empirical genome-wide significances reported in table 3 are quite large. For example, using the recommended^{23, 24} approach by Wilson,²⁵ the corresponding 95% confidence intervals for empirical p values $p < 0.01$, $p = 0.08$, and $p = 0.10$ are 0 to 0.037, 0.041 to 0.150, and 0.055 to 0.174, respectively.

DISCUSSION

Many loci have been implicated for linkage to stature in a number of genome-wide screens,^{3, 26–33} with some chromosomal regions such as 5q31, 6q25, and 7q35–36 being reported by multiple authors (see table 4). We combined the primary genotype and phenotype data of four genome scans comprising 1417 individuals from 277 families in study samples ascertained for type II diabetes (Botnia and Helsinki samples), familial combined hyperlipidaemia (FCHL sample), and low high-density lipoprotein cholesterol (HDL-C sample), and performed variance components linkage analysis for stature (fig 1). Pooling of primary data usually yields greater statistical power and also fewer false-positives than meta-analysis of initial results because of the reduced number of statistical tests performed.^{34, 35} Since some genetic effect on stature is expected to be sex-specific, we also analysed males and females separately in our combined sample.

The original genome-wide screens differed in their choice of genetic markers, so prior to combining the genotype data it was necessary to place all the markers genotyped in the original samples in a shared marker map. Because the power and accuracy of multipoint linkage analysis to map trait loci is highly dependent on the genetic map used,^{36–38} we made every effort to utilise the most current data in the construction of the marker map. Since the order of markers in genetic maps is obtained from observed meiotic information by statistical inference, physical evidence of marker order

Table 3 Chromosomal regions showing multipoint LOD scores ≥ 2.0					
Gender	Chromosome	Peak marker	deCODE (cM)	Multipoint LOD score	Empirical genome-wide significance*
Males and females	4	D4S426	202.69	2.18	0.33
	22	D22S282	50.81	2.85	0.08
Females	13	D13S221	16.26	2.76	0.10
Males	1	D1S1631	125.75	4.25	<0.01
	9	D9S2169	12.55†	2.57	0.19
	18	D18S60	86.76†	2.39	0.33

*Determined from 100 simulated genome scans; †genetic distance interpolated from the neighbouring deCODE markers.

obtained from the genome sequence may be more accurate. Therefore, we carried out the construction of the common marker map in three hierarchical stages utilising a custom-made software tool: (1) we ordered the markers according to their physical location obtained from the July 2003 release of the human genome sequence stored in the UCSC database (<http://genome.ucsc.edu/>); (2) we obtained the genetic positions of the markers from the deCODE genetic map; and (3) for those markers which were not included in the deCODE marker map or were in disagreement with the sequence information, we used linear interpolation for genetic location estimation using the physical and genetic locations of the immediately neighbouring deCODE markers.

This approach for constructing genetic maps based on sequence information is similar to the strategy applied by Nievergelt *et al.*³⁹ However, to our knowledge this study is the first in which a sequence based genetic map has been used to guide the multipoint linkage analyses.

We found suggestive linkage on 4q35 and 22q13 when we used phenotypic information from both males and females adding sex and age as covariates in the variance components model; however, neither of these was significant according to simulations. Our results added additional support for the linkage on 4q35 reported by Hirschhorn *et al.*²⁷ who noted multipoint LOD scores of 1.73 at marker D4S1540 in their Finnish cohort (included in our combined sample and

Figure 1 Results of the multipoint variance components linkage analyses. The y axis represents the multipoint LOD score and the x axis the genetic location according to the deCODE genetic map. Black lines show the results using all available phenotypes, red lines the results using only female phenotypes, and blue lines the results using only male phenotypes.

Table 4 Overview of results from genome-wide QTL screens for stature where multipoint LOD scores over 2.0 have been reported

