

SHORT REPORT

Combined analysis of genomewide scans for adult height: results from the NHLBI Family Blood Pressure Program

Xiaodong Wu¹, Richard S. Cooper¹, Eric Boerwinkle², Stephen T. Turner³, Steve Hunt⁴, Richard Myers⁵, Richard A Olshen⁶, David Curb⁷, Xiaofeng Zhu¹, Donghui Kan¹ and Amy Luke¹

¹Department of Preventive Medicine and Epidemiology, Loyola University Medical Center, Maywood, IL, USA; ²Institute for Molecular Medicine and Human Genetics Center, University of Texas Houston Health Science Center, Houston, TX, USA; ³Division of Hypertension, Mayo Clinic, Rochester, MN, USA; ⁴Cardiovascular Genetics Research Clinic, University of Utah School of Medicine, Salt Lake City, UT, USA; ⁵Section of Neurogenetics, Boston University Medical Center, Boston, MA, USA; ⁶Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA, USA; ⁷Pacific Health Research Institute, Honolulu, HI, USA

A combined analysis of genome scans was performed for adult height in the NHLBI Family Blood Pressure Program. Height data were available on 6752 individuals. Linkage analysis was performed first separately for each of the eight ethnic groups in the four networks using the variance component method. To increase the power to detect the common genetic components affecting height for all the individuals, a linkage analysis was performed subsequently for the combined data set by pooling the average allele-sharing IBD (\bar{I}) for all groups. By combining the data, we replicated evidence for a QTL influencing adult height on chromosome 7 (7q31) (LOD = 2.46), which has been reported in two previous studies. Suggestive linkage (LOD > 1) was found in another six regions in our combined analysis. Evidence for linkage for two of these regions (2p12, 20p11) has also been reported previously.

European Journal of Human Genetics (2003) 11, 271–274. doi:10.1038/sj.ejhg.5200952

Keywords: height; BMI; genome scan; linkage

Introduction

Adult height is a complex genetic trait with high heritability (>80%).^{1–5} Genetic studies of adult height have revealed that it is normally distributed, suggesting that height could be affected by the interaction of multiple genetic and environmental factors. Compared with other complex traits, adult height can be measured easily and accurately, reducing error variance in the phenotype. In

most genetic studies for other complex human traits, height has been collected along with other data. Obviously, this raises the possibility of pooling data. Several recent genome scans for adult height revealed significant and suggestive evidence for linkage on several chromosomes.^{4–7} However, few of these regions were replicated among different studies. This may not be surprising considering that height is a complex genetic trait affected by multiple genes, each having a small effect, as well as by environmental factors. An obvious step to increase power and decrease false-positive results in linkage analysis of complex traits is to pool data across studies, so long as one is careful regarding the heterogeneity among different data sets.^{8,9} Here we report the results of our combined analysis of genome scans for adult height in the ongoing NHLBI

*Correspondence: Dr X Wu, Department of Preventive Medicine and Epidemiology, Loyola University Medical Center, 2160 S. First Ave, Maywood, IL 60153, USA. Tel: +1 708 327 9070; Fax: +1 708 327 9009; E-mail: xwu@apache.medctr.luhs.org
Received 14 May 2002; revised 22 October 2002; accepted 19 December 2002

Family Blood Pressure Program (FBPP). We first performed the genome scan for adult height within each ethnic group of the four networks in FBPP separately. We then performed a genome scan using information from all the groups by combining IBD sharing estimated from each group separately.

Materials and methods

Subjects were selected from the four component networks (GenNet, GENOA, HyperGEN, and SAPPHiRe) of the FBPP. All the families were ascertained by hypertensive probands. The ascertainment scheme and program design have been reported elsewhere.¹⁰ In brief, GenNet sampled African-American and European-American nuclear families through identification of a young-middle-aged proband with elevated blood pressure (BP). Both GENOA and HyperGEN sampled African-American and European-American sibships containing sibpairs with essential hypertension. GENOA also sampled Mexican-American sibships containing sibpairs with hypertension, as well as some non hypertensive sibs in all three ethnic groups. SAPPHiRe recruited groups of Japanese and Chinese sibpairs that were concordant for hypertension or extremely discordant. In each network, we combined the data according to ethnic background, including European Americans and African Americans from GenNet; European Americans, African Americans and Mexican Americans from GENOA; European Americans and African Americans from HyperGEN; and Asians from SAPPHiRe.

