BMB
reports

Analysis of copy number variation in 8,842 Korean individuals reveals 39 genes associated with hepatic biomarkers AST and ALT

Hyo Young Kim[1], Seoae Cho[2], Jeongmi Yu[1], Samsun Sung[1] & Heebal Kim[1,]*

[1]Department of Agricultural Biotechnology and Research Institute for Agriculture and Life Sciences, Seoul National University,
[2]Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul 151-742, Korea

Biochemical tests such as aspartate aminotransferase (AST) and alanine aminotransferase (ALT) are useful for diagnosing patients with liver disease. In this study, we tested the association between copy number variation and the hepatic biomarkers AST and ALT based on 8,842 samples from population-based cohorts in Korea. We used Affymetrix Genome-Wide Human 5.0 arrays and identified 10,534 CNVs using HelixTree software. Of the CNVs tested using univariate linear regression, 100 CNVs were significant for AST and 16 were significant for ALT ($P < 0.05$). We identified 39 genes located within the CNV regions. *DKK1* and *HS3ST3B1* were shown to play roles in heparan sulfate biosynthesis and the Wnt signaling pathway, respectively. *NAF1* and *NPY1R* were associated with glycoprotein processes and neuropeptide Y receptor activity based on GO categories. *PTER*, *SOX14* and *TM7SF4* were expressed in liver. *DPYS* and *CTSC* were found to be associated with dihydropyrimidinuria and Papillon-Lefèvre syndrome phenotypes using *OMIM*. *NPY5R* was found to be associated with dyslipidemia using the *Genetic Association Database*. [BMB reports 2010; 43(8): 547-553]

INTRODUCTION

The liver is the largest glandular organ in the human body. The organ has many functions, including storage of glycogen, filtration of harmful substances such as alcohol, and maintenance of normal glucose concentration. The liver also produces urea and the majority of cholesterol in the body (about 80% of the body) (http://www.mamashealth.com/) (1-3). Biochemical tests for liver function are commonly used to diagnose patients with liver disease. Aspartate aminotransferase (AST) and alanine aminotransferase (ALT) are two such biochemicals that are used as markers for evaluating liver injury (4-6). The ratio of serum levels of AST/ALT is used as an indicator for the evaluation of alcoholic liver disease. An AST/ALT ratio of less than one implies moderate liver disease, such as nonalcoholic fatty liver disease (NAFLD). In contrast, an AST/ALT ratio greater than one implies severe liver disease, such as chronic hepatitis or alcoholic fatty liver disease (7). Two loci (10q24.2 and 22 q13.31) have been identified as influencing the plasma levels of ALT in three European populations (Switzerland: n = 5,636, Italy: n = 1,200, London: n = 879) (8).

Genetic differences that exist between various human genomes are attributable in part to large-scale structural variations between individuals. Differences in copy number contribute to changes in gene expression. Hence, DNA copy number variations (CNVs) contribute to genomic variation between humans (9, 10). While many CNVs have been identified in human genomes, no studies have reported CNVs that are correlated with liver disease. Research aimed at evaluating the genetic variation between humans could contribute to our understanding of dosage effect and disease susceptibility (11, 12). Many DNA structural variations in human genomes have been identified in various populations (13, 14). Data from these studies can be found in public databases such as the Database of Genomic Variants (http://projects.tcag.ca/variation/; structural variation in "normal" populations) (15). However, CNV association studies have been hindered in part by inadequate standardization of CNV detection methods. Further, while annotated CNVs have been reported for Caucasians, Africans, Chinese and Japanese, they may not appropriately reflect the CNVs in the genomes of other ethnic groups (16).

While many studies have examined the biology of liver disease in humans, few have focused on the identification of liver-associated CNVs; moreover, CNVs have not been identified in Koreans. Accordingly, our study addresses three main questions. First, there are many tests that evaluate biochemical liver function or liver enzymes. Therefore, what are the liver-associated CNVs that are correlated with these biochemicals? Second, while many CNVs have been annotated in Caucasians, why have so few CNVs been reported in Koreans? Third, if liver-associated CNVs in the genomes of Koreans are identi-

*Corresponding author. Tel: 82-2-880-4803; Fax: 82-2-883-8812; E-mail: heebal@snu.ac.kr

Received 17 April 2010, Accepted 16 July 2010

Keywords: Alanine aminotransferase, Aspartate aminotransferase, Copy number variation, Liver, Univariate linear regression

fied, what are their biological significance? We also tried to identify liver-associated CNVs in Koreans and determine their biological significance. Therefore, we analyzed 8,842 unrelated Korean (Ansung and Ansan) individuals using an Affymetrix Genome-Wide Human 5.0 array and identified 10,534 CNVs using the Golden Helix software. Using univariate linear regression, we screened 39 genes within CNVs associated with the hepatic biomarkers AST and ALT. Data obtained from the Korean Genome Association Study combined with the results of this report provide valuable CNV-related information associated with liver disease.