Locus	Peak marker	deCODE cM	LOD score	Population	Reference
1p21	D1S1631	125.75	2.25	GENOA African American	Wu <i>et al</i> ⁶
2q11	D2S113	112.79*	2.23	Finland (Botnia)	Hirschhorn <i>et al</i> ²⁷
3p26	D3S1297–D3S1304	4.94*–19.61	3.17	United Kingdom	Wiltshire <i>et al</i> ³⁰
3p14	D3S1766	79.08	2.31	Finland	Hirschhorn <i>et al</i> ²⁷
3q23	D3S1764	145.53	2.03	Multiple	Wu <i>et al</i> ⁶
3q26	D3S1763	168.52	2.06	GENOA European American	Wu <i>et al</i> ⁶
4q25	D4S1564	111.18	2.28	Finland (Botnia)	Hirschhorn <i>et al</i> ²⁷
5p14–p13	D5S2845–D5S1470	41.83*–53.87	2.04	Dutch	Willemsen <i>et al</i> ⁴¹
5q31	D5S816	138.04	2.26	HyperGen European American	Wu <i>et al</i> ⁶
5q31	144 cM (ABI)†	NA	2.14	Nebraska	Deng <i>et al</i> ³⁹
6q12	D6S1053	81.09	2.66	GENOA European American	Wu <i>et al</i> ⁶
6q12–q14	D6S1053–D6S1031	81.09*–89.14	2.32	Netherlands	Willemsen <i>et al</i> ⁴¹
6q25	D6S2436	161.27	3.06	Netherlands	Xu <i>et al</i> ³¹
6q25	D6S1007	166.61	3.85	Finland (Botnia)	Hirschhorn <i>et al</i> ²⁷
7q21	D7S669–D7S630	89.39*–99.93	2.26	United Kingdom	Wiltshire <i>et al</i> ³⁰
7q35	D7S2195	152*	3.40	Sweden	Hirschhorn <i>et al</i> ²⁷
7q36	165 cM (Marshfield)†	NA	2.91	Finland	Perola <i>et al</i> ²⁸
7q36	D7S3058	175.87*	2.46	Multiple	Wu <i>et al</i> ⁶
8q24	D8S1100–D8S373	148.81–164.47	2.52	Finland	Hirschhorn <i>et al</i> ²⁷
9q21	D9S1868	53.04	2.01	Finland (Botnia)	Hirschhorn <i>et al</i> ²⁷
9p1	D9S301	66.99	2.09	Netherlands	Xu <i>et al</i> ³¹
9q34	159 cM (Marshfield)†	NA	2.61	Finland	Perola <i>et al</i> ²⁸
12p13	D12S341	0.68*	2.07	Finland	Hirschhorn <i>et al</i> ²⁷
12q13	D12S1090–D12S398	NS–67.42	3.35	Finland	Hirschhorn <i>et al</i> ²⁷
13q33	D13S779–D13S797	95.76–105.22	3.56	Finland	Hirschhorn <i>et al</i> ²⁷
14q11	GATA74E02A	13.41	2.45	Framingham Heart Study, 70 year olds	Beck <i>et al</i> ⁴²
14q23	D14S592	60.5	3.67	GENOA European American	Wu <i>et al</i> ⁶
17q21	D17S958	74.32	2.69	Finland (Botnia)	Hirschhorn <i>et al</i> ²⁷
20p12	D20S66–D20S98	38.23–40.76	3.00	Pima Indians	Thompson <i>et al</i> ³⁶
20q13	D20S96	66.55	2.51	Finland (Botnia)	Hirschhorn <i>et al</i> ²⁷

*Genetic distance interpolated using the neighbouring deCODE markers; †authors did not report peak marker.

NA: deCODE genetic location could not be obtained; NS: sequence information not available in the UCSC database.

labelled “Helsinki” here) and 1.89 at markers D4S3051–D4S426 in their Botnia cohort (also included in our combined sample). In our combined sample the multipoint LOD scores were 1.55 for D4S1540, 2.12 for D4S3051, and 2.18 for D4S426, providing additional support for linkage to stature at the latter locus. Also, the linkage reported by Hirschhorn *et al* at marker D22S282 (multipoint LOD score 1.10) in the Botnia cohort was further corroborated by our findings; in our combined sample the multipoint LOD score at D22S282 was 2.85. Although neither of these loci seems significant according to our simulations (genome-wide significance 0.23 and 0.06 for D4S426 and D22S282, respectively), it is noteworthy that evidence for linkage was increased in our combined sample, providing further support for the contribution of these loci in the determination of adult height.

We identified a suggestive locus linked to female stature in our combined analysis of 13q12 at marker D13S221 (multipoint LOD score 2.66, empiric *p* value 0.06). This genome region was also implicated by Hirschhorn *et al* in their Finnish cohort (included in our sample and labelled “Helsinki” here). They reported a LOD score of 1.01 at markers D13S221–GGAA29H03 when the analyses were performed using the phenotypes from both genders (388 females and 314 males). They did not perform sex-limited analyses and their study sample contained approximately half as many females as our sample (388 *v* 758). This would probably diminish the power of linkage assuming that the locus truly is female-specific.

