Association of copy number variation in the AHI1 gene with risk of obesity in the Chinese population

Liansha Huang, Dacai Teng1, Hao Wang2, Guoqing Sheng2 and Tonghua Liu

Department of Science and Technology, Beijing University of Chinese Medicine, 11 Bei San Huan Dong Lu, Chaoyang District, Beijing 100029, People’s Republic of China, 1The Research Center of Neurobiology, Xuzhou Medical College, Xuzhou, People’s Republic of China and 2CAS Key Laboratory of Regenerative Biology, South China Institute for Stem Cell Biology and Regenerative Medicine, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, People’s Republic of China

(Correspondence should be addressed to G Sheng; Email: lisabucm@gmail.com; T Liu; Email: thliu@vip.163.com)

Abstract

Objective: The prevalence of obesity has increased dramatically over the past decade. Gene copy number variants (CNVs) have been recognized as a hereditable source of susceptibility in human complex diseases including obesity. Recent studies have shown that Abelson helper integration site 1 (Ahi1) gene has a significant contribution in the homeostasis regulation in mouse models of obesity. A study was therefore carried out to investigate whether CNVs in AHI1 gene contribute to human obesity.

Subjects and methods: We analyzed samples from 70 Chinese overweight adults and 74 healthy controls for DNA copy number change using the Affymetrix single-nucleotide polymorphism (SNP) 6.0 array. Validation of CNVs of AHI1 was achieved by real-time PCR using the ΔΔCt method.

Results: Copy number gain analysis revealed significant gains (P < 0.0017) of AHI1 gene copy number in 17 of 70 (24.3%) samples but only four of 74 (5.4%) controls overall. Then we studied the frequency distribution of CNVs in AHI1 gene according to body mass index (BMI) grade. Five out of 28 (18.5%) at-risk obese, six out of 26 (23.1%) moderate obese, and six out of 17 (29.4%) severe obese subjects studied showed increased AHI1 gene copy number.

Conclusions: The result suggested that there was a significant linear trend for increasing AHI1 gene copy number frequencies with increasing BMI.

European Journal of Endocrinology 166 727–734

Introduction

Obesity is a major public health and economic burden, but the underlying genetic factors remain largely elusive (1, 2). Copy number variation (CNV) of human DNA segments is an important source of genetic diversity, and increasing evidence indicates that CNV may underlie disease susceptibility (2, 3, 4, 5, 6).

CNVs captured 17.7% of the total detected genetic variation in gene expression by disrupting coding sequences, perturbing long-range gene regulation, or altering gene dosage (7). In the last few years, the role of CNVs in metabolic disorders such as obesity and diabetes has been investigated (8, 9, 10). It has become increasingly clear that, collectively, CNVs significantly contribute to the etiology of these diseases. Despite the large amount of information, the overall association between gene CNVs and the pathophysiological mechanism of obesity remains to be characterized. In addition, unlike Caucasian populations, genetic factors contributing to the risk of obesity are not well studied in Asian populations. In light of this, and the fact that CNV is emerging as a new way to understand human genomic variation, the objective of this study was to identify obesity-associated CNV in a Chinese cohort.

The Abelson helper integration site 1 (AHI1) gene (accession number NM_017651.1) is located at 6q23.3. It covers from 135 605 109–135 818 902 bp. A number of functional motifs have been predicted for the AHI1 gene including signal transduction, RNA processing, transcriptional regulation, cytoskeleton assembly, vesicle trafficking, and cell division, and it has been proposed that AHI1 might be involved in insulin-mediated feeding control (11, 12, 13).

In order to evaluate the significance of AHI1 in obesity, we have analyzed the DNA copy number in normal controls and overweight subjects using quantitative Affymetrix single-nucleotide polymorphism (SNP) 6.0 platform and real-time PCR. Using this approach, we have identified frequent gains in AHI1 in overweight subjects. Further, we have identified that gain in AHI1 copy number is associated with increased body mass index (BMI). The data reported here increase our understanding of the biology of obesity and will allow the development of more targeted therapies for this disease.
Subjects and methods