RESULTS AND DISCUSSION

Analysis of serum liver enzymes

We first tried to identify CNVs associated with hepatic biomarkers (AST and ALT) within the Korean population. For these studies, the value of ALT was transformed to 1/square root(y) while AST was transformed to 1/(y). The value of AST ranged from 0.001 to 0.091 IU/L with a mean of 0.04 ± 0.01, and the value of ALT ranged from 0.03 to 0.27 IU/L with a mean of 0.21 ± 0.05. A beanplot is suitable for visualizing distributions of individual observations on a one-dimensional scatterplot (17). We therefore used beanplots to compare the distributions of AST and ALT between Korean subpopulations, Ansan and Ansung (Fig. 1). Using the beanplot, we did not observe a difference between the two populations. However, we did see a difference between men and women. Based on these observations, we corrected for location, age and gender in our statistical analysis model. Further, we computed Pearson's correlation coefficients to evaluate whether or not AST and ALT shared a conserved relationship. Our results showed that AST had a significant ($P < 0.05$) positive correlation with ALT (cor = 0.73).

Thirty-nine genes are associated with hepatic biomarkers

We identified 10,534 CNVs with a median size of 112 Kb from 8,842 unrelated Korean individuals using the Golden Helix software. These results were very different compared to the size distributions and counts of CNVs in a previous study (16). However, this result was similar to a previous result based on a threshold of ±0.4. This discrepancy was simple because CNVs defining criterion are very diverse depending on the program or algorithm. Although univariate linear regression is commonly used to identify single nucleotide polymorphisms (SNPs) (18), there have been no reports on its use to discover CNVs. Therefore, a univariate linear regression model was fitted to explain the impact of single CNV regions on each quantitative trait. The positive β values of AST and ALT were 4,200 and 5,384, respectively, whereas the negative β values were 6,334 and 5,150, respectively. Highly significant CNVs were visualized in Manhattan plots, including 100 loci for AST and 16 loci for ALT (Fig. 2). We detected 39 genes that were completely located within the CNV regions. The gene lists, beta-coefficients and liver-associated phenotypes are summarized in Table 1. Ten of the genes were associated with liver phenotypes or diseases described in prior studies.

Heparan sulfate biosynthesis and Wnt signaling pathway are enriched in AST and ALT

To examine the functional implications of our CNVs, we functionally annotated genes located in our CNVs using the

