Nat Rev Genet 7: 745-758

Feuk L, Carson AR, Scherer SW (2006) Structural variation in the human genome. Nat Rev Genet 7: 85-97

Fujimoto A, Kimura R, Ohashi J, Omi K, Yuliwulandari R, Batubara L, Mustofa MS, Samakkarn U, Settheetham-Ishida W, Ishida T, Morishita Y, Furusawa T, Nakazawa M, Ohtsuka R, Tokunaga K (2008) A scan for genetic determinants of human hair morphology: *EDAR* is associated with Asian hair thickness. Hum Mol Genet 17: 835-843

Fullerton SM, Bartoszewicz A, Ybazeta G, Horikawa Y, Bell GI, Kidd KK, Cox NJ, Hudson RR, Di Rienzo A (2002) Geographic and haplotype structure of candidate type 2 diabetes susceptibility variants at the *calpain-10* locus. Am J Hum Genet 70: 1096-1106

Gauderman WJ (2002a) Sample size requirements for matched case-control studies of gene-environment interaction. Stat Med 21: 35-50

Gauderman WJ (2002b) Sample size requirements for association studies of gene-gene interaction. Am J Epidemiol 155: 478-484

Gauderman WJ, Witte JS, Thomas DC (1999) Family-based association studies. J Natl Cancer Inst Monogr: 31-37

Ghanem N, Uring-Lambert B, Abbal M, Hauptmann G, Lefranc MP, Lefranc G (1988) Polymorphism of MHC class III genes: definition of restriction fragment linkage groups and evidence for frequent deletions and duplications. Hum Genet 79: 209-218

Gilles F, Goy A, Remache Y, Manova K, Zelenetz AD (2000) Cloning and characterization of a Golgin-related gene from the large-scale polymorphism linked to the PML gene. Genomics 70: 364-374

Göring HH, Curran JE, Johnson MP, Dyer TD, Charlesworth J, Cole SA, Jowett JB, Abraham LJ, Rainwater DL, Comuzzie AG, Mahaney MC, Almasy L, MacCluer JW, Kissebah AH, Collier GR, Moses EK, Blangero J (2007) Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. Nat Genet 39: 1208-1216

Groot PC, Mager WH, Frants RR (1991) Interpretation of polymorphic DNA patterns in the human alpha-amylase multigene family. Genomics 10: 779-785

Gudbjartsson DF, Jonasson K, Frigge ML, Kong A (2000). Allegro, a new computer program for multipoint linkage analysis. Nat Genet 25: 12-13

Hamblin MT, Thompson EE, Di Rienzo A. (2002) Complex signatures of natural selection at the Duffy blood group locus. Am J Hum Genet 70: 369-383

Hartl DL, Clark AG (1997) Principles of population genetics. Sinauer Associates, Sunderland

Helgason A, Pálsson S, Thorleifsson G, Grant SF, Emilsson V, Gunnarsdottir S, Adeyemo A, Chen Y, Chen G, Reynisdottir I, Benediktsson R, Hinney A, Hansen T, Andersen G, Borch-Johnsen K, Jorgensen T, Schäfer H, Faruque M, Doumatey A, Zhou J, Wilensky RL, Reilly MP, Rader DJ, Bagger Y, Christiansen C, Sigurdsson G, Hebebrand J, Pedersen O, Thorsteinsdottir U, Gulcher JR, Kong A, Rotimi C, Stefánsson K (2007) Refining the impact of *TCF7L2* gene variants on type 2 diabetes and adaptive evolution. Nat Genet 39: 218-225

Hinds DA, Kloek AP, Jen M, Chen X, Frazer KA (2006) Common deletions and SNPs are in linkage disequilibrium in the human genome. Nat Genet 38: 82-85

Hoh J, Wille A, Ott J (2001) Trimming, weighting, and grouping SNPs in human case-control association studies. Genome Res 11: 2115-2119

Hollox EJ, Armour JA, Barber JC (2003) Extensive normal copy number variation of a beta-defensin antimicrobial-gene cluster. Am J Hum Genet 73: 591-600

Horai S, Hayasaka K, Kondo R, Tsugane K, Takahata N (1995) Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. Proc Natl Acad Sci USA 92: 532-536

Horng JT, Hu KC, Wu LC, Huang HD, Lin FM, Huang SL, Lai HC, Chu TY (2004) Identifying the combination of genetic factors that determine susceptibility to cervical cancer. IEEE Trans Inf Technol Biomed 8:59-66

Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW, Lee C (2004) Detection of large-scale variation in the human genome. Nat Genet 36: 949-951