Experimental subjects

All subjects were unrelated Eastern Han Chinese individuals and were normal and healthy defined by a comprehensive suite of exclusion criteria (14). The exclusion criteria for the study subjects were a history of i) serious residual effects of cerebral vascular disease; ii) type 2 diabetes mellitus, except for easily controlled (defined as adult asymptomatic hyperglycemia controlled by diet or oral agents); iii) chronic renal disease manifested by serum creatinine > 1.9 mg/dl; iv) chronic liver disease or alcoholism; v) significant chronic lung disease; vi) > 6 months of corticosteroid therapy at pharmacological levels; vii) > 6 months of treatment with anticonvulsant therapy; viii) evidence of other metabolic disease, such as hyper- or hyperparathyroidism; and ix) hyperthyroidism. The exclusion criteria were mainly based on nurse-administered medical history self-declaration and/or examination of medical records and were applied most rigorously to potential subjects contacted between the ages of 25 and 50 years. About 7% of the total people screened were excluded from our study sample because they met at least one of the exclusion criteria. Finally, of the 155 volunteers initially screened, 70 Chinese overweight adults and 74 healthy controls that met our stringent selection criteria were selected for the analysis of CNVs using the Affymetrix SNP 6.0 array.

Definitions of at-risk obese, moderate obese, and severe obese in Asian adults were according to WHO criteria, i.e. 23–24.9, 25–29.9, and ≥ 30 kg/m² respectively (15).

The analysis was performed on the Chinese cohort comprising subjects (age 18.5–72.3 years) who were overweight (BMI ≥ 23.0 kg/m²) and normal controls (age 19.2–71.4 years) who had never been overweight (BMI always < 23.0 kg/m²). Blood was obtained in the sitting position after a 12 h overnight fast and glucose and lipids were measured. BMI was calculated as body weight (in kilograms) divided by the square of height (in meters). Weight was measured in light indoor clothing without shoes, using a calibrated balance beam scale, and height was measured using a calibrated stadiometer.

The studies had been approved by the Local Ethics Committee. Each subject gave informed written consent to the study. A summary of our study cohort is provided in Table 1. A second cohort (Supplementary Table 2, (see section on supplementary data given at the end of this article)) of 50 Chinese overweight adults and 44 healthy controls meeting the same criteria was used to validate the SNP array result by performing quantitative real-time PCR (qRT-PCR).

DNA extraction and array hybridization

A Qiagen DNeasy kit was used to extract genomic DNA (gDNA) from blood samples. DNA digestion, labeling, and hybridization were performed according to the manufacturer’s instructions. Briefly, total gDNA (500 ng) is digested with NspI and StyI restriction enzymes and ligated to adaptors that recognize the cohesive 4 bp overhangs. All fragments resulting from restriction enzyme digestion, regardless of size, were substrates for adaptor ligation. A generic primer that recognizes the adaptor sequence was used to amplify adaptor-ligated DNA fragments. PCR conditions were optimized to preferentially amplify fragments in the 200–1100 bp size range. PCR amplification products for each restriction enzyme digest were combined and purified using polystyrene beads. The amplified DNA was then fragmented, labeled, and hybridized using Affymetrix Genome-Wide Human SNP Array 6.0. After washing and staining, the arrays were scanned for data analysis.

Verification of copy numbers by qRT-PCR

To validate CNVs identified by SNP arrays, we performed qRT-PCR for AHI1 using gDNA from 40 controls and 32 overweight subjects. The copy number assays AHI1 (Hs00984704_cn, Hs01171572_cn and Hs00684966_cn) and RNAseP (4403326) were obtained from Applied Biosystems. RNAseP (accession number, NM_006413) gene is always present in two copies and is used in many studies as a reference gene to standardize the amount of template DNA added to the reaction. In our study, the RNAseP gene was co-amplified with the AHI1 gene and served as an internal standard (two copies present in a subject with or without increase in the AHI1 gene). The formula of ‘intermediate’ measurement from manufacturer’s protocol (Applied Biosystems pre-design TaqMan copy number assays kit) and other studies was used to analyze the estimated copy number (16, 17, 18). It is equal to 2 × 2^{∆ΔCt} and represents the transformation of the relative abundance of the gene of interest in a given sample (ΔΔCt) from the ‘Ct’ scale to copy number scale. The detailed procedure is as follows.