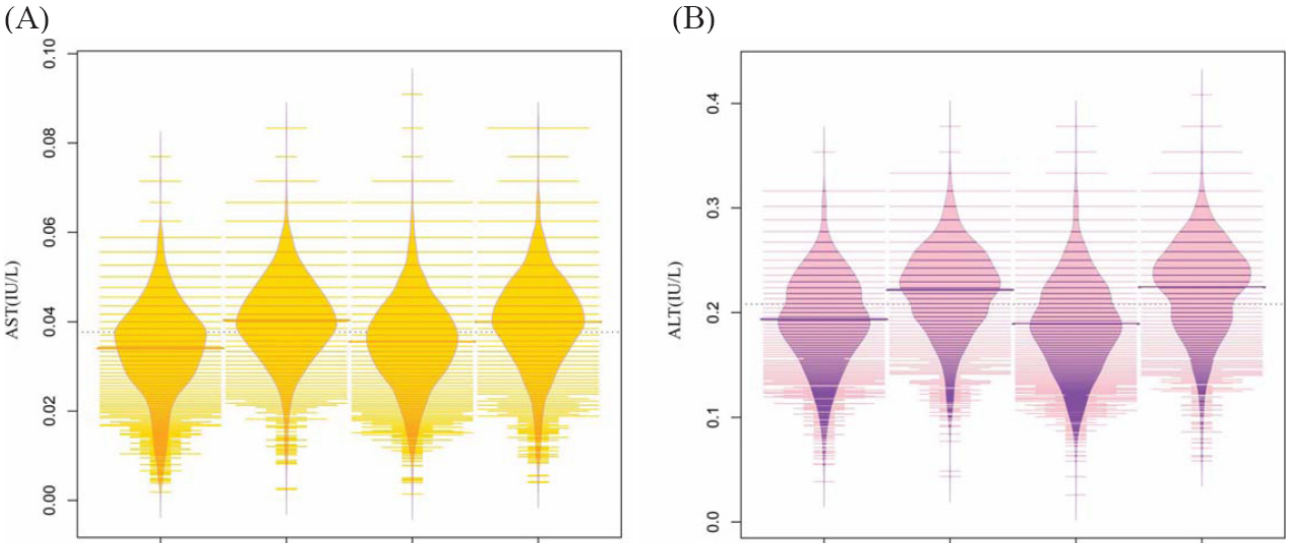

Fig. 1. Beanplots of the distributions of AST and ALT in 8,842 unrelated Korean (Ansung and Ansan) individuals. Thick lines denote the average values of AST (A) and ALT (B). AST, aspartate aminotransferase; ALT, alanine aminotransferase.

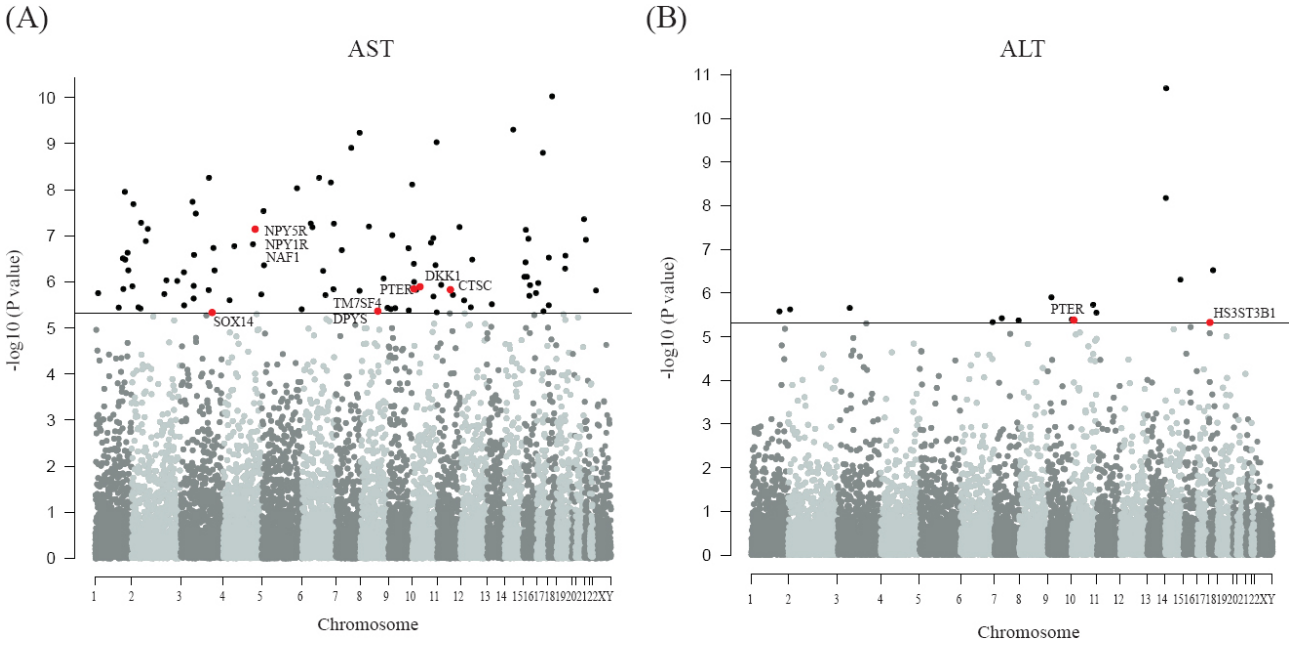

Fig. 2. Visualization of genome-wide association data Manhattan plots of the degree of association ($-\log_{10}$ P value) between 10,534 CNVs for AST (A) and ALT (B) traits among 8,842 individuals. Black dots represent CNVs with P values < 0.05. Red dots represent liver-associated genes identified in previous studies.

DAVID functional annotation tool (19). We identified several biochemical pathways and Gene Ontology (GO) annotations relevant to AST and ALT. Four genes (*DKK1, DPYS, HS3ST3B1* and *MAP3K7*) had KEGG pathway information. The genes annotated using KEGG represented 10 biochemical pathways, including that of heparan sulfate biosynthesis, pyrimidine and beta-alanine metabolism and Wnt signaling. The *HS3ST3B1* gene was found to be involved in the biosynthesis of heparin sulfate, which is a polysaccharide complex synthesized in most mammalian cells (20). The *3OST3B* gene shows wide expression of multiple transcripts and is most abundant in the liver (21, 20). *DKK1* and *MAP3K7* were found to be involved in the Wnt signaling pathway, which plays an important role in developing and regenerating the liver (22). *DKK1* expression was down-regulated in fetal liver and inhibits Wnt signaling in mammalian cells (23).