International HapMap Consortium (2005) A haplotype map of the human genome. Nature 437: 1299-1320

International HapMap Consortium, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, Parkin M, Roy J, Stahl E, Winchester E, Ziaugra L, Altshuler D, Shen Y, Yao Z, Huang W, Chu X, He Y, Jin L, Liu Y, Shen Y, Sun W, Wang H, Wang Y, Wang Y, Xiong X, Xu L, Waye MM, Tsui SK, Xue H, Wong JT, Galver LM, Fan JB, Gunderson K, Murray SS, Oliphant AR, Chee MS, Montpetit A, Chagnon F, Ferretti V, Leboeuf M, Olivier JF, Phillips MS, Roumy S, Sallée C, Verner A, Hudson TJ, Kwok PY, Cai D, Koboldt DC, Miller RD, Pawlikowska L, Taillon-Miller P, Xiao M, Tsui LC, Mak W, Song YQ, Tam PK, Nakamura Y, Kawaguchi T, Kitamoto T, Morizono T, Nagashima A, Ohnishi Y, Sekine A, Tanaka T, Tsunoda T, Deloukas P, Bird CP, Delgado M, Dermitzakis ET, Gwilliam R, Hunt S, Morrison J, Powell D, Stranger BE, Whittaker P, Bentley DR, Daly MJ, de Bakker PI, Barrett J, Chretien YR, Maller J, McCarroll S, Patterson N, Pe'er I, Price A, Purcell S, Richter DJ, Sabeti P, Saxena R, Schaffner SF, Sham PC, Varilly P, Altshuler D, Stein LD, Krishnan L, Smith AV, Tello-Ruiz MK, Thorisson GA, Chakravarti A, Chen PE, Cutler DJ, Kashuk CS, Lin S, Abecasis GR, Guan W, Li Y, Munro HM, Qin ZS, Thomas DJ, McVean G, Auton A, Bottolo L, Cardin N, Eyheramendy S, Freeman C, Marchini J, Myers S, Spencer C, Stephens M, Donnelly P, Cardon LR, Clarke G, Evans DM, Morris AP, Weir BS, Tsunoda T, Mullikin JC, Sherry ST, Feolo M, Skol A, Zhang H, Zeng C, Zhao H, Matsuda I, Fukushima Y, Macer DR, Suda E, Rotimi CN, Adebamowo CA, Ajayi I, Aniagwu T, Marshall PA, Nkwodimmah C, Royal CD, Leppert MF, Dixon M, Peiffer A, Qiu R, Kent A, Kato K, Niikawa N, Adewole IF, Knoppers BM, Foster MW, Clayton EW, Watkin J, Gibbs RA, Belmont JW, Muzny D, Nazareth L, Sodergren E, Weinstock GM, Wheeler DA, Yakub I, Gabriel SB, Onofrio RC, Richter DJ, Ziaugra L, Birren BW, Daly MJ, Altshuler D, Wilson RK, Fulton LL, Rogers J, Burton J, Carter NP, Clee CM, Griffiths M, Jones MC, McLay K, Plumb RW, Ross MT, Sims SK, Willey DL, Chen Z, Han H, Kang L, Godbout M, Wallenburg JC, L'Archevêque P, Bellemare G, Saeki K, Wang H, An D, Fu H, Li Q, Wang Z, Wang R, Holden AL, Brooks LD, McEwen JE, Guyer MS, Wang VO, Peterson JL, Shi M, Spiegel J, Sung LM, Zacharia LF, Collins FS, Kennedy K, Jamieson R, Stewart J (2007) A second generation human haplotype map of over 3.1 million SNPs. Nature

449: 851-861

Jefreys H (1961) Theory of Probability (3rd edn). Oxford University Press, Oxford

Ji Y, Eichler EE, Schwartz S, Nicholls RD (2000) Structure of chromosomal duplicons and their role in mediating human genomic disorders. Genome Res 10: 597-610

Jobling MA, Hurles ME, Tyler-Smith C (2004) Human evolutionary genetics: Origins, peoples and disease. Garland Publishing, India

John S, Shephard N, Liu G, Zeggini E, Cao M, Chen W, Vasavda N, Mills T, Barton A, Hinks A, Eyre S, Jones KW, Ollier W, Silman A, Gibson N, Worthington J, Kennedy GC (2004) Whole-genome scan, in a complex disease, using 11,245 single-nucleotide-polymorphisms: comparison with microsatellites. Am J Hum Genet 75: 54-64

Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D (1992) Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science 258: 818-821