Each 20 μl assay contained 10 ng gDNA, 900 nM each of forward and reverse primers for the reference gene (RNAseP), and for the target gene 250 nM each of the VIC dye (reference) and FAM dye (target)-labeled gene-specific probe in 1× TaqMan Gene Expression

| Table 1 Basic phenotypical characteristics of the subjects. |
|-------------------------|---------|---------|
| Traits                  | Normal  | Overweight |
| n                       | 74      | 70       |
| Age (years)             | 41.6±13.5 | 38.5±12.8 |
| Gender (M/F)            | 38/35   | 36/34    |
| Triglycerol (mmol/l)    | 1.16±0.42 | 1.04±0.36 |
| Total cholesterol (mmol/l) | 4.37±0.78 | 4.86±0.68 |
| LDL-cholesterol (mmol/l) | 2.37±0.66 | 2.57±0.97 |
| HDL-cholesterol (mmol/l) | 1.42±0.45 | 1.40±0.38 |
| Fasting glucose (mmol/l) | 4.67±0.58 | 4.87±0.56 |
| Weight (kg)             | 57.6±3.4 | 75.5±5.6 |
| Height (m)              | 1.68±0.81 | 1.67±0.72 |
Master Mix. Individual samples were run in triplicate. Thermal-cycling conditions were as follows: 2 min at 50 °C, 10 min at 95 °C, followed by 40 cycles of 15 s at 92 °C, and 60 s at 60 °C. Real-time data were collected using the OPTICON Software (Hoofddorp, North Holland, Netherlands). QRT-PCR data were analyzed using the relative quantification or ΔΔCt method (16) based on DNA copy number ratio (R) of a target gene vs reference gene in a given patient sample relative to a matched normal control sample. Each replicate was normalized to RNaseP to obtain a ΔCt (FAM dye Ct – VIC dye Ct) and then an average ΔCt for each sample (from the three replicates) was calculated. All samples were then normalized to an average ratio established from 40 control samples to determine ΔΔCt. Relative quantity (RQ) is 2–ΔΔCt, and copy number is 2× RQ. Relative quantification is easier to perform than absolute quantification because a calibration curve is not necessary. It is based on the copy number expression levels of a target gene vs a housekeeping gene (reference or control gene). RNaseP was selected as the internal reference gene because its copy number changes have rarely been observed. We used 1 as default amplification efficiency in our comparative Ct quantification. This mathematical model for copy number measurement will greatly increase the accuracy and reproducibility.

Data analyses

The scanned images of SNP arrays were analyzed using the dChip Software (19). The dChip program allows the copy number analysis against user-defined reference or match-paired samples. Data were normalized to a baseline array with a median signal intensity by the invariant set normalization method. The model-based (PM/MM) method was used to obtain the signal values after normalization in dChip. To infer copy numbers of AHI1 gene, we used the normal sample as the reference and the median smoothing method with a window size of 5. A signal value is computed for each SNP (Supplementary Table 1, see section on supplementary data given at the end of this article) in each sample and is analogous to the expression value in expression array analysis. Based on these signal values, the raw copy number for an SNP in a sample is computed as (signal/(mean signal of normal samples at this SNP)×2), and log2 ratio is computed as log2 (signal/mean signal of normal samples at this SNP). The raw copy numbers and log2 ratio data can be viewed in the chromosome view.

Statistical analyses

GraphPad Prism 5.0 Software (GraphPad Inc., La Jolla, CA, USA) was used for statistical analysis. The nonparametric Mann–Whitney U test and χ2 test were used to compare the distribution of the AHI1 copy number between overweight and controls. P value < 0.05 were considered statistically significant.

Results

CNV represents a form of structural genomic variation in which the number of copies of a DNA segment, ranging from 1 kb to several Mb, varies in different individuals when compared with a reference genome (20). We calculated quality control measures on our Affymetrix SNP 6.0 genome-wide association study (GWAS) data based on statistical distributions to exclude poor-quality DNA samples and false-positive CNVs. The first threshold is the percentage of attempted SNPs that were successfully genotyped. Only samples with call rate > 98% were included. The genome-wide intensity signal must have as little noise as possible. Only samples with the s.d. of normalized intensity (log R ratio (LRR)) < 0.25 were included.

In an effort to identify novel copy number alterations that drive the pathogenesis of obesity, we initially investigated gDNA derived from 70 overweight subjects and 74 healthy controls using Affymetrix SNP 6.0 microarrays, as described in Subjects and methods section. Genomic copy numbers for more than 900 000 probes were determined by calculating the median signal intensities of the healthy controls and overweight subjects with respect to normal reference DNAs. The result displayed as copy number heatmap in Fig. 1a indicated increased copy number at the AHI1 encompassing segment at the 6q23.2 chromosomal locus spanning ~210 kb. Also, the mean number of AHI1 copies estimated by the SNP intensity was significantly increased (20%) in the overweight group compared with controls (P = 0.0002, Fig. 1b).

