Published in final edited form as:

Obesity (Silver Spring). 2011 June ; 19(6): 1229–1234. doi:10.1038/oby.2010.323.

Copy Number Variations at the Prader–Willi Syndrome Region on Chromosome 15 and associations with Obesity in Whites

Yuan Chen¹, Yong-Jun Liu², Yu-Fang Pei^{1,2}, Tie-Lin Yang^{1,2}, Fei-Yan Deng², Xiao-Gang Liu¹, Ding-You Li³, and Hong-Wen Deng^{1,2,4,5}

¹The Key Laboratory of Biomedical Information Engineering of Ministry of Education and Institute of Molecular Genetics, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi, The People's Republic of China

²School of Medicine, University of Missouri-Kansas City, Kansas City, Missouri, USA

³Department of Pediatrics, Children's Mercy Hospital, Kansas City, Missouri, USA

⁴Center of System Biomedical Sciences, Shanghai University of Science and Technology, Shanghai, The People's Republic of China

⁵College of Life Sciences and Engineering, Beijing Jiao Tong University, Beijing, The People's Republic of China

Abstract

Obesity is a serious health problem with strong genetic determination. Copy number variation (CNV) is a common type of genomic variant associated with some complex human diseases. However, it is not clear how CNVs contribute to the etiology of obesity. In this study, we examined 1,000 unrelated US whites to search for CNVs that may predispose to obesity. We focused our analyses on the Prader–Willi syndrome (PWS) critical region (chromosome 15q11–q13), because the PWS region is a hotspot for CNV generation and obesity is one of the major clinical manifestations for chromosome abnormalities at this region. We constructed a map containing 39 CNVs at the PWS critical region with CNV occurrence rates higher than 1%. Among them, three CNVs were significantly associated with body fat mass ($P < 0.05$), with a higher copy number (CN) associated with an increase of 5.08–9.77 kg in body fat mass. These three CNVs are close to two known PWS genes, *NDN* (necdin homolog) and *C15orf2* (chromosome 15 open reading frame 2), and partially overlap with another obesity gene *PWRN1* (Prader–Willi region nonprotein-coding RNA 1). Interestingly, our recently published whole genome association scan study using the same sample by examining single-nucleotide polymorphisms (SNPs) did not find any significant associations at these CNV regions, suggesting the importance of examining both CNVs and SNPs for better understanding of genetic basis of obesity. Further studies are warranted to validate these CNVs and their importance to obesity.

Correspondence: Hong-Wen Deng (dengh@umkc.edu).

The first two authors contributed equally to this work.

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at <http://www.nature.com/oby>

DISCLOSURE

The authors declared no conflict of interest.

INTRODUCTION

Obesity is characterized by an excess of body fat due to an imbalance between energy intake and energy expenditure (1). Evidence from twin, family, and adoption studies demonstrate that obesity has strong genetic determination with heritability estimates of body fat mass ranging from 0.4 to 0.7 (2,3). Extensive efforts have been made to uncover the genetic basis of obesity through conventional linkage and candidate gene association studies (4–6), single gene mutation animal models, and Mendelian disorders with obesity as a major clinical feature (e.g., Prader–Willi syndrome (PWS; OMIM 176270)) (7). Recent more powerful genome-wide association studies (GWASs) identified novel genes or single-nucleotide polymorphisms (SNPs) for obesity (8–10). We published a pilot GWAS in 1,000 US whites and found the *CTNBL1* (catenin (cadherin-associated protein), β -like 1) gene is important for obesity (9). Despite the extensive efforts, all the genes/SNPs identified so far, collectively, explain only a small portion of heritability of obesity. This situation suggests that additional genes or different types of genetic variants await for discovery.

Recent studies of structural variation in the human genome demonstrate that copy number variation (CNV), including duplications or deletions, is a common type of genomic variability (11–13). It has also been shown that CNV play an important role in predisposing to human diseases, such as glomerulonephritis (14), HIV-1 AIDS (15), autism (16), and systemic lupus erythematosus (17). CNV is defined as a DNA segment (ranging from 1 kilobase (kb) to several megabases (Mb)) with variable copy number (CN) in comparison with a reference genome (11). CNV may account for about 20% of the total detected genetic variation in gene expression by disrupting coding sequences, perturbing long-range gene regulation, or altering gene dosage (18). Extensive distribution of CNV in the genomes of normal individuals implies that CNV may be a driving force for phenotypic diversity (19–21).