In males, the three loci 1p21, 9p24, and 18q21 showed some evidence for linkage to stature. Marker D1S1631 was significantly linked to male stature according to our simulations (multipoint LOD score 4.25, empiric *p* value <0.01). This locus has been implicated previously for suggestive linkage to stature by Wu *et al*⁶ who reported a LOD score of 2.25 for stature on 1p21 at the same marker D1S1631 in an

African American population ascertained for essential hypertension. However, in their sample only 34.5% were males and they did not perform sex-limited analyses. Our linkage result on 9p24 (multipoint LOD score 2.57, empiric *p* value 0.15) is also highly interesting since Mukhopadhyay *et al*,³³ who analysed sex-limited effects in a manner comparable to ours, reported a male-specific linkage finding on chromosome 9, approximately 40 cM away from our linkage peak (multipoint LOD score 1.65 at marker D9S319). It is also interesting that the locus linked to male stature on 18q21 (multipoint LOD score 2.39 at marker D18S60) is located ~20 cM upstream of linkage peaks reported by Hirschhorn *et al*: D18S541–D18S1121 in their Botnia cohort (multipoint LOD score 1.58) and D18S1121 in their Finnish cohort (multipoint LOD score 1.77).

Although a substantial amount of our dataset (416 individuals from 58 Botnia families and 668 individuals from 165 Helsinki families) overlaps the study sample analysed by Hirschhorn *et al*, we were unable to further corroborate other loci identified by them in our combined study sample. Previous reports of QTL genome scans have demonstrated that linkage peaks found in a subset of samples may not be identified in a combined sample.^{3, 40} There are at least two likely reasons for this result. First, our combined analysis used a different marker map and second, there may be locus heterogeneity between the study samples included in our combined analysis. The chance of the latter occurring does not seem very likely since all the study individuals originate from the fairly homogeneous Finnish population.

Perola *et al* reported two suggestive linkage peaks for stature on 7q36 and 9q34 in a combined sample of five Finnish genome-wide scans with total of 580 individuals.²⁸ One of these study samples, the FCHL sample, was also included in our combined sample. In our analyses, these loci did not provide evidence for linkage; the maximum LOD

scores on these chromosomes were 0.50 on 7q21 and 0.75 on 9q21. This is not too surprising since the relatively small FCHL sample did not significantly contribute to the earlier results (personal communication). The region on 7q was also reported by Hirschhorn *et al* in a Swedish sample, which was not included in our combined analysis, and by Wu *et al*³ in a combined sample of multiple populations. Even though our results did not yield substantial evidence for linkage to 7q, there is accumulative evidence that there are true loci on 7q that contribute to the variation in adult stature.

In our combined sample we observed some interesting minor peaks on chromosome 6 which would provide some support for previous reports of putative QTLs for stature. Xu *et al*³¹ reported a LOD score of 3.06 at marker D6S2436 in a Dutch sample, and Hirschhorn *et al* reported a LOD score of 3.85 at marker D6S1007 located ~5 cM downstream of D6S2436 in a Botnia population. Both of these markers were included in our combined sample and the respective LOD scores obtained were 0.79 and 1.05. Interestingly, when we analysed only males, the corresponding LOD scores were 1.19 and 1.49 at these markers. Another region on chromosome 6 has also been indicated for stature by two other studies by Willemsen *et al*⁴¹ and Wu *et al* which yielded overlapping results at marker D6S1053 with LOD scores of 2.32 and 2.66, respectively. This marker showed no evidence for linkage to stature in our sample (multipoint LOD score 0.05), but we observed suggestive linkage ~50 cM upstream of this marker where the highest LOD score obtained was 1.80 at marker D6S1959. Since it is known that the confidence intervals for QTLs implicated in genome-wide screens may cover wide chromosomal regions,⁴² it is possible that our finding represents the same underlying QTL for stature.