As shown in Table 2, the frequency of increased copy number (CN > 2) of AHI1 gene was significantly higher (P = 0.0017) in the overweight group (17/70, 24.3%) than in the healthy controls (4/74, 5.4%). However, there is no significant difference in frequency of decreased copy number (CN < 2) between overweight subjects (3/70, 4.3%) and healthy controls (8/74, 10.8%; Table 2).

Therefore, the result from microarrays suggested that the mean and distribution of AHI1 copies in the overweight subjects were significantly increased in the control group.

To validate the CNVs identified by SNP array, we performed qPCR analysis for the AHI1 gene using gDNA from 30 overweight subjects and 42 normal controls. Quantification of copy numbers for the AHI1 gene using qPCR coincided with the copy number derived from the Affymetrix SNP array analysis. The average copy number for the AHI1 gene was also increased by 20% in the group of overweight subjects, which is in complete agreement with the SNP data (P < 0.0001, Fig. 2). Frequency distribution of AHI1 copy number gain analysis showed >2 copies in ten of 30 (33.3%) overweight subjects (median three copies, range 2–5) compared with only five of 42 (11.9%) controls (median two copies, range 2–4). Thus, evidence from qPCR
confirms the association of \textit{AHI1} copy number gains with increased risk of obesity ($P=0.0394$).

Then we studied the frequency distribution of CNVs in the \textit{AHI1} gene in obese samples. All the overweight samples were classified into three groups according to common criteria of overweight severity: at-risk obese (BMI $<25$ kg/m$^2$, $n=27$), moderate obese (BMI $<30$ kg/m$^2$, $n=26$), and severe obese (BMI $\geq 30$ kg/m$^2$, $n=17$). Five out of 28 (18.5%) at-risk obese, six out of 26 (26.9%) moderate obese, and six out of 17 (29.4%) severe obese subjects studied showed increased \textit{AHI1} gene copy number (Table 3). Thus, the highest incidence of copy number increase in \textit{AHI1} was observed in the highest BMI level. According to the linear regression analysis, a robust correction was observed for increasing copy number frequencies with increasing BMI ($r=0.952$, $P<0.0002$).

**Discussion**

The obesity epidemic is an enormous global problem. Obesity is increasing at an alarming rate worldwide and is an important risk factor for type 2 diabetes, cardiovascular diseases, hypertension, and other chronic diseases.

GWASs led to the identification of various common SNPs for human obesity (21, 22, 23, 24, 25, 26). In particular, a robust association between the fat mass and obesity-associated gene (\textit{FTO}) and BMI has been observed in a recent GWAS (27).

In addition, CNV studies also found rare variants of extreme obesity in adults and children (10, 28). Currently, two common CNV regions have been described for BMI and obesity. At first, association of a common deletion near the neuronal growth regulator 1 (\textit{NEGR1}) gene with BMI was reported (26). Some CNVs are in strong linkage disequilibrium with adjacent SNPs; therefore, this CNV was detected in a meta-analysis of SNP data of 15 GWASs comprising more than 32,000 individuals. Recently, another common CNV (chromosome 10p11.22) was shown to be associated with BMI in a Chinese sample (nominal $P$ value $=0.011$) (29).

Recent development in the array technology enables the detection of small DNA copy number changes throughout the entire genome, which may mark the locus of putative obesity genes. SNP arrays have many advantages over more conventional methods of genome feature analysis in terms of efficiency, precision, and minimal DNA requirements and may well become the dominant technology for measuring genome-wide copy

### Table 2: Association of \textit{AHI1} copy number with the development of overweight.

<table>
<thead>
<tr>
<th>Copy number</th>
<th>Overweight ($n/N$, %)</th>
<th>Control ($n/N$, %)</th>
<th>OR (95% CI)</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$&lt;2$</td>
<td>3/70 (4.3)</td>
<td>8/74 (10.8)</td>
<td>2.71 (0.69–10.65)</td>
<td>0.2102</td>
</tr>
<tr>
<td>$&gt;2$</td>
<td>17/70 (24.3)</td>
<td>4/74 (5.4)</td>
<td>0.18 (0.06–0.56)</td>
<td>0.0017</td>
</tr>
</tbody>
</table>

$P$ value was calculated using the Fisher’s exact test.
number. Although the use of high-density SNP arrays to provide clinically relevant information has great appeal, this strategy must be first validated using newer generation SNP arrays containing probes for more SNP markers with higher informative rates. With this technique, recurrent as well as novel aberrations can be readily identified with high resolution. Therefore, CNV of AHI1 gene could be discovered although a relatively small sample size was used in our study.