In this study, we aimed to investigate whether CNVs was associated with risk of obesity. Based on our Affymetrix 500K SNP array data in a sample of 1,000 US whites, we performed association analyses at the PWS critical region (15q11–q13), where most PWS genes located (22). This is because obesity is one of the major clinical manifestation of PWS and the PWS critical region is enriched in segmental duplication (23). Previous studies have found that chromosome abnormalities, segmental duplication (24), or deletion (25) at this genomic region can result in obesity. Using our study sample, we constructed a map containing 39 CNVs at the PWS critical region with CNV occurrence rates higher than 1%. Our analyses demonstrated that three CNVs were significantly associated with body fat mass, with a higher CN of these CNVs leading to 5.08–9.77 kg increase in body fat mass.

METHODS AND PROCEDURES

Subjects

The study was approved by the involved institutional review boards. Signed informed consents were obtained from all study participants before they entered the study. Based on our sample pool which had more than 6,000 subjects recruited from the Midwestern United

States, a total of 1,000 unrelated subjects, aged 19.1–87.2 years at the time of survey, were randomly selected for this study. They were whites of European origin and were normal healthy subjects defined by a comprehensive suite of exclusion criteria (4). Briefly, individuals with chronic diseases and conditions that might potentially affect fat metabolism were excluded, mainly based on nurse-administrated medical history self-declaration and/or examination of medical records. These diseases/conditions included chronic disorders involving vital organs (heart, lung, liver, kidney, and brain) and severe endocrinological, metabolic, and nutritional diseases (diabetes, hypoand hyper-parathyroidism, hyperthyroidism, morbid obesity, etc.), chronic use of drugs (anticonvulsant, corticosteroid), recent (within the past year) major gastrointestinal disease, such as peptic ulcer, malabsorption, chronic ulcerative colitis, regional enteritis, or any significant chronic diarrhea state. Body fat mass, body lean mass, and trunk fat mass were measured by Hologic QDR 4500 dual-energy X-ray absorptiometry machines that were calibrated daily, using standard protocols for all the study subjects. Anthropometric measures and a structured questionnaire covering lifestyle, diet, family information, medical history, and others were obtained for all the study subjects. 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 short-term reproducibility (coefficient of variation) of BMI, fat mass, lean mass, and trunk fat mass measurements is on average 0.2%, 1.1%, 1.0%, and 1.0%, respectively. The basic characteristics of the study subjects are summarized in Table 1.

Genomic DNA

Genomic DNA was extracted from whole human blood using a commercial isolation kit (Gentra systems, Minneapolis, MN) following the protocols detailed in the kit. DNA concentration was assessed by a DU530 ultraviolet/visible spectroscopy spectrophotometer (Beckman Coulter, Fullerton, CA). Before advancing to the hybridization step, the measurements of DNA concentration were double-checked by pico-green analysis that can detect fluorescent signal enhancement of PicoGreen dsDNA Quantitation Reagent, which selectively binds to dsDNA (26,27).

Mapping array hybridization

Affymetrix mapping 500K array (Affymetrix, Santa Clara, CA) was used in this study. Array hybridization was performed at the Vanderbilt Microarray Shared Resource at Vanderbilt University Medical Center, Nashville, TN using the standard protocol recommended by the manufacturer. Basically, for each array, 250 ng of genomic DNA was digested with either Nsp1 or Sty1 and ligated to adapters that recognize the cohesive four base pair overhangs. A generic primer that recognizes the ligated adapter sequence was used to amplify the ligation products in a PCR. The amplified DNA was assayed by agarose gel electrophoresis to verify an average size distribution of 250 bp to 1500 bp. The amplified DNA was purified, fragmented, and stained using immunopure streptavidin (Pierce, Milwaukee, WI), biotinylated antistreptavidin antibody (Vector Labs, Burlingame, CA) and R-phycoerythrin streptavidin (Invitrogen, Carlsbad, CA), and hybridized to the array. Fluorescence intensities were quantitated using an Affymetrix array scanner 3000-7G.

Genotyping data management and processing were performed using Affymetrix GeneChip Operating System. A total of 15 subjects (six males and nine females) with genome-wide SNP genotyping call rates less than 93% were excluded from further analysis. The call rate was assessed using the Bayesian Robust Linear Model with Mahalanobis (BRLMM) algorithm (28).

CNV and CNV region determination

Affymetrix GeneChip Chromosome Copy Number Analysis Tool Software ver.4.0.1 was used to detect autosomal CNV with default settings. Briefly, samples were normalized by quantile normalization on perfect match probes (29). Then the 94 starting samples in every sample plate were assigned with equal probabilities into test and reference sets. This randomization process yielded 47 ± 2 subjects in a reference set. A Hidden Markov Model based algorithm was implemented to calculate CN changes based on SNP hybridization signal-intensity data from experimental sample relative to the intensity distributions of the randomization derived reference data set in the same plate. The Hidden Markov Model has five hidden states representing different biological phenomena: homozygous deletion (CN = 0), heterozygous deletion (CN = 1), normal diploid (CN = 2), single copy gain (CN = 3), and amplification (CN = 4). In our analyses, homozygous deletion and heterozygous deletion were defined as “loss”, whereas single copy gain and amplification were defined as “gain”.