Glycoprotein process and neuropeptide Y receptor activity are enriched in AST and ALT

The biological characteristics of 39 genes were detected using GO categories. The analyzed functions are summarized in Supplementary Table 1, and the enriched GO terms are visualized in Fig. 3. Enriched functions associated with liver biology were found to be related to glycoprotein biosynthetic and metabolic processes (Fig. 3A). The liver produces the glycoprotein hormone that regulates production of bone marrow platelets (http://review-center.net/metabolism/liver-metabolism-pathways-and-its-disorders/). Several glycoproteins, including fibronectin, hyalurinic, laminin, merosin, nidogen and tenascin, are expressed in fibrotic livers (24, 25). One such glycoprotein, GP73 (Golgi protein), is up-regulated upon hepatitis viral infection (26, 27). Enriched molecular functions associated with liver biology include neuropeptide Y receptor activity (Fig. 3B). Neuropeptide Y was identified in human liver where it regulates blood flow and secretion in the liver (28, 29).

Ten genes are associated with liver disease

Using the *Genetic Association Database* (GAD), we detected one gene associated with liver disease, *NPY5R*. Neuropeptide Y receptor Y5 (*NPY5R*) is known to be associated with dyslipidemia, a fatty liver disease. Marceau *et al*. (2010) showed dyslipidemia is an important risk factor for fatty liver disease (30). Five genes (*CTSC, DPYS, HS3ST3B1, PRM3* and *SPATA7*) were shown to be correlated with human disease states using *OMIM*. The genes annotated using *OMIM* represent nine disease phenotypes, including dihydropyrimidinuria, Papillon-Lefèvre syndrome (PLS) and Haim-Munk syndrome. *DPYS* and *CTSC* were found to be associated with dihydropyrimidinuria, a deficiency in dihydropyrimidinase (DHP), and PLS phenotypes, respectively. The activity of DHP, which is exclusively expressed in the liver, is characterized by increased excretion of dihydrothymine and dihydrouracil (31, 32). Mutations in the cathepsin C (*CTSC*) gene cause Haim-Munk syndrome and PLS, a rare autosomal-recessive disease characterized by juve-

Table 1. Thirty-nine genes associated with serum liver enzymes (AST and ALT)