The one LOD drop critical region for the male-specific linkage peak on 1p21 spans approximately 121–131 cM in the deCODE genetic map, and is flanked by markers DIS2753 (120 cM) and DIS418 (133.99 cM). This region contains tens of known genes as well as numerous predicted genes whose function is currently unknown. Not surprisingly, this region harbours three candidate genes of special interest for stature: *EXTL2*, *COL11A1*, and *CH13L2*. *EXTL2* is an alpha-1, 4-N-acetylhexosaminyltransferase involved in the biosynthesis of heparin/heparan sulfate and is a member of the hereditary multiple exostoses gene family of tumour suppressors.^{43–45} Mutations in two other members of this gene family, *EXT1* and *EXT2*, have been shown to cause hereditary multiple exostosis (HME) syndrome types I (*EXT1*; MIM 133700) and II (*EXT2*; MIM 133701). Approximately 50% of patients with HME show short stature, linking this gene to abnormal growth.^{45–46} Interestingly, Wicklund *et al*⁴⁷ showed a significant excess of male patients (104:76) and male probands (28:15) with HME, suggesting some sex-specific effect. Another interesting positional candidate gene *COL11A1* encodes the alpha-1 subunit of collagen XI and is mostly expressed in cartilage tissue. Li *et al*⁴⁸ showed that the limb bones of newborn *COL11A1* null mice are wider at the metaphyses than normal bones and only about half the normal length, thus concluding that this gene is essential for normal skeletal morphogenesis. In humans, a splicing defect in *COL11A1* has been shown to associate with Marshall syndrome (MIM 154780) which includes symptoms such as short stature and a stocky build relative to unaffected family members.⁴⁹ Mutations in *COL11A1* have also been shown to cause a related condition, Stickler syndrome, type II (*STL2*; MIM 604841) with hallmark features of skeletal manifestations. A third positional candidate gene, *CH13L2*, encodes a chondrocyte protein which is highly expressed in cartilage chondrocytes. Benz *et al*⁵⁰ showed that the up regulation of this gene characterises the stimulation of chondrocyte growth. Since all these genes are likely to play important

roles in the normal function of growth plates of long bones, they also represent excellent functional candidate genes for fine mapping and gene identification efforts for human growth genes. The character of this seemingly male-specific locus on 1p21 influencing stature in our sample remains to be elucidated. It is possible that some gene(s) in the region are regulated via sex hormone related signalling pathways or have epistatic interactions between genes that are expressed differentially in males and females. We made an effort to localise binding sites for transcription factors that would link these candidate genes to sex hormone mediated pathways. We used phylogenetic foot printing methods with pair wise comparisons between human and mouse and between human and dog, using two independent libraries of documented transcription factor binding sites (TRANSFAC⁵¹ and JASPAR⁵²) and four different analysis algorithms (CONREAL,⁵³ CONREAL and LAGAN,⁵⁴ ConSite,⁵⁵ and rVista⁵⁶). In addition to these phylogenetic foot printing methods, we also searched for nuclear hormone receptor binding sites utilising NHR-scan⁵⁷ which implements a Hidden Markov Model approach for detecting putative binding sites. Using these computational methods, we found a high number of potential binding sites within these genes and their putative promoter regions (data not shown). However, due to the well known false-positive rate of contemporary bioinformatics methods including computational binding site prediction, these sites would have to be experimentally verified, which falls beyond the scope of this study.

For traits such as stature where there is a marked difference in the trait mean between sexes, it may be warranted to carry out sex-limited analyses in order to locate loci with sex-specific effects. So far, this type of approach is fairly uninvestigated in genome-wide scans aimed at complex traits. The X chromosome has been implicated in at least one genome-wide scan for stature²⁹ as well as by sex chromosome disorders such as Turner's syndrome patients and XYY men where stature is affected. Unfortunately, since our samples were genotyped only with autosomal markers, we could not investigate linkage of loci on sex chromosomes to stature. However, studies of mother-daughter correlations for females affected with Turner's syndrome have shown that it is likely that autosomes also contain loci contributing to the normal variation in stature⁵⁸ and it is possible that some of these loci are expressed differentially in males and females. In this study we were able to find significant evidence for a male-specific locus, but only suggestive evidence for female-specific loci linked to stature. Albeit assortative mating is clearly an important confounding factor affecting heritability estimates of height⁵⁹ and might mask the effects of environment, no sex-specific effects for assortative mating were seen in a previous Finnish study,⁶⁰ so its effect probably does not explain the male-specific finding here. There may, however, be stronger male-specific genetic determinants controlling stature as suggested by the finding that in females the effect of environmental factors may contribute more to variation in stature compared to men.⁸ The locus on 1p21 identified in this study may represent such a QTL, but further studies are needed to confirm this.

ACKNOWLEDGEMENTS

We would like to thank all the patients and their families for participation in this study. Also we wish to thank Professor JD Terwilliger for his helpful comments in preparing this manuscript, S Anjala, P Ellonen, H Keränen, and M Levander for their technical assistance, J Saharinen and T Ripatti for their expert advice on bioinformatics as well as the Finnish IT centre for science (CSC) for providing us with sufficient computational power for the determination of empirical p values.

ELECTRONIC-DATABASE INFORMATION

Cartographer is available from <http://www.bioinfo-helsinki.fi/cartographer> and the UCSC database can be found at <http://genome.ucsc.edu/>.