The exact boundaries of the CNVs cannot be obtained from the Affymetrix 500K SNP genotyping data. In this study, the boundary of the CNVs were approximated by physical positions of the probe pairs having the maximal distance within a CNV, which yielded relatively shorter CNVs sizes compared to the actual sizes. For association analyses, we used a procedure to redefine the genomic regions with CN change. As illustrated in Figure 1, CNVs were merged to CNV regions, which are stretches of genomic DNA covering overlapping CNVs. A CNV region started with two overlapping CNVs and was bounded by the two SNPs with maximal distance in the two CNVs. The CNV region would be extended when the next overlapping CNV exceeded the previous boundaries. For association analyses, we divided complex CNV regions which comprised individual CNVs with inconsistent boundaries across the entire sample, into several sub-regions to ensure that (i) each CNV subregion has the same boundary among subjects and (ii) each subject will be either gain or loss at each CNV sub-region. CNV sub-regions obtained from the above procedure with CNV frequencies higher than 1% were selected for association analyses. A count of all CNVs occurring in the same CNV sub-region was used to describe the frequency.

Association analysis

After adjusting for the significant covariates of age and sex ($P < 0.05$), the residuals of fat mass of all the subjects followed the normal distribution. *F*-test was performed to test for association between CNV sub-regions and the residuals of fat mass using the software package Helixtree (Golden Helix, Bozeman, MT). The test is based on whether the distributions of the dependent variable within each category are significantly different between the various categories of the predictor variable. The independent variable was CNV sub-region, which was divided into three levels according to its CN (gain: CN >2 ; normal:

CN = 2 and loss: CN < 2). P value < 0.05 was considered to be nominally significant, which was further subject to Bonferroni correction to account for multiple comparisons. Since all the CNV sub-regions were derived from five CNV regions and the sub-regions within each specific CNV region were closely adjacent to each other, we used 0.01 as the study-wise significance level (i.e., $0.05/N$, where $N = 5$).

To better understand the contribution of the identified CNVs to body composition, one sided F -test was performed to assess whether the CNVs were associated with other body composition indexes, such as body lean mass, BMI, and trunk fat mass, in the direction as expected. In addition, we tested their effects in the gender-stratified subgroups. In the total sample, age, and sex were adjusted as covariates. In the gender-stratified analyses, age was adjusted as a covariate. The adjusted phenotypes, if not following normal distribution, were further subject to a Box–Cox transformation to follow a normal distribution.

RESULTS

Characterization of CNVs and CNV regions

At the PWS critical region, a total of 94 distinct CNVs were detected, ranging from 19.8 kb to 1.6 Mb in size. The estimated median and mean sizes of them were 390.2 kb and 491.1 kb, respectively. The 94 distinct CNVs were merged into five CNV regions, which had the median and mean sizes of 563.6 kb and 778.2 kb, respectively. Each CNV region consisted of multiple probes (median 68, mean 65). Consequently, the genome coverage of PWS critical region by CNV regions was approximately 3.9 Mb (27.0%). Finally, a total of 39 CNV subregions were assigned for subsequent association analyses (see Supplementary Table S1 online).

Association analyses

Table 2 shows the information of three CNVs (CNV_031, CNV_034 and CNV_035) achieving nominal P values < 0.05. Their frequencies were 2.84%, 1.22%, and 1.12%, respectively. Among them, CNV_034 ($P = 0.0042$) and CNV_035 ($P = 0.0043$) showed highly significant associations with body fat mass variation. Their significance remained even after corrections for multiple testing based on the number of CNV regions ($P < 0.01$). Nominally significant association was also observed for CNV_031 ($P = 0.0420$). At locus CNV_031, CNV_034, and CNV_035, compared with the normal CN groups, the groups carrying CN gains had 5.08 kg, 9.79 kg, and 9.77 kg higher fat mass, respectively.

Using one-sided F test, we further assessed the effects of CNV_031, CNV_034, and CNV_035 on body lean mass, BMI, and trunk fat mass. Gender-stratified analyses were also performed for body fat mass, body lean mass, BMI, and trunk fat mass (Table 3). In the total sample, all these three CNVs were consistently associated with higher BMI and trunk fat mass at the significance level of 0.1. CNV_031 was significantly associated with higher body fat mass, BMI, and trunk fat mass in men but not in women. In women, significant increase in body fat mass, BMI, and trunk fat mass were observed for CNV_034 and CNV_035. No evidence of association was found for any CNVs with body lean mass in the total sample or the gender-stratified subsamples.