Trait	CNVR	Gene[a]	Beta-coefficient	P value	Liver-associated phenotype	Literature (year)
AST	Chr4:164012707-164647495	*NPY5R**	−0.0126	7.43E-04	Dyslipidemia-caused fatty liver disease	Marceau *et al*. (2010)
	Chr4:164012707-164647495	*NPY1R**	−0.0126	7.43E-04	Neuropeptide Y receptor activity	GO
	Chr4:164012707-164647495	*NAF1**	−0.0126	7.43E-04	Glycoprotein process	GO
	Chr10:53731444-53843163	*DKK1**	−0.0071	1.38E-02	Wnt signaling inhibitor	Fedi *et al*. (1999)
	Chr11:87666857-87892347	*CTSC**	−0.0117	1.55E-02	Papillon−Lefèvre syndrome	Almuneef *et al*. (2003)
	Chr8:105247995-105727124	*TM7SF4**	−0.0126	4.45E-02	Highest expression in liver	Staege *et al*. (2001)
	Chr8:105247995-105727124	*DPYS**	−0.0126	4.45E-02	Dihydropyrimidinuria	Nyhan (2005)
	Chr3:138809756-139122947	*SOX14**	−0.0113	4.80E-02	Lower level in adult liver	Arsic *et al*. (1998)
	Chr6:91257342-91444831	*MAP3K7*	−0.0108	5.85E-05		
	Chr4:164012707-164647495	*TKTL2*	−0.0126	7.43E-04		
	Chr15:54342294-54871765	*ZNF280D*	−0.0152	1.22E-03		
	Chr15:54342294-54871765	*TEX9*	−0.0152	1.22E-03		
	Chr15:54342294-54871765	*MNS1*	−0.0152	1.22E-03		
	Chr2: 64521312-64568911	*HSPC159*	−0.0043	1.38E-03		
	Chr15:46811866-47497399	*SHC4*	−0.0176	8.10E-03		
	Chr15:46811866-47497399	*COPS2*	−0.0176	8.10E-03		
	Chr15:46811866-47497399	*GALK2*	−0.0176	8.10E-03		
	Chr15:46811866-47497399	*SECISBP2L*	−0.0176	8.10E-03		
	Chr15:46811866-47497399	*CEP152*	−0.0176	8.10E-03		
	Chr15:46811866-47497399	*EID1*	−0.0176	8.10E-03		
	Chr2:224474138-224753508	*SERPINE2*	−0.0123	1.00E-02		
	Chr2:224474138-224753508	*MRPL44*	−0.0123	1.00E-02		
	Chr16:11261998-11292512	*TNP2*	−0.0026	1.11E-02		
	Chr16:11261998-11292512	*PRM2*	−0.0026	1.11E-02		
	Chr16:11261998-11292512	*PRM3*	−0.0026	1.11E-02		
	Chr16:11261998-11292512	*PRM1*	−0.0026	1.11E-02		
	Chr12:29089080-29470913	*ERGIC2*	−0.01	2.62E-02		
	Chr12:29089080-29470913	*FAR2*	−0.01	2.62E-02		
	Chr8:105247995-105727124	*LRP12*	−0.0126	1.45E-02		
AST/ALT	Chr10:16371836-16615099	*PTER**	0.0143	1.50E-02	Low expression in liver	Hou *et al*. (1996)
	Chr10:16371836-16615099	*C1QL3*	0.0143	1.50E-02		
ALT	Chr17:14014164-14613835	*HS3ST3B1**	−0.009	4.97E-02	Abundant in liver	Shworak *et al*. (1999)
	Chr14:87637635-88374957	*KCNK10*	0.0142	5.18E-03		
	Chr14:87637635-88374957	*ZC3H14*	0.0142	5.18E-03		
	Chr14:87637635-88374957	*PTPN21*	0.0142	5.18E-03		
	Chr14:87637635-88374957	*SPATA7*	0.0142	5.18E-03		
	Chr14:87637635-88374957	*EML5*	0.0142	5.18E-03		
	Chr17:14014164-14613835	*CDRT15*	−0.009	4.97E-02		
	Chr17:14014164-14613835	*MGC12916*	−0.009	4.97E-02		

[a]There are 39 genes (P value < 0.05) significantly selected for each trait; *The 10 genes identified as liver-associated in previous studies are indicated by asterisks (*).

nile periodontitis. Pyogeneic liver abscesses are a well recognized complication of neutrophil dysfunction in PLS (33). Four genes (*CTSC, DPYS, GALK2* and *PTER*) were found to be actively expressed in human liver using *BioGPS* (34). This is evident from gene expression patterns produced by the GeneAtlas U133A data sets. Further, Pter expression was down-regulated in mouse liver tissue (35).

MATERIALS AND METHODS

Study subjects

The Korea Association Resource (KARE) project was initiated in 2007 for the large-scale genome-wide association analysis of 10,038 unrelated Korean individuals. Among them, we selected 8,842 chips for CNV analysis in order to obtain genotyping data for quality control (QC) purposes (36). Samples

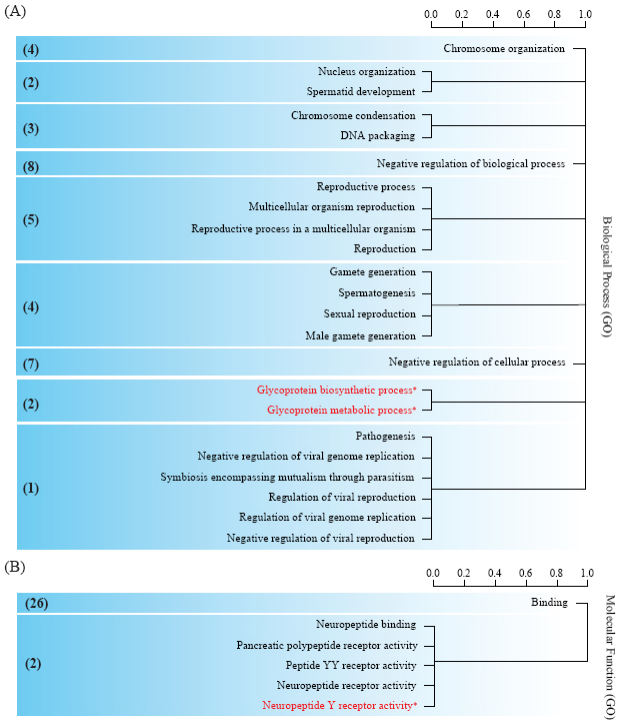

Fig. 3. Tree views of enriched GO categories Enriched GO categories are visualized for 39 genes found within CNVs associated with AST and ALT. Numbers in parentheses at the left denote the gene count within GO groups. The terms identified as liver-associated in previous studies are indicated by asterisks (*).