Authors' affiliations

S Sammalisto, T Hiekkalinna, E Suviolahti, K Sood, A Metzidis, H E Lilja, Department of Molecular Medicine, National Public Health Institute, Helsinki, Finland

P Pajukanta, Department of Human Genetics, David Geffen School of Medicine at the University of California, Los Angeles, CA, USA

A Soro-Paavonen, M-R Taskinen, T Tuomi, Department of Medicine, Helsinki University Central Hospital, Helsinki, Finland

P Almgren, M Orho-Melander, L Groop, Department of Endocrinology, Wallenberg Laboratory, University Hospital MAS, Lund University, Malmö, Sweden

L Peltonen, M Perola, Department of Molecular Medicine, National Public Health Institute, Biomedicum Helsinki and Department of Medical Genetics, University of Helsinki, Finland

This work was funded by the Center for Excellence in Disease Genetics of the Academy of Finland and Biocentrum Helsinki (SS, TH, ES, KS, AM, HEL, LP, and MP), NIH grant R01HL70150-OIA1 (LP), EC GenomEII twin project grant QL62-CT-2002-01254 (LP, MP, and SS), the Helsinki Biomedical Graduate School (SS), and the Emil Aaltonen Foundation (TH). LCG, TT, and the Botnia Study are principally supported by the Sigrid Juselius Foundation, the Academy of Finland, the Finnish Diabetes Research Foundation, The Folkhalsan Research Foundation, the EC (BM4-CT95-0662, GIFT), the Swedish Medical Research Council, the JDF Wallenberg Foundation, Novo Nordisk Foundation, Whitehead Institute, and Millennium Pharmaceuticals.