DISCUSSION

To our knowledge, this is the first study reporting CNVs important for common obesity in the general white population. In this study, we constructed a map at the PWS critical region (15q11–q13), which yielded five CNV regions (containing 94 distinct CNVs and 39 commonly occurring CNV subregions with consistent boundaries). We identified significant associations between three CN gains (CNV_031, CNV_034, and CNV_035) and higher body fat mass.

The PWS critical region is one of seven chromosomes in the human genome with a high rate of segmental duplication (30). At this region, we observed five CNV regions, which was consistent with a previous report that CN changes were usually enriched in regions of segmental duplications or low-copy repeats (13). In terms of CNV construction, the five CNV regions were partially matched with previous observations in 506 unrelated healthy individuals from Northern Schleswig-Holstein (Northern Germany) and 270 HapMap individuals (13,31), suggesting the reliability of the constructed CNV regions in our study. We also compared our CNV map with that constructed in PWS patients. Except for sporadic partially overlapped CN loss events, most PWS-causing deletions did not exist in this general healthy white population (32).

Variants of chromosome 15 aberrations present at the PWS critical region have been implicated in human disorders, such as morbid obesity (24) and PWS (32,33). PWS, with obesity, muscle hypotonia, hyperphagia, short stature, small hands and feet, behavioral issues, hypogonadism, and characteristic facial appearance as major clinical manifestations, is a rare genetic disorder. It is normally caused by deficiency of several paternally expressed imprinted transcripts within chromosome 15q11–q13, a region includes *NDN* ((necdin homolog (mouse)), *C15orf2* (chromosome 15 open reading frame 2), *SNURF-SNRPN* (*SNRPN* upstream reading frame-small nuclear ribonucleoprotein polypeptide N) and the *SNORD* genes (C/D box small nucleolar RNAs, C/D box snoRNAs) (34–36). However, the special effects of these PWS genes on obesity was not clear. Deletions encompasses noncoding small nucleolar RNAs (including *SNORD116*) were found to cause obesity in humans (35,37). While, neither *SNORD116*-deficient nor *NDN*-deficient mice developed obesity (38,39). The potential contributions of these PWS genes to obesity, especially to variation in body composition in the generally healthy human population, remained to be revealed.

In this study, CN loss at the PWS critical region did not affect body fat mass. We could not exclude the possibility that cis- or trans-acting regulatory effects of these CN loss may be balanced by the contribution of other CNVs such as ungenotyped translocations. The results of earlier studies have been inconsistent regarding the relationship between deficiency of PWS critical region and obesity (36–39). We identified three CN gains, CNV_031, CNV_034, and CNV_035, which contributed to higher body fat mass. They map to chromosome 15q11.2 at a 1Mb region between two well known PWS candidate genes, *NDN* and *C15orf2*. Butting *et al.* (34) studied the pattern of genomic sequence duplication at the 1Mb region between *NDN* and *C15orf2* genes and identified a new *PWRNI* (Prader–Willi region nonprotein-coding RNA 1) gene, which was found to represent an alternative 5' part

of *SNRPN* gene (40). CNV_031 and CNV_035 each map to two 11.3 kb intervals, with contig positions of 1173038...1184380 and 1023825...1035124, respectively (Contig No. NT_026446.13). Through BLAST search of the human genome (NCBI Build 36.2, Reference Sequence) against exon sequences and exonintron boundaries for *PWRNI* gene, we could annotate CNV_031 to a genomic region containing *PWRNI* exons 19 and 20 duplicons, only 77 bp away from an exon 13 duplicon (Contig position for the three exon duplicons: 1023508... 1027452). We annotated CNV_035 to a genomic region containing both *PWRNI* exon 1 and partial *PWRNI* intron 1 (Contig position for exon1 and intron1: 1173083...1187284) (Figure 2). In addition, CNV_034 was mapped between *PWRNI* exon 1 and an exon 1 duplicon.

PWRNI, whose expression in the brain is imprinted, is considered to be an important candidate gene for PWS and common obesity (34,40). Since the genomic position of the three identified CNVs (CNV_031, CNV_034, and CNV_035) are partially overlapped with the *PWRNI* gene, we speculate that the CN changes in *PWRNI* may influence gene function through misalignment and unequal crossing over (23,25,41), and eventually contribute to phenotypic variation in body fat mass. It is also possible that these CN changes may cause variability in the presence or absence of regulatory elements that influence the expression of other nearby genes such as *NDN*, *C15orf2*, *SNRPN*, and *SNORD* genes (18,40). Two (CNV_031 and CNV_035) of the three CNVs are encompassed with several *PWRNI* encoding sequences (exon 1, exon 19, and exon 20), they may also affect alternative splicing which was observed in its EST clones (34). Still, to understand how these CNVs work, more work including expression studies targeted on *PWRNI* and its nearby genes in the apparently healthy subjects and/or severe obese people are warranted.