By performing Affymetrix SNP 6.0 microarray platform, we found that the distribution of copy number of AHI1 differed significantly between controls and overweight subjects. The qPCR validation also showed a significant genetic gain of AHI1 in overweight cases, which was a similar frequency to those from SNP arrays. These findings suggested that CNV of AHI1 might be a risk factor for the development of obesity.

Interestingly, we found that the frequency of DNA copy number gains was more frequent in advanced obesity, which suggested that CNV of AHI1 might play a role in obesity progression and contribute to metabolic consequence.

These findings demonstrate that an increased copy number of AHI1 could be a prevalent genetic change in the majority of obese cases. The discrimination of obese category based on copy number aberrations will lead us to investigate the possible functional associations of AHI1 with obese severity.

AHI1 protein contains several distinct protein domains, which include six WD40 domains, one SH3 domain, potential SH3 binding sites, and an N-terminal coiled–coil domain. Structure–function analysis indicates that AHI1 might interact with other proteins to act as a complex function (30). Many groups have identified that AHI1 mutations are a frequent cause of disease in patients with specific forms of Joubert syndrome (31, 32, 33), an autosomal recessive neurodevelopmental disorder. Distribution of AHI1 in the embryonic and postembryonic mouse tissues has been reported, and both protein and mRNA studies have shown that AHI1 is highly distributed in several brain areas implicated in feeding and metabolic regulation, including the hypothalamus and nucleus tractus solitarius in the brain stem (31, 34, 35, 36).

Nevertheless, AHI1 is a poorly characterized gene and little is known about its potential role in endocrinological diseases. However, two recent reports have demonstrated increased expression of AHI1 protein in the hypothalamus of insulin-treated mice, suggesting a role for AHI1 as a putative central regulator for glucose metabolism and energy homeostasis (13, 36). Also, one group reported a significant association between variants in the AHI1 gene and BMI in a Caucasian population of type 2 diabetes patients (12). These data suggest that alterations in AHI1 directly contribute to the pathogenesis of obesity.

How increased activity of AHI1 possibly collaborates with other network genes to induce obesity and obesity-related metabolic complications remains to be determined.

It is possible that these copy number changes may cause variability in the presence or absence of regulatory elements that influence the expression of other associated genes. For example, it has been found that AHI1-interacting protein, Huntington-associated protein 1, is involved in the central regulation of food intake and body weight (37). Therefore, to understand how these CNVs work, more work including expression studies targeted on AHI1 and its associated genes in the apparently healthy subjects and/or severe obese people are warranted. Similarly, large data sets would also be necessary to investigate the possibility that these CNVs may exert their effects through interaction with other genetic variants. Further analysis of much larger sample sizes of normal and overweight individuals will be required to validate this conclusion.

Understanding the contribution of altered AHI1 activity to obesity will not only provide valuable mechanistic insights but also help to determine whether

**Table 3** Frequency distribution of AHI1 copy number gain in normal, at-risk obese, moderate obese, and severe obese*.

<table>
<thead>
<tr>
<th>BMI grade</th>
<th>Copy number gain (n, %)</th>
<th>Total (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal (18.5–22.9 kg/m²)</td>
<td>4 (5.4)</td>
<td>74</td>
</tr>
<tr>
<td>At-risk obese (23–24.9 kg/m²)</td>
<td>5 (18.5)</td>
<td>27</td>
</tr>
<tr>
<td>Moderate obese (25–29.9 kg/m²)</td>
<td>6 (26.9)</td>
<td>26</td>
</tr>
<tr>
<td>Severe obese (≥30 kg/m²)</td>
<td>6 (29.4)</td>
<td>17</td>
</tr>
</tbody>
</table>

*P = 0.0053.
Supplementary data

This is linked to the online version of the paper at http://dx.doi.org/10.1530/EJE-11-0999.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This work was supported in part by NSFC30700213, 2011DEA30920, U0972001, and 2011CB504004.

References


www.eje-online.org