Competing interests: none declared

REFERENCES

- Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996;**273**(5281):1516-7.
- Altmüller J, Palmer LJ, Fischer G, Scherb H, Wjst M. Genomewide scans of complex human diseases: true linkage is hard to find. *Am J Hum Genet* 2001;**69**(5):936-50.
- Wu X, Cooper RS, Boerwinkle E, Turner ST, Hunt S, Myers R, Olshen RA, Curb D, Zhu X, Kan D, Luke A. Combined analysis of genomewide scans for adult height: results from the NHLBI Family Blood Pressure Program. *Eur J Hum Genet* 2003;**11**(3):271-4.
- Cavanagh J. International collaboration provides convincing linkage replication in complex disease through analysis of a large pooled data set: Crohn disease and chromosome 16. *Am J Hum Genet* 2001;**68**(5):1165-71.
- Cox NJ, Wapelhorst B, Morrison VA, Johnson L, Pinchuk L, Spielman RS, Todd JA, Concannon P. Seven regions of the genome show evidence of linkage to type 1 diabetes in a consensus analysis of 767 multiplex families. *Am J Hum Genet* 2001;**69**(4):820-30.
- Pajukanta P, Allayee H, Krass KL, Kuraishy A, Soro A, Lilja HE, Mar R, Taskinen MR, Nuotio I, Laakso M, Rotter JL, de Bruin TW, Cantor RM, Lusa AJ, Peltonen L. Combined analysis of genome scans of Dutch and Finnish families reveals a susceptibility locus for high-density lipoprotein cholesterol on chromosome 16q. *Am J Hum Genet* 2003;**72**(4):903-17.
- Carmichael CM, McGue M. A cross-sectional examination of height, weight, and body mass index in adult twins. *J Gerontol A Biol Sci Med Sci* 1995;**50**(4):B237-44.
- Silventoinen K, Sammalisto S, Perola M, Boomsma DI, Cornes BK, Davis C, Dunkel L, De Lange M, Harris JR, Hjelmberg JV, Luciano M, Martin NG, Mortensen J, Nistico L, Pedersen NL, Skytte A, Spector TD, Stazi MA, Willemsen G, Kaprio J. Heritability of adult body height: a comparative study of twin cohorts in eight countries. *Twin Res* 2003;**6**(5):399-408.
- Parker A, Meyer J, Lewitzky S, Rennich JS, Chan G, Thomas JD, Orho-Melander M, Lehtovirta M, Forsblom C, Hyrkko A, Carlsson M, Lindgren C, Groop LC. A gene conferring susceptibility to type 2 diabetes in conjunction with obesity is located on chromosome 18p11. *Diabetes* 2001;**50**(3):675-80.
- Lindgren CM, Mahtani MM, Widen E, McCarthy ML, Daly MJ, Kirby A, Reeve MP, Kruglyak L, Parker A, Meyer J, Almgren P, Lehto M, Kanninen T, Tuomi T, Groop LC, Lander ES. Genomewide search for type 2 diabetes mellitus susceptibility loci in Finnish families: the Botnia study. *Am J Hum Genet* 2002;**70**(2):509-16.
- Pajukanta P, Nuotio I, Terwilliger JD, Porkka KV, Ylitalo K, Pihlajamäki J, Suomalainen AJ, Syvanen AC, Lehtimäki T, Viikari JS, Laakso M, Taskinen MR, Ehnholm C, Peltonen L. Linkage of familial combined hyperlipidaemia to chromosome 1q21-q23. *Nat Genet* 1998;**18**(4):369-73.
- Pajukanta P, Terwilliger JD, Perola M, Hiekkalinna T, Nuotio I, Ellonen P, Parkkonen M, Hartiala J, Ylitalo K, Pihlajamäki J, Porkka K, Laakso M, Viikari J, Ehnholm C, Taskinen MR, Peltonen L. Genomewide scan for familial combined hyperlipidemia genes in Finnish families, suggesting multiple susceptibility loci influencing triglyceride, cholesterol, and apolipoprotein B levels. *Am J Hum Genet* 1999;**64**(5):1453-63.
- Lilja HE, Soro A, Ylitalo K, Nuotio I, Viikari JS, Salomaa V, Vartiainen E, Taskinen MR, Peltonen L, Pajukanta P. A candidate gene study in low HDL-cholesterol families provides evidence for the involvement of the APOA2 gene and the APOA1C3A4 gene cluster. *Atherosclerosis* 2002;**164**(1):103-11.
- Soro A, Pajukanta P, Lilja HE, Ylitalo K, Hiekkalinna T, Perola M, Cantor RM, Viikari JS, Taskinen MR, Peltonen L. Genome scans provide evidence for low-HDL-C loci on chromosomes 8q23, 16q24.1-24.2, and 20q13.11 in Finnish families. *Am J Hum Genet* 2002;**70**(5):1333-40.
- Silventoinen K, Kaprio J, Lahelma E. Genetic and environmental contributions to the association between body height and educational attainment: a study of adult Finnish twins. *Behav Genet* 2000;**30**(6):477-85.
- Mahtani MM, Widen E, Lehto M, Thomas J, McCarthy M, Brayer J, Bryant B, Chan G, Daly M, Forsblom C, Kanninen T, Kirby A, Kruglyak L, Munnely K, Parkkonen M, Reeve-Daly MP, Weaver A, Bretin T, Dyk G, Lander ES, Groop LC. Mapping of a gene for type 2 diabetes associated with an insulin secretion defect by a genome scan in Finnish families. *Nat Genet* 1996;**14**(1):90-4.