It is known that paternally and maternally derived duplication of 15q11–q13 may have different clinical outcomes (24,42–45). Patients with maternally transmitted duplications at dup (15) (q11.2–12) may result in developmental delay and those with paternal origin developed obesity (24). Considering this, there may be parent-of-origin effects for the three identified CNVs in this study. However, our study sample of unrelated subjects does not allow further exploration of the origin (inherited from father/mother or be de novo germline mutation) and the inheritance pattern of these CNVs, while studies by using family data can be helpful to address this issue.

In our recently published whole genome association study in the same sample by analyzing SNPs (9), no probes in CNV_031, CNV_034, or CNV_035 showed any significant association with body fat mass. This implies that CNV analyses may provide further insight into the genetic basis of obesity and suggests that integration for both types of variants could be helpful for better understanding of the etiology of obesity. It was notable that the identified three CN gains had gender-specific effects on both trunk and total fat mass accumulation. Significant associations were found only in the male subgroup for CNV_031 and in the female subgroup for CNV_034 and CNV_035. These findings supported the hypothesis regarding the sex-specific genetic architecture of fat mass and the gender-specificity in the association between identified three CN gains and fat mass accumulation. No evidence was found to show the three CNVs would lead to reduced lean mass.

In our study, the genome coverage of PWS critical region was 27%. However, the estimate of CNV coverage of PWS critical region based on the Database of Genomic Variants (<http://projects.tcag.ca/variation/?source=hg18>), a continuously updating catalog of structural variations identified in healthy human populations, was 74.8%. This coverage may be overestimated, as a recent report based on higher-resolution arrays showed that nonCNV genome may be included in this database (46). It was also noted that this database summarized various studies in which different SNP array platforms and/ or technologies were involved. To our knowledge, none of the current available technologies alone (oligonucleotide arrays, microarray based comparative genomic hybridization, etc.) may capture all variation in the human genome.

A limitation of this study is that Affymetrix 500K array is primarily designed to detect SNPs. It does not include specific probes for CNV detection. Therefore, the accurate boundaries, frequencies, and population-genetic properties of some CNVs may not be determined unambiguously (46). With rapid advancement in biotechnology, more powerful platforms are now available for CNV detection, including Affymetrix SNP Array 6.0, Agilent high-density array comparative genomic hybridization array, nucleotide scale CNV detection through fosmid end-sequence pair approach and so on.

Another limitation is that a dysmorphic examination was not performed on the subjects to rule out obesity-related syndromes due to genetics or endocrine disturbances. Instead, we ruled out them mainly based on a nurse-administrated medical history self-declaration. The agreement between declaration and dysmorphic examination was unclear. Therefore, strictly speaking, we could not completely rule out the disturbant factors that may affect fat distribution.

In summary, we identified three CNVs at the targeted PWS region that may be important for common obesity in the general white populations. Future efforts in other independent populations defined by better exclusion criteria and using more powerful platforms for CNV studies are necessary to confirm and validate our findings.

A total of 39 CNV IDs were assigned. The Start and End Physical Position for the approximate boundary probe pairs were annotated based on NCBI Build 36.2 human genome assembly. A length of “-” means the CNV was assigned by only one Probe Set thus the length was undefined. “# Probeset” is the number of probeset in the CNV. “# Loss”, “# Normal” and “# Gain” are the number of individuals carrying the specific CN state. “CNV Frequency” is the percentage of subjects who do not have a normal CNV among the sample.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Megan N. Mayer for helping edit the manuscript. The study was benefited from support from National Natural Science Foundation of China (31000554), Shanghai Leading Academic Discipline Project (Project Number: S30501) and the Ministry of Education of China. H.W.D. was partially supported by grants from NIH

(P50AR055081, R01AG026564, R01AR050496, RC2DE020756, R01AR057049, and R03TW008221) and Franklin D. Dickson/Missouri Endowment.