with high heterozygosity, high missing genotype call rate, gender inconsistencies, and individuals with cancer were excluded. The average pair-wise identity-by-state values of identical or related individuals were higher than the computed values of Korean sib-pair samples. As a result, 8,842 chips were obtained for the Ansung (2,374 men and 2,263 women) and Ansan (1,809 men and 2,396 women) population-based cohorts. The mean age was 52.2 years (standard deviation 8.94). Regarding the participants, genomic DNA was isolated from peripheral blood.

CNVs discovery

Genome-wide variations were measured using an Affymetrix Genome-Wide Human 5.0 array. We identified CNV regions using copy-number analysis module (CNAM) and HelixTree software version 7.0 (Golden Helix Inc., Bozeman, MT, USA) (37). Copy number analysis was performed to create normalized log2 ratios. Speicifically, the CNAM module was used to read the Affymetrix CEL intensity files, normalize intensity values against reference samples and import log2 ratios. We used all 8,842 chips as a reference. It should be mentioned that generating a reference from other ethnic chips results in increased variability due to systematic differences. Therefore, instead of using another ethnic sample or a small sample as a reference, referencing all of the samples can decrease variability. The analysis parameters included a multivariate algorithm, moving 5,000 window sizes, a maximum of 100 segments per window, a minimum of 10 markers per segment, and P values less than 0.01 for pair-wise permutations (n = 1,000). The multivariate algorithm segmented all samples simultaneously, making it possible to perform the CNV association study for all samples.

Screening CNV association tests

To explain the impact of the single CNV regions on each quantitative trait, we performed univariate linear regression (38). The additive genetic models included correlations for sex, age and location of individuals. For multiple testing, significance was determined using a P value and a Bonferroni cutoff of 0.05. The intensity of each CNV associated with continuous response variables was tested via the following univariate linear regression model.

For continuous variables,

$$Y = \beta_0 + \beta_1 CNV + \beta_2 Area + \beta_3 Age + \beta_4 Gender + \varepsilon$$

where is the p-vector of the coefficients. Statistical analyses were performed using the statistical package R (http://www.r-project.org/) (ver. 2.9).

Gene enrichment analysis

RefGene was downloaded from the UCSC Genome Browser (ver. hg18; http://genome.ucsc.edu/). We detected genes that were completely located within our CNV regions. To functionally analyze genes located in our CNVs, we performed two analyses using the DAVID tool (ver. 6.7 Beta; http://david.abcc.ncifcrf.gov/). The two functional tools we used included Gene Ontology (39) and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis (40). Gene-disease associations were obtained using OMIM (http://www.ncbi.nlm.nih.gov/sites/entrez?db=omim), the Genetic Association Database (http://geneticassociationdb.nih.gov/cgi-bin/index.cgi) and BioGPS (http://biogps.gnf.org/#goto=welcome). All parsing was performed using Python software (ver. 2.5).

Acknowledgements

The consortium for large-scale genotyping data was supported by the Korea Association Resource (KARE) project funded by the Korean National Institute of Health, Republic of Korea.

REFERENCES

1. Gitzelmann, R., Spycher, M., Feil, G., Muller, J., Seilnacht, B., Stahl, M. and Bosshard, N. (1996) Liver glycogen synthase deficiency: a rarely diagnosed entity. *Eur. J. Pediatr.* **155**, 561-567.
2. Pocai, A., Lam, T., Obici, S., Gutierrez-Juarez, R., Muse, E., Arduini, A. and Rossetti, L. (2006) Restoration of hypo-