- Sheffield VC, Weber JL, Buetow KH, Murray JC, Even DA, Wiles K, Gastier JM, Pulido JC, Yandava C, Sundén SL, et al. A collection of tri- and tetranucleotide repeat markers used to generate high quality, high resolution human genome-wide linkage maps. *Hum Mol Genet* 1995;**4**(10):1837-44.
- Kong A, Gudbjartsson DF, Sainz J, Jonsson GM, Gudjonsson SA, Richardson B, Sigurdardottir S, Barnard J, Hallbeck B, Masson G, Shlien A, Palsson ST, Frigge ML, Thorgeirsson TE, Gulcher JR, Stefansson K. A high-resolution recombination map of the human genome. *Nat Genet* 2002;**31**(3):241-7.
- Abecasis GR, Cherny SS, Cookson WO, Cardon LR. GRR: graphical representation of relationship errors. *Bioinformatics* 2001;**17**(8):742-3.
- O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 1998;**63**(1):259-66.
- Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002;**30**(1):97-101.
- Hiekkalinna T, Terwilliger JD, Sammalisto S, Peltonen L, Perola M. AUTOGEN: powerful tools for automated genome-wide linkage and linkage disequilibrium analysis. *Twin Res Hum Genet* 2005;**8**(1):16-21.
- Agresti A, Coull BA. Approximate is better than 'exact' for interval estimation of binomial proportions. *Am Statistician* 1998;**52**(2):119-26.
- Brown LD, Cai TT, DasGupta A. Interval estimation for a binomial proportion. *Statist Sci* 2001;**16**(2):101-13.
- Wilson EB. Probable inference, the law of succession, and statistical inference. *J Am Statist Assoc* 1927;**11**:241-47.
- Thompson DB, Ossowski V, Janssen RC, Knowler WC, Bogardus C. Linkage between stature and a region on chromosome 20 and analysis of a candidate gene, bone morphogenetic protein 2. *Am J Med Genet* 1995;**59**(4):495-500.
- Hirschhorn JN, Lindgren CM, Daly MJ, Kirby A, Schaffner SF, Burt NP, Altshuler D, Parker A, Rioux JD, Platko J, Gaudet D, Hudson TJ, Groop LC, Lander ES. Genomewide linkage analysis of stature in multiple populations reveals several regions with evidence of linkage to adult height. *Am J Hum Genet* 2001;**69**(1):106-16.
- Perola M, Ohman M, Hiekkalinna T, Leppavuori J, Pajukanta P, Wessman M, Koskenvuo M, Palotie A, Lange K, Kaprio J, Peltonen L. Quantitative-trait-locus analysis of body-mass index and of stature, by combined analysis of genome scans of five Finnish study groups. *Am J Hum Genet* 2001;**69**(1):117-23.
- Deng HW, Xu FH, Liu YZ, Shen H, Deng H, Huang QY, Liu YJ, Conway T, Li JL, Davies KM, Recker RR. A whole-genome linkage scan suggests several genomic regions potentially containing QTLs underlying the variation of stature. *Am J Med Genet* 2002;**113**(1):29-39.
- Willshire S, Frayling TM, Hattersley AT, Hitman GA, Walker M, Levy JC, O'Rahilly S, Groves CJ, Menzel S, Cardon LR, McCarthy ML. Evidence for linkage of stature to chromosome 3p26 in a large U.K. family data set ascertained for type 2 diabetes. *Am J Hum Genet* 2002;**70**(2):543-6.
- Xu J, Blecker ER, Jongepier H, Howard TD, Koppelman GH, Postma DS, Meyers DA. Major recessive gene(s) with considerable residual polygenic effect regulating adult height: confirmation of genomewide scan results for chromosomes 6, 9, and 12. *Am J Hum Genet* 2002;**71**(3):646-50.
- Beck SR, Brown WM, Williams AH, Pierce J, Rich SS, Langefeld CD. Age-stratified QTL genome scan analyses for anthropometric measures. *BMC Genet* 2003;**4**(suppl 1):S31.
- Mukhopadhyay N, Finegold DN, Larson MG, Cupples LA, Myers RH, Weeks DE. A genome-wide scan for loci affecting normal adult height in the Framingham Heart Study. *Hum Hered* 2003;**55**(4):191-201.
- Allison DB, Heo M. Meta-analysis of linkage data under worst-case conditions: a demonstration using the human OB region. *Genetics* 1998;**148**(2):859-65.
- Guerra R, Etzel CJ, Goldstein DR, Sain SR. Meta-analysis by combining p-values: simulated linkage studies. *Genet Epidemiol* 1999;**17**(suppl 1):S605-9.
- Halpern J, Whittemore AS. Multipoint linkage analysis. A cautionary note. *Hum Hered* 1999;**49**(4):194-6.
- Daw EW, Thompson EA, Wijsman EM. Bias in multipoint linkage analysis arising from map misspecification. *Genet Epidemiol* 2000;**19**(4):366-80.
- Goring HH, Terwilliger JD. Linkage analysis in the presence of errors III: marker loci and their map as nuisance parameters. *Am J Hum Genet* 2000;**66**(4):1298-309.
- Nievergelt CM, Smith DW, Kohlenberg JB, Schork NJ. Large-scale integration of human genetic and physical maps. *Genome Res* 2004;**14**(6):1199-205.
- Wu X, Cooper RS, Borecki I, Hanis C, Bray M, Lewis CE, Zhu X, Kan D, Luke A, Curb D. A combined analysis of genomewide linkage scans for body mass