References

1. Spiegelman BM, Flier JS. Obesity and the regulation of energy balance. *Cell*. 2001; 104:531–543. [PubMed: 11239410]
2. McPherson R. Genetic contributors to obesity. *Can J Cardiol*. 2007; 23 (Suppl A):23A–27A.
3. Luke A, Guo X, Adeyemo AA, et al. Heritability of obesity-related traits among Nigerians, Jamaicans and US black people. *Int J Obes Relat Metab Disord*. 2001; 25:1034–1041. [PubMed: 11443503]
4. Deng HW, Deng H, Liu YJ, et al. A genomewide linkage scan for quantitative-trait loci for obesity phenotypes. *Am J Hum Genet*. 2002; 70:1138–1151. [PubMed: 11923910]
5. Peeters AV, Beckers S, Verrijken A, et al. Association of SIRT1 gene variation with visceral obesity. *Hum Genet*. 2008; 124:431–436. [PubMed: 18820948]
6. Rouskas K, Paletas K, Kalogeridis A, et al. Association between BBS6/ MKKS gene polymorphisms, obesity and metabolic syndrome in the Greek population. *Int J Obes (Lond)*. 2008; 32:1618–1625. [PubMed: 18813213]
7. Rankinen T, Zuberi A, Chagnon YC, et al. The human obesity gene map: the 2005 update. *Obesity (Silver Spring)*. 2006; 14:529–644. [PubMed: 16741264]
8. Lindgren CM, Heid IM, Randall JC, et al. Wellcome Trust Case Control Consortium; Procardis Consortia; Giant Consortium. Genome-wide association scan meta-analysis identifies three Loci influencing adiposity and fat distribution. *PLoS Genet*. 2009; 5:e1000508. [PubMed: 19557161]
9. Liu YJ, Liu XG, Wang L, et al. Genome-wide association scans identified *CTNBL1* as a novel gene for obesity. *Hum Mol Genet*. 2008; 17:1803–1813. [PubMed: 18325910]
10. Norris JM, Langefeld CD, Talbert ME, et al. Genome-wide association study and follow-up analysis of adiposity traits in Hispanic Americans: the IRAS Family Study. *Obesity (Silver Spring)*. 2009; 17:1932–1941. [PubMed: 19461586]
11. Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. *Nat Rev Genet*. 2006; 7:85–97. [PubMed: 16418744]
12. Lupski JR. An evolution revolution provides further revelation. *Bioessays*. 2007; 29:1182–1184. [PubMed: 18008371]
13. Redon R, Ishikawa S, Fitch KR, et al. Global variation in copy number in the human genome. *Nature*. 2006; 444:444–454. [PubMed: 17122850]
14. Aitman TJ, Dong R, Vyse TJ, et al. Copy number polymorphism in *Fcgr3* predisposes to glomerulonephritis in rats and humans. *Nature*. 2006; 439:851–855. [PubMed: 16482158]
15. Gonzalez E, Kulkarni H, Bolivar H, et al. The influence of *CCL3L1* gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science*. 2005; 307:1434–1440. [PubMed: 15637236]
16. Sebat J, Lakshmi B, Malhotra D, et al. Strong association of de novo copy number mutations with autism. *Science*. 2007; 316:445–449. [PubMed: 17363630]
17. Yang Y, Chung EK, Wu YL, et al. Gene copy-number variation and associated polymorphisms of complement component C4 in human systemic lupus erythematosus (SLE): low copy number is a risk factor for and high copy number is a protective factor against SLE susceptibility in European Americans. *Am J Hum Genet*. 2007; 80:1037–1054. [PubMed: 17503323]
18. Stranger BE, Forrest MS, Dunning M, et al. Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science*. 2007; 315:848–853. [PubMed: 17289997]
19. Eichler EE, Sankoff D. Structural dynamics of eukaryotic chromosome evolution. *Science*. 2003; 301:793–797. [PubMed: 12907789]
20. Stankiewicz P, Shaw CJ, Withers M, Inoue K, Lupski JR. Serial segmental duplications during primate evolution result in complex human genome architecture. *Genome Res*. 2004; 14:2209–2220. [PubMed: 15520286]
21. Sebat J, Lakshmi B, Troge J, et al. Large-scale copy number polymorphism in the human genome. *Science*. 2004; 305:525–528. [PubMed: 15273396]