thalamic lipid sensing normalizes energy and glucose homeostasis in overfed rats. *J. Clin. Invest.* **116**, 1081-1091.
3. Zhang, X. and Beynen, A. (2007) Influence of dietary fish proteins on plasma and liver cholesterol concentrations in rats. *Br. J. Nutr.* **69**, 767-777.
4. Bathum, L., Petersen, H., Rosholm, J., Hyltoft Petersen, P., Vaupel, J. and Christensen, K. (2001) Evidence for a substantial genetic influence on biochemical liver function tests: results from a population-based Danish twin study. *Clin. Chem.* **47**, 81-87.
5. Hanley, A., Williams, K., Festa, A., Wagenknecht, L., D'Agostino, R. and Haffner, S. (2005) Liver markers and development of the metabolic syndrome. *Diabetes* **54**, 3140-3147.
6. Sattar, N., Scherbakova, O., Ford, I., O'Reilly, D., Stanley, A., Forrest, E., MacFarlane, P., Packard, C., Cobbe, S. and Shepherd, J. (2004) Elevated alanine aminotransferase predicts new-onset type 2 diabetes independently of classical risk factors, metabolic syndrome, and C-reactive protein in the west of Scotland coronary prevention study. *Diabetes* **53**, 2855-2860.
7. Clemenz, M., Frost, N., Schupp, M., Caron, S., Foryst-Ludwig, A., Bohm, C., Hartge, M., Gust, R., Staels, B. and Unger, T. (2008) Liver-specific peroxisome proliferator-activated receptor α target gene regulation by the angiotensin type 1 receptor blocker telmisartan. *Diabetes* **57**, 1405-1413.
8. Yuan, X., Waterworth, D., Perry, J., Lim, N., Song, K., Chambers, J., Zhang, W., Vollenweider, P., Stirnadel, H. and Johnson, T. (2008) Population-based genome-wide association studies reveal six loci influencing plasma levels of liver enzymes. *Am. J. Hum. Genet.* **83**, 520-528.
9. Iafrate, A., Feuk, L., Rivera, M., Listewnik, M., Donahoe, P., Qi, Y., Scherer, S. and Lee, C. (2004) Detection of large-scale variation in the human genome. *Nature Genetics* **36**, 949-951.
10. Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P., Maner, S., Massa, H., Walker, M. and Chi, M. (2004) Large-scale copy number polymorphism in the human genome. *Science* **305**, 525-528.
11. Feuk, L., Carson, A. and Scherer, S. (2006) Structural variation in the human genome. *Nature Reviews Genetics* **7**, 85-97.
12. Redon, R., Ishikawa, S., Fitch, K., Feuk, L., Perry, G., Andrews, T., Fiegler, H., Shapero, M., Carson, A. and Chen, W. (2006) Global variation in copy number in the human genome. *Nature* **444**, 444-454.
13. de Stahl, T., Sandgren, J., Piotrowski, A., Nord, H., Andersson, R., Menzel, U., Bogdan, A., Thuresson, A., Poplawski, A. and von Tell, D. (2008) Profiling of copy number variations (CNVs) in healthy individuals from three ethnic groups using a human genome 32 K BAC-clone-based array. *Human Mutation* **29**, 398-408.
14. McCarroll, S., Kuruvilla, F., Korn, J., Cawley, S., Nemesh, J., Wysoker, A., Shapero, M., De Bakker, P., Maller, J. and Kirby, A. (2008) Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet* **40**, 1166-1174.
15. Marshall, C., Noor, A., Vincent, J., Lionel, A., Feuk, L., Skaug, J., Shago, M., Moessner, R., Pinto, D. and Ren, Y. (2008) Structural variation of chromosomes in autism spectrum disorder. *Am. J. Hum. Genet.* **82**, 477-488.
16. Yim, S., Kim, T., Hu, H., Kim, J., Kim, B., Lee, J., Han, B., Shin, S., Jung, S. and Chung, Y. (2009) Copy number variations in East-Asian population and their evolutionary and functional implications. *Human Molecular Genetics*. **19**, 1001-1008.
17. Kampstra, P. (2008) Beanplot: a boxplot alternative for visual comparison of distributions, Jour. **28**, co1.
18. Cooper, G., Johnson, J., Langaee, T., Feng, H., Stanaway, I., Schwarz, U., Ritchie, M., Stein, C., Roden, D. and Smith, J. (2008) A genome-wide scan for common genetic variants with a large influence on warfarin maintenance dose. *Blood* **112**, 1022-1027.
19. Dennis, Jr, G., Sherman, B., Hosack, D., Yang, J., Gao, W., Lane, H. and Lempicki, R. (2003) DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol.* **4**, P3.
20. Lyon, M., Deakin, J. and Gallagher, J. (1994) Liver heparan sulfate structure. A novel molecular design. *J. Biol. Chem.* **269**, 11208-11215.
21. Shworak, N., Liu, J., Petros, L., Zhang, L., Kobayashi, M., Copeland, N., Jenkins, N. and Rosenberg, R. (1999) Multiple isoforms of heparan sulfate D-glucosaminyl 3-O-sulfotransferase. *J. Biol. Chem.* **274**, 5170-5184.
22. Armengol, C., Cairo, S., Fabre, M. and Buendia, M. (2009) Wnt signaling and hepatocarcinogenesis: the hepatoblastoma model. *Int. J. Biochem. Cell Biol*. (in press)
23. Fedi, P., Bafico, A., Soria, A., Burgess, W., Miki, T., Bottaro, D., Kraus, M. and Aaronson, S. (1999) Isolation and biochemical characterization of the human Dkk-1 homologue, a novel inhibitor of mammalian Wnt signaling. *J. Biol. Chem.* **274**, 19465-19472.
24. Li, D. and Friedman, S. (1999) Liver fibrogenesis and the role of hepatic stellate cells: new insights and prospects for therapy. *J. Gastroenterol. Hepatol.* **14**, 618-633.
25. Tsukada, S., Parsons, C. and Rippe, R. (2006) Mechanisms of liver fibrosis. *Clinica Chimica Acta* **364**, 33-60.
26. Kladney, R., Bulla, G., Guo, L., Mason, A., Tollefson, A., Simon, D., Koutoubi, Z. and Fimmel, C. (2000) GP73, a novel Golgi-localized protein upregulated by viral infection. *Gene* **249**, 53-65.
27. Block, T., Comunale, M., Lowman, M., Steel, L., Romano, P., Fimmel, C., Tennant, B., London, W., Evans, A. and Blumberg, B. (2005) Use of targeted glycoproteomics to identify serum glycoproteins that correlate with liver cancer in woodchucks and humans. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 779-784.
28. DING, W., Fujimura, M., Mori, A., Tooyama, I. and Kimura, H. (1991) Light and electron microscopy of neuropeptide Y-containing nerves in human liver, gallbladder, and pancreas. *Gastroenterology* **101**, 1054-1059.
29. El-Salhy, M. (2000) Neuropeptide levels in murine liver and biliary pathways. *Ups. J. Med. Sci.* **105**, 207-213.
30. Marceau, P., Biron, S., Hould, F., Marceau, S., Simard, S., Thung, S. and Kral, J. (1999) Liver pathology and the metabolic syndrome X in severe obesity. *J. Clin. Endocrinol. Metab.* **84**, 1513-1517.
31. Van Gennip, A., De Abreu, R., Van Lenthe, H., Bakkeren, J., Rotteveel, J., Vreken, P. and Van Kuilenburg, A. (1997)