- index from the National Heart, Lung, and Blood Institute Family Blood Pressure Program. *Am J Hum Genet* 2002;**70**(5):1247–56.
- 41 **Willemsen G**, Boomsma DI, Beem AL, Vink JM, Slagboom PE, Posthuma D. QTLs for height: results of a full genome scan in Dutch sibling pairs. *Eur J Hum Genet* 2004;**12**(10):820–8.
 - 42 **Roberts SB**, MacLean CJ, Neale MC, Eaves LJ, Kendler KS. Replication of linkage studies of complex traits: an examination of variation in location estimates. *Am J Hum Genet* 1999;**65**(3):876–84.
 - 43 **Wuyts W**, Van Hul W. Characterization and genomic localization of the mouse *Extl2* gene. *Cytogenet Cell Genet* 2000;**89**(3–4):185–8.
 - 44 **Wuyts W**, Van Hul W, Hendrickx J, Speleman F, Wauters J, De Boule K, Van Roy N, Van Agtmael T, Bossuyt P, Willems PJ. Identification and characterization of a novel member of the EXT gene family, *EXTL2*. *Eur J Hum Genet* 1997;**5**(6):382–9.
 - 45 **Wuyts W**, Van Hul W, Wauters J, Nemtsova M, Reyniers E, Van Hul EV, De Boule K, de Vries BB, Hendrickx J, Herrygers I, Bossuyt P, Balemans W, Franssen E, Vits L, Coucke P, Nowak NJ, Shows TB, Mallet L, van den Ouweland AM, McGaughan J, Halley DJ, Willems PJ. Positional cloning of a gene involved in hereditary multiple exostoses. *Hum Mol Genet* 1996;**5**(10):1547–57.
 - 46 **Ahn J**, Ludecke HJ, Lindow S, Horton WA, Lee B, Wagner MJ, Horsthemke B, Wells DE. Cloning of the putative tumour suppressor gene for hereditary multiple exostoses (*EXT1*). *Nat Genet* 1995;**11**(2):137–43.
 - 47 **Wicklund CL**, Pauli RM, Johnston D, Hecht JT. Natural history study of hereditary multiple exostoses. *Am J Med Genet* 1995;**55**(1):43–6.
 - 48 **Li Y**, Lacerda DA, Warman ML, Beier DR, Yoshioka H, Ninomiya Y, Oxford JT, Morris NP, Andrikopoulos K, Ramirez F, et al. A fibrillar collagen gene, *Col11a1*, is essential for skeletal morphogenesis. *Cell* 1995;**80**(3):423–30.
 - 49 **Griffith AJ**, Sprunger LK, Sirko-Osada DA, Tiller GE, Meisler MH, Warman ML. Marshall syndrome associated with a splicing defect at the *COL11A1* locus. *Am J Hum Genet* 1998;**62**(4):816–23.
 - 50 **Benz K**, Breit S, Lukoschek M, Mau H, Richter W. Molecular analysis of expansion, differentiation, and growth factor treatment of human chondrocytes identifies differentiation markers and growth-related genes. *Biochem Biophys Res Commun* 2002;**293**(1):284–92.
 - 51 **Matys V**, Fricke E, Geffers R, Gossling E, Haubrock M, Hehl R, Hornischer K, Karas D, Kel AE, Kel-Margoulis OV, Kloos DU, Land S, Lewicki-Potapov B, Michael H, Munch R, Reuter I, Rotert S, Saxel H, Scheer M, Thiele S, Wingender E. TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res* 2003;**31**(1):374–8.
 - 52 **Sandelin A**, Alkema W, Engstrom P, Wasserman WW, Lenhard B. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res* 2004;**32**(Database issue):D91–4.
 - 53 **Berezikov E**, Guryev V, Plasterk RH, Cuppen E. CONREAL: conserved regulatory elements anchored alignment algorithm for identification of transcription factor binding sites by phylogenetic footprinting. *Genome Res* 2004;**14**(1):170–8.
 - 54 **Brudno M**, Do CB, Cooper GM, Kim MF, Davydov E, Green ED, Sidow A, Batzoglu S. LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. *Genome Res* 2003;**13**(4):721–31.
 - 55 **Lenhard B**, Sandelin A, Mendoza L, Engstrom P, Jareborg N, Wasserman WW. Identification of conserved regulatory elements by comparative genome analysis. *J Biol* 2003;**2**(2):13.
 - 56 **Loots GG**, Ovcharenko I. rVISTA 2.0: evolutionary analysis of transcription factor binding sites. *Nucleic Acids Res* 2004;**32**(Web Server issue):W217–21.
 - 57 **Sandelin A**, Wasserman WW. Prediction of nuclear hormone receptor response elements. *Mol Endocrinol* 2004;**19**(3):595–606.
 - 58 **Brook CG**, Murset G, Zachmann M, Prader A. Growth in children with 45,XO Turner's syndrome. *Arch Dis Child* 1974;**49**(10):789–95.
 - 59 **Spuhler JN**. Assortative mating with respect to physical characteristics. *Soc Biol* 1982;**29**(1–2):53–66.
 - 60 **Silventoinen K**, Kaprio J, Lahelma E, Viken RJ, Rose RJ. Assortative mating by body height and BMI: Finnish twins and their spouses. *Am J Human Biol* 2003;**15**(5):620–7.