22. Mutirangura A, Greenberg F, Butler MG, et al. Multiplex PCR of three dinucleotide repeats in the Prader-Willi/Angelman critical region (15q11–q13): molecular diagnosis and mechanism of uniparental disomy. *Hum Mol Genet.* 1993; 2:143–151. [PubMed: 8499903]
23. Zody MC, Garber M, Sharpe T, et al. Analysis of the DNA sequence and duplication history of human chromosome 15. *Nature.* 2006; 440:671–675. [PubMed: 16572171]
24. Mao R, Jalal SM, Snow K, et al. Characteristics of two cases with dup(15) (q11. 2–q12): one of maternal and one of paternal origin. *Genet Med.* 2000; 2:131–135. [PubMed: 11397326]
25. Christian SL, Fantes JA, Mewborn SK, Huang B, Ledbetter DH. Large genomic duplicons map to sites of instability in the Prader-Willi/Angelman syndrome chromosome region (15q11–q13). *Hum Mol Genet.* 1999; 8:1025–1037. [PubMed: 10332034]
26. Ahn SJ, Costa J, Emanuel JR. PicoGreen quantitation of DNA: effective evaluation of samples pre- or post-PCR. *Nucleic Acids Res.* 1996; 24:2623–2625. [PubMed: 8692708]
27. Singer VL, Jones LJ, Yue ST, Haugland RP. Characterization of PicoGreen reagent and development of a fluorescence-based solution assay for double-stranded DNA quantitation. *Anal Biochem.* 1997; 249:228–238. [PubMed: 9212875]
28. Rabbee N, Speed TP. A genotype calling algorithm for affymetrix SNP arrays. *Bioinformatics.* 2006; 22:7–12. [PubMed: 16267090]
29. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics.* 2003; 19:185–193. [PubMed: 12538238]
30. Bailey JA, Gu Z, Clark RA, et al. Recent segmental duplications in the human genome. *Science.* 2002; 297:1003–1007. [PubMed: 12169732]
31. Pinto D, Marshall C, Feuk L, Scherer SW. Copy-number variation in control population cohorts. *Hum Mol Genet.* 2007; 16(Spec No. 2):R168–R173. [PubMed: 17911159]
32. Butler MG, Fischer W, Kibiryeva N, Bittel DC. Array comparative genomic hybridization (aCGH) analysis in Prader-Willi syndrome. *Am J Med Genet A.* 2008; 146:854–860. [PubMed: 18266248]
33. Wang YM, Chuang L, Wang BT, Kuo PL. Maternal uniparental disomy in a patient with Prader-Willi syndrome with an additional small inv dup(15) chromosome. *J Formos Med Assoc.* 2004; 103:943–947. [PubMed: 15624046]
34. Buiting K, Nazlican H, Galetzka D, et al. C15orf2 and a novel noncoding transcript from the Prader-Willi/Angelman syndrome region show monoallelic expression in fetal brain. *Genomics.* 2007; 89:588–595. [PubMed: 17337158]
35. Sahoo T, del Gaudio D, German JR, et al. Prader-Willi phenotype caused by paternal deficiency for the HBII-85 C/D box small nucleolar RNA cluster. *Nat Genet.* 2008; 40:719–721. [PubMed: 18500341]
36. Kanber D, Giltay J, Wiczorek D, et al. A paternal deletion of *MKRN3*, *MAGEL2* and *NDN* does not result in Prader-Willi syndrome. *Eur J Hum Genet.* 2009; 17:582–590. [PubMed: 19066619]
37. de Smith AJ, Purmann C, Walters RG, et al. A deletion of the *HBII-85* class of small nucleolar RNAs (snoRNAs) is associated with hyperphagia, obesity and hypogonadism. *Hum Mol Genet.* 2009; 18:3257–3265. [PubMed: 19498035]
38. Skryabin BV, Gubar LV, Seeger B, et al. Deletion of the *MBII-85* snoRNA gene cluster in mice results in postnatal growth retardation. *PLoS Genet.* 2007; 3:e235. [PubMed: 18166085]
39. Tsai TF, Armstrong D, Beaudet AL. Necdin-deficient mice do not show lethality or the obesity and infertility of Prader-Willi syndrome. *Nat Genet.* 1999; 22:15–16. [PubMed: 10319852]
40. Wawrzik M, Spiess AN, Herrmann R, Buiting K, Horsthemke B. Expression of *SNURF-SNRPN* upstream transcripts and epigenetic regulatory genes during human spermatogenesis. *Eur J Hum Genet.* 2009; 17:1463–1470. [PubMed: 19471314]
41. Emanuel BS, Shaikh TH. Segmental duplications: an ‘expanding’ role in genomic instability and disease. *Nat Rev Genet.* 2001; 2:791–800. [PubMed: 11584295]
42. Bolton PF, Dennis NR, Browne CE, et al. The phenotypic manifestations of interstitial duplications of proximal 15q with special reference to the autistic spectrum disorders. *Am J Med Genet.* 2001; 105:675–685. [PubMed: 11803514]

43. Veltman MW, Thompson RJ, Craig EE, et al. A paternally inherited duplication in the Prader-Willi/Angelman syndrome critical region: a case and family study. *J Autism Dev Disord*. 2005; 35:117–127. [PubMed: 15796127]
44. Cook EH Jr, Lindgren V, Leventhal BL, et al. Autism or atypical autism in maternally but not paternally derived proximal 15q duplication. *Am J Hum Genet*. 1997; 60:928–934. [PubMed: 9106540]
45. Mohandas TK, Park JP, Spellman RA, et al. Paternally derived de novo interstitial duplication of proximal 15q in a patient with developmental delay. *Am J Med Genet*. 1999; 82:294–300. [PubMed: 10051161]
46. McCarroll SA, Kuvshinov FG, Korn JM, et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet*. 2008; 40:1166–1174. [PubMed: 18776908]

Figure 1.