Kehrer-Sawatzki H, Cooper DN (2007) Understanding the recent evolution of the human genome: insights from human-chimpanzee genome comparisons. Hum Mutat 28: 99-130

Kelley JL, Madeoy J, Calhoun JC, Swanson W, Akey JM (2006) Genomic signatures of positive selection in humans and the limits of outlier approaches. Genome Res 16: 980-989

Kennedy GC, Matsuzaki H, Dong S, Liu WM, Huang J, Liu G, Su X, Cao M, Chen W, Zhang J, Liu W, Yang G, Di X, Ryder T, He Z, Surti U, Phillips MS, Boyce-Jacino MT, Fodor SP, Jones KW. (2003) Large-scale genotyping of complex DNA. Nat Biotechnol 21:1233-1237

Kimchi-Sarfaty C, Oh JM, Kim IW, Sauna ZE, Calcagno AM, Ambudkar SV, Gottesman MM (2007) A "silent" polymorphism in the *MDR1* gene changes substrate specificity. Science 315: 525-528

Kimura M (1983) The neutral theory of molecular evolution. Cambridge University Press, Cambridge

Kimura M, Ota T (1973) The age of a neutral mutant persisting in a finite population. Genetics 75: 199-212

Kimura R, Fujimoto A, Tokunaga K, Ohashi J (2007) A practical genome scan for population-specific strong selective sweeps that have reached fixation. PLoS ONE 2: e286

Kiyosawa H, Lensch MW, Chance PF (1995) Analysis of the CMT1A-REP repeat: mapping crossover breakpoints in CMT1A and HNPP. Hum Mol Genet 4: 2327-2334

Klein J, Takahata N (2002) Where do we come from? The molecular evidence for human descent. Springer-Verlag Berlin Heidelberg, Heidelberg

Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, Henning AK, SanGiovanni JP, Mane SM, Mayne ST, Bracken MB, Ferris FL, Ott J, Barnstable C, Hoh J (2005) Complement factor H polymorphism in age-related macular degeneration. Science 308: 385-389

Kleinjan DA, van Heyningen V (2005) Long-range control of gene expression: emerging mechanisms and disruption in disease. Am J Hum Genet 76: 8-32

Kruglyak L (1997) The use of a genetic map of biallelic markers in linkage studies. Nat Genet 17: 21-24

Kruglyak L (1997) What is significant in whole-genome linkage disequilibrium studies? Am J Hum Genet 61: 810-812

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 58: 1347-1363

Kruglyak L, Lander ES (1995) Complete multipoint sib-pair analysis of qualitative and quantitative traits. Am J Hum Genet 57: 439-454

Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP (2001) A forkhead-domain gene is mutated in a severe speech and language disorder. Nature 413: 519–523

Lamason RL, Mohideen MA, Mest JR, Wong AC, Norton HL, Aros MC, Jurynec MJ, Mao X, Humphreville VR, Humbert JE, Sinha S, Moore JL, Jagadeeswaran P, Zhao W, Ning G, Makalowska I, McKeigue PM, O'donnell D, Kittles R, Parra EJ, Mangini NJ, Grunwald DJ, Shriver MD, Canfield VA, Cheng KC (2005) SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. Science 310: 1782-1786

Lander ES (1996) The new genomics: global views of biology. Science 274: 536–539

Lander ES, Green P (1987) Construction of multilocus genetic linkage maps in humans. Proc Nat Acad Sci U S A 84: 2363-2367

Linardopoulou EV, Williams EM, Fan Y, Friedman C, Young JM, Trask BJ (2005) Human subtelomeres are hot spots of interchromosomal recombination and segmental duplication. Nature 437: 94-100

Liu PY, Zhang YY, Lu Y, Long JR, Shen H, Zhao LJ, Xu FH, Xiao P, Xiong DH, Liu YJ, Recker RR, Deng HW (2005) A survey of haplotype variants at several disease candidate genes: the importance of rare variants for complex diseases. J Med Genet 42: 221-227

Lucek PR, Ott J (1997) Neural network analysis of complex traits. Genet Epidemiol 14: 1101-1106

Lucito R, West J, Reiner A, Alexander J, Esposito D, Mishra B, Powers S, Norton L, Wigler M (2000) Detecting gene copy number fluctuations in tumor cells by microarray analysis of genomic representations. Genome Res 10: 1726-1736

Lunetta KL, Hayward LB, Segal J, van Eerdewegh P (2004) Screening large-scale association study data: exploiting interactions using random forests. BMC Genet 5:32

Lunn DJ, Whittaker JC, Best N (2006) A Bayesian toolkit for genetic association studies. Genet Epidemiol 30: 231-247