Dihydropyrimidinase deficiency: confirmation of the enzyme defect in dihydropyrimidinuria. *J. Inher. Metab. Dis.* **20**, 339-342.
32. Nyhan, W. (2005) Disorders of purine and pyrimidine metabolism. *Mol. Genet. Metab.* **86**, 25-33.
33. Almuneef, M., Al Khenaizan, S., Al Ajaji, S. and Al-Anazi, A. (2003) Pyogenic liver abscess and Papillon-Lefevre syndrome: not a rare association. *Pediatrics* **111**, e85.
34. Wu, C., Orozco, C., Boyer, J., Leglise, M., Goodale, J., Batalov, S., Hodge, C., Haase, J., Janes, J. and Huss, J. (2009) BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biology* **10**, R130.
35. Hou, X., Maser, R., Magenheimer, B. and Calvet, J. (1996) A mouse kidney-and liver-expressed cDNA having homology with a prokaryotic parathion hydrolase (phosphotriesterase)-encoding gene: abnormal expression in injured and polycystic kidneys. *Gene* **168**, 157-163.
36. Cho, Y., Go, M., Kim, Y., Heo, J., Oh, J., Ban, H., Yoon, D., Lee, M., Kim, D. and Park, M. (2009) A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. *Nature Genetics*. **41**, 527-534.
37. Lambert, C. (2005) HelixTree® Genetics Analysis Software. Golden Helix. *Inc. http//wwwgoldenhelixcom*.
38. McMurray, A., Pearson, P., Pace, R. and Scott, D. (2004) Research: a commonsense approach. *Social Science Press*.
39. Harris, M., Clark, J., Ireland, A., Lomax, J., Ashburner, M., Foulger, R., Eilbeck, K., Lewis, S., Marshall, B. and Mungall, C. (2004) The Gene Ontology (GO) database and informatics resource. *Nucleic. Acids Research* **32**, D258-261.
40. Kanehisa, M., Goto, S., Kawashima, S. and Nakaya, A. (2002) The KEGG databases at GenomeNet. *Nucleic. Acids Research* **30**, 42-46.