Defining copy number variation (CNV) regions, and CNV sub-region. CNV was determined by the start and end points based on physical position of Affymetrix 500K single-nucleotide polymorphisms (SNPs) on chromosomes. CNV region is the extending genomic region by overlapped CNVs.

Figure 2.

Schematic overview of 1Mb Prader-Willi syndrome (PWS) region between *NDN* and *C15orf2* genes containing the *PWRN1* gene with its partial duplications and CNV_031, CNV_034 and CNV_035. Open box (the first from right): *PWRN1* gene; other three open boxes: partial duplications of *PWRN1* gene. Except one for CNV_034, other black arrows point to *PWRN1* exons and their duplicated copies. Contig positions are shown at the side of bidirectional arrows. Not drawn to scale. CNV, copy number variation.

Table 1

Basic characteristics of the subjects

Trait	Male (<i>n</i> = 492)	Female (<i>n</i> = 493)
Age (years)	50.49 (18.87)	50.30 (17.65)
Height (cm)	177.78 (7.06)	163.80 (6.53)
Weight (kg)	88.86 (14.88)	71.25 (15.88)
Fat mass (kg)	23.29 (8.87)	26.61 (10.41)

Data presented are unadjusted means (s.d.).

Table 2

A ssociation analysis of CNVs with body fat mass

CNV_ID	Splitter	ID_dbSNP	Phy_position ^a	Chromosome	Cytoband	Probe number	Freq ^b (%)	P value ^c
CNV_034	SNP_A-2090337	rs7496441	22278324	15	q11.2	1	1.22	0.0042
CNV_035	SNP_A-2290456	rs11634882	22289332	15	q11.2	3	1.12	0.0043
CNV_031	SNP_A-2192997	rs2883495	22140119	15	q11.2	3	2.84	0.0420

CNV, copy number variation; SNP, single-nucleotide polymorphisms.

^aPhysical position.

^bFrequency.

^cOnly the nominally significant results are listed.

Table 3

A ssociation analysis of CNV_034, CNV_035 with body fat mass and other body composition indexes

Traits ^a	CNV_031		CNV_034		CNV_035	
	Normal	Gain	Normal	Gain	Normal	Gain
Fat mass (kg)						
T	24.81 (9.72)	29.88 (11.58) ^b	24.80 (9.68)	34.59 (14.30) ^b	24.82 (9.69)	34.59 (14.30) ^b
F	26.49 (10.33)	30.76 (12.70)	26.43 (10.27)	37.89 (14.10) ^b	26.47 (10.29)	37.89 (14.10) ^b
M	23.16 (8.79)	28.71 (10.51) ^b	23.20 (8.78)	29.64 (15.08)	23.20 (8.78)	29.64 (15.08)
Lean mass (kg)						
T	53.59 (12.57)	53.78 (11.14)	53.61 (12.58)	53.44 (9.55)	53.61 (12.57)	53.44 (9.55)
F	43.40 (6.68)	46.00 (6.69)	43.39 (6.68)	47.29 (6.49)	43.41 (6.69)	47.29 (6.49)
M	63.55 (8.24)	64.15 (6.10)	63.59 (8.22)	62.66 (3.86)	63.59 (8.22)	62.66 (3.86)
BMI (kg/m ²)						
T	27.28 (5.26)	29.38 (6.29) ^b	27.28 (5.23)	31.30 (7.97) ^b	27.29 (5.25)	31.30 (7.97) ^b
F	26.51 (5.89)	28.58 (6.95)	26.49 (5.86)	31.77 (8.37) ^c	26.52 (5.9)	31.77 (8.37) ^c
M	28.04 (4.41)	30.44 (5.49) ^c	28.06 (4.39)	30.61 (8.54)	28.06 (4.39)	30.61 (8.54)
Trunk fat mass (kg)						
T	12.21 (5.75)	14.21 (6.57) ^c	12.19 (5.73)	17.28 (7.44) ^b	12.21 (5.74)	17.28 (7.44) ^b
F	12.03 (5.88)	13.49 (6.93)	11.98 (5.86)	17.81 (6.93) ^b	12.01 (5.89)	17.81 (6.93) ^b
M	12.38 (5.62)	15.18 (6.31) ^b	12.40 (5.60)	16.49 (9.20)	12.40 (5.6)	16.49 (9.2)

CN, copy number; F, female; M, male; T, total sample.

^aTrait values presented are unadjusted means (s.d.).

^bStatistically higher than the CN normal group at the 5% significance level using one-sided *F*-test.

^cStatistically higher than the CN normal group at the 10% significance level using one-sided *F*-test.