Maraganore DM, de Andrade M, Lesnick TG, Strain KJ, Farrer MJ, Rocca WA, Pant PV, Frazer KA, Cox DR, Ballinger DG (2005) High-resolution whole-genome association study of Parkinson disease. Am J Hum Genet 77: 685-693

Marchini J, Howie B, Myers S, McVean G, Donnelly P (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 39: 906-913

Marjoram P, Tavaré S (2006) Modern computational approaches for analysing molecular genetic variation data. Nat Rev Genet 7: 759-770

McCarroll SA, Hadnott TN, Perry GH, Sabeti PC, Zody MC, Barrett JC, Dallaire S, Gabriel SB, Lee C, Daly MJ, Altshuler DM (2006) Common deletion polymorphisms in the human genome. Nat Genet 38: 86-92

Moore JH (2003) The ubiquitous nature of epistasis in determining susceptibility to common human diseases. Hum Hered 56: 73-82

Moran MD (2003) Arguments for rejecting the sequential Bonferroni in ecological studies. Oikos 100: 403-405

Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, Spielman RS, Cheung VG (2004) Genetic analysis of genome-wide variation in human gene expression. Nature 430: 743-747

Morris AP (2006) A flexible Bayesian framework for modeling haplotype association with disease, allowing for dominance effects of the underlying causative variants. Am J Hum Genet 79: 679-694

Motsinger AA, Lee SL, Mellick G, Ritchie MD (2006) GPNN: power studies and applications of a neural network method for detecting gene-gene interactions in studies of human disease. BMC Bioinformatics 7: 39

Motulsky H 数学いらずの医科統計学 メディカル・サイエンスインターナショナル

Nachman MW, Crowell SL (2000) Estimate of the mutation rate per nucleotide in humans. Genetics 156: 297-304

Nakajima T, Wooding S, Sakagami T, Emi M, Tokunaga K, Tamiya G, Ishigami T, Umemura S, Munkhbat B, Jin F, Guan-Jun J, Hayasaka I, Ishida T, Saitou N, Pavelka K, Lalouel JM, Jorde LB, Inoue I (2004) Natural selection and population history in the human angiotensinogen gene (*AGT*): 736 complete *AGT* sequences in chromosomes from around the world. Am J Hum Genet 74: 898-916

Neel JV (1962) Diabetes mellitus: a "thrifty" genotype rendered detrimental by "progress" ? Am J Hum Genet 14: 353-362

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Nelson MR, Kardia SL, Ferrell RE, Sing CF (2001) A combinatorial partitioning method to identify multilocus genotypic partitions that predict quantitative trait variation. Genome Res 11: 458-470

Neuhäuser M (2004) Testing whether any of the significant tests within a table are indeed significant. Oikos 106: 409-410

Newman T, Trask BJ (2003) Complex evolution of 7E olfactory receptor genes in segmental duplications. Genome Res 13: 781-93

Ng PC, Henikoff S (2002) Accounting for human polymorphisms predicted to affect protein function. Genome Res 12: 436-446

Nielsen R, Bustamante C, Clark AG, Glanowski S, Sackton TB, Hubisz MJ, Fledel-Alon A, Tanenbaum DM, Civello D, White TJ, J Sninsky J, Adams MD, Cargill M (2005) A scan for positively selected genes in the genomes of humans and chimpanzees. PLoS Biol 3: e170

Nielsen R, Hellmann I, Hubisz M, Bustamante C, Clark AG (2007) Recent and ongoing selection in the human genome. Nat Rev Genet 8: 857-868

Nyholt DR (2004) A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. Am J Hum Genet 74: 765-769

Ohta T (2002) Near-neutrality in evolution of genes and gene regulation. Proc Natl Acad Sci USA 99: 16134-16137

Onda H, Kasuya H, Yoneyama T, Takakura K, Hori T, Takeda J, Nakajima T, Inoue I (2001) Genomewide-linkage and haplotype-association studies map intracranial aneurysm to chromosome 7q11. Am J Hum Genet 69: 804-819

Ozaki K, Ohnishi Y, Iida A, Sekine A, Yamada R, Tsunoda T, Sato H, Sato H, Hori M, Nakamura Y, Tanaka T (2002) Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. Nat Genet 32: 650-654

Palmer LJ, Cardon LR (2005) Shaking the tree: mapping complex disease genes with linkage disequilibrium. Lancet 366: 1223-1234

Parra I, Windle B (1993) High resolution visual mapping of stretched DNA by fluorescent hybridization. Nat Genet 5: 17-21

Peiffer DA, Le JM, Steemers FJ