DNA was extracted from whole blood by standard methods at each of the four networks and was sent to the Mammalian Genotyping Service in Marshfield, WI for analysis. Screening Set 8 (372 highly polymorphic microsatellite markers) was used for all four networks. This screening set has an average heterozygosity of 80%, an average inter-marker distance of 10 cm, and covers 95% of the human genome. Some aspects of the marker data upon which this paper is based are under review by members of the FBPP group. However, no edits of the data contem-

plated thus far have any bearing on the inferences of this paper.

Individuals younger than 20 years old were excluded from the analysis because growth may not have reached a maximum. Height was adjusted by age for each gender separately in each of the eight groups and the residuals were standardized to obtain *Z* scores. These were the phenotypes in our analyses. Heritability for height was estimated for each group separately using the variance component method. The genome scans were performed using the multipoint variance component method in SOLAR.¹¹ The VC method specifies the expected genetic covariances between relatives as a function of the estimated proportion of alleles sharing IBD at a marker locus. The IBD probabilities were estimated using a multipoint approach that considers all available genotypes. The likelihood ratio test was applied to test the null hypothesis of no additive genetic variance due to a quantitative trait locus (QTL) at a particular location. To obtain the empiric genomewide significance of this result, we performed simulations by randomly generating genotypes while retaining the phenotypes and pedigrees under the hypothesis of no linkage. Maximum LOD score (MLS) from each simulation was recorded and the *P*-value for our genome scan was obtained by calculating the frequency of MLS larger than the tested LOD score.

It is well known that allele frequencies between different human populations can be quite different. The estimates of allele frequencies could be biased if we used the combined data set directly; in turn, this could bias the estimates of IBD. To get an unbiased estimate of IBD, we first estimated multiple IBD sharing within each of eight groups separately using SOLAR and then combined all the IBD sharing for further analysis using the variance components method implemented in SOLAR for the combined data.

Results

Data on height were available from 6752 individuals in 2508 families from the FBPP study. The characteristics of the family members for each group are presented in Table 1.

Table 1 Characteristics of families

Population	% of male	Number of height	Age (years)	Height (cm)	
				Female	Male
GenNet European American	47.2	598	45.0 (14.0)	163.2 (6.3)	178.1 (7.1)
GenNet African American	40.3	595	40.5 (11.6)	164.1 (7.2)	176.8 (7.1)
GENOA European American	47.5	749	56.3 (9.8)	162.4 (6.0)	176.7 (6.6)
GENOA African American	34.5	611	56.5 (9.4)	164.8 (6.2)	177.7 (6.3)
GENOA Mexican American	38.8	778	55.8 (11.6)	157.9 (5.6)	170.6 (7.0)
HyperGEN European American	47.5	1100	60.9 (8.7)	161.2 (6.2)	174.9 (6.3)
HyperGEN African American	30.8	1252	50.7 (11.0)	162.5 (6.1)	175.6 (6.8)
SAPPHiRe Asian	45.3	1069	54.0 (12.2)	155.9 (5.2)	167.5 (6.0)
Combined	41.4	6752	53.6 (12.3)	161.3 (6.7)	174.5 (7.5)

Table 2 Results of linkage analysis of individual genome scans

Chromosome	Marker	Location	Region	LOD	Study
1	D1S1631	136	115–143	2.25	GENOA African American
3	D3S1763	175	160–181	2.06	GENOA European American
5	D5S816	137	127–178	2.26	HyperGen European American
	D5S816	134	123–148	1.48	SAPPHIRE Asian
6	D6S1053	78	62–94	2.66	GENOA European American
14	D14S592	67	58–92	3.67	GENOA European American
	D14S592	67	48–96	1.60	SAPPHIRE Asian

The height for females (161.3 cm) and males (174.5) was significantly different ($P < 0.0001$) from each other in the combined sample. The number of families in each group is from 605 to 1410 and the total number of families is 2508 families in the combined data. The mean sibship size in each group is from 1.7 to 3.6, and the mean sibship size is 2.8 in the combined data. The heritability of height for the eight groups varied from 0.75 to 0.98, which is compatible with previous reports.^{4–6} For individual genome scans, the region showing the strongest evidence for linkage was found in GENOA European Americans, centered at 14q21.1 (MLS = 3.67 at marker D14S592) (Table 2). The same region also showed evidence of linkage in SAPPHIRE Asians with a LOD = 1.60. To obtain the genomewide empirical P -value, we performed 250 simulations. We obtained $MLS > 3.67$ 19 times, giving an empirical genomewide P -value of 0.08. Four other regions with a $LOD > 2.0$ (nominal $P = 0.0012$) were found in individual genome scans. These regions are at 1p11, 3q25, 5q23, and 6q13 (Table 2).

The genome scan result for the combined IBD analysis is presented in Figure 1. Seven regions with $LOD > 1$ (nominal $P = 0.016$) are also listed in Table 3 as demonstrating suggestive evidence for linkage. The strongest support for linkage was found at 7q36.1 region (174 cm, marker D7S3058) with multipoint $LOD = 2.46$.

Discussion

Hirschhorn *et al*⁴ and Perola *et al*⁵ reported strong evidence for linkage on chromosome 7 (150 cm, marker D7S2195, $LOD = 3.40$; and marker 164 cm, D7S1826, $LOD = 2.91$). Encouragingly, the most significant result in our combined analysis was also found around the same region (174 cm, marker D7S3058, $LOD = 2.46$), which makes it likely a true positive result.

Another two regions showing suggestive evidence in our study have been reported previously. Our chromosome 2 region (2p12, 92 cm, marker D2S1394, $LOD = 1.27$) has been reported by Hirschhorn *et al*⁴ (104 cm, marker D2S113, $LOD = 2.23$). Thompson *et al*¹² have reported linkage on chromosome 20 (20p11, 34.22 cm, marker D20S66,

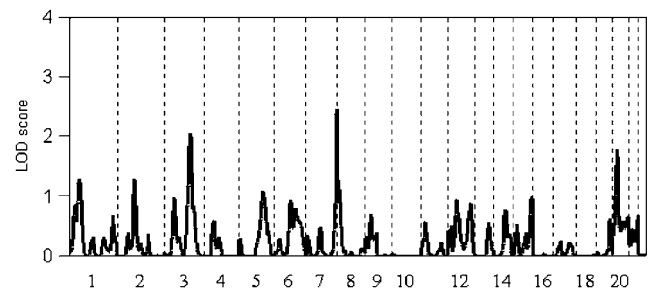


Figure 1 Genome-scan results for combined linkage analysis. The X-axis is the chromosome location and the Y-axis is the LOD score. Each chromosome is scaled according to its length from the genetic map.

Table 3 Evidence of linkage analysis in combined samples

Chromosome	Marker	Location	Region	LOD
1	D1S1622	56	48–63	1.27
2	D2S1394	92	87–99	1.27
3	D3S1764	146	134–158	2.03
5	D5S1505	135	131–138	1.08
7	D7S3058	174	167–182	2.46
8	D8S277	5	1–15	1.26
20	D20S604	29	20–35	1.77

$LOD = 3.0$). Although none of the linkage studies reported previously replicated this finding, we did find suggestive evidence at this region (29 cm, marker D20S604, $LOD = 1.77$).

The strongest finding in all the individual studies was on 14q21.1 (MLS = 3.67) with a genomewide empirical P -value of 0.08, which is not significant at the $\alpha = 0.05$ level. None of the linkage studies reported previously has shown evidence for linkage at this region. More evidence from future studies is needed to clarify this result.

It has been shown by simulation studies that the power to detect a contributing locus with a small effect in a genome-scan study with moderate sample size is very

limited when an acceptable type I error rate is maintained.⁴ In a recent review paper, Altmüller *et al*¹³ compared 31 whole-genome scans for different human complex diseases and found that the most obvious difference influencing success in finding linkage across studies was sample size. Pooling data from multiple studies will greatly increase power assuming the variations among studies are due to random sampling error. Pooling data also will reduce the type I error rates resulting from many independent tests of common hypotheses. By performing linkage analysis in our combined data, we found suggestive evidence for linkage at three regions reported previously for linkage with adult height.

Power calculations for this study have been undertaken under a simplified scenario using a method developed by Sham *et al*¹⁴. For a diallelic QTL with equal allele frequency, we assumed a model with 5% QTL additive variance, 70% residual shared variance and a recombination rate between marker and QTL of 5%. Considering a sample size of 6700 with sibship size of 3 with, the power to achieve LOD = 3 is 97%. If the sibship size decreases to 2, the power will decrease to only 25% in the same sample. If we increase QTL additive variance from 5 to 10% and keep other parameters the same, the power to detect the QTL with LOD = 3 will increase to 97% even with sibship size of 2. Although the average sibship size of the FBPP study is close to 3 (2.8), the variation of pedigree structure among different networks could decrease the power under an ideal situation. Based on the above discussion, it is reasonable to expect a high power to detect a QTL with an effect size of 10% in the combined sample.

No obvious candidate gene for adult height was found for the chromosome 7 region (7q31) in an initial search in NCBI databases. For the chromosome 2 region (2p12), a candidate gene is bone morphogenetic protein 10 (BMP10). This gene is a member of TGF- β family of growth factors. For the region at chromosome 20 (20p11), an obvious candidate gene is BMP2. Like BMP10, BMP2 also belongs to the TGF- β family and induces bone formation. In an association study by Thompson *et al*,¹² no significant association was found with BMP2 in a subset of 20 tallest and 20 shortest individuals selected from a sample of over 500 Pima Indians. However, the power to detect any association is low with such a small sample size.

In summary, by combining the data sets from eight studies in the FBPP, we replicated evidence for a QTL influencing adult height in chromosome 7 (7q31) (LOD = 2.46), which has been reported in two studies previously. Suggestive linkage (LOD > 1) was found in another six regions in our combined analysis. Evidence for linkage for two of these regions (2p12, 20p11) has also been reported previously.

Acknowledgements

We thank Thomas Dyer from Southwest Foundation for Biomedical Research, San Antonio, TX, USA for his helpful suggestions using SOLAR package. We are grateful to two anonymous reviewers whose constructive comments were very helpful in preparing the paper for publication. We thank Christy Chang of Johns Hopkins University for her discussion of data cleaning. This work was supported by grants from the National Heart, Lung and Blood Institute (UO1 HL54485; HL54466; HL65702; HL54481; HL45508; HL47910; HL51021; HL54464; HL54473; HL54496; HL54472; HL54515; HL54495; HL54471; HL54509; 2HHZ598; HL65702; HL54527).

References

- 1 Carmichael CM, McGue MA: Cross-sectional examination of height, weight and body mass index in adult twins. *J Gerontol A: Biol Sci Med Sci* 1995; **50**: B237–B244.
- 2 Preece MA: The genetic contribution to stature. *Horm Res* 1996; **45**: 56–58.
- 3 Silventoinen K, Kaprio J, Lahelma E, Loskenvuo M: Relative effect of genetic and environmental factors on body height: differences across birth cohorts among Finnish men and women. *Am J Public Health* 2000; **90**: 627–630.
- 4 Hirschhorn JN, Lindgren CM, Daly MJ *et al*: Genomewide linkage analysis of stature in multiple populations reveals several regions with evidence of linkage to adult height. *Am J Hum Genet* 2001; **69**: 106–116.
- 5 Perola M, Ohman M, Hiekkalinna T *et al*: 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**: 117–123.
- 6 Wiltshire S, Frayling TM, Hattersley AT *et al*: 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**: 543–546.
- 7 Xu J, Bleecker ER, Jongepier H *et al*: 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**: 646–650.
- 8 Allison DB, Heo M: Meta-analysis of linkage data under worst-case conditions: a demonstration using the human OB region. *Genetics* 1998; **148**: 859–865.
- 9 Palmer LJ, Barnes KC, Burton PR *et al*: Meta-analysis for linkage to asthma and atopy in the chromosome 5q31–33 candidate region. *Hum Mol Genet* 2001; **10**: 891–899.
- 10 The FBPP Investigators. A multi-center genetic study of hypertension. The Family Blood Pressure Program (FBPP). *Hypertension* 2002; **39**: 3–9.
- 11 Almasy L, Blangero J: Multipoint quantitative trait linkage analysis in general pedigrees. *Am J Hum Genet* 1998; **62**: 1198–1211.
- 12 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**: 495–500.
- 13 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**: 936–950.
- 14 Sham PC, Cherny SS, Purcell S, Hewitt JK: Power of linkage versus association analysis of quantitative traits, by use of variance-components models, for sibship data. *Am J Hum Genet* 2000; **66**: 1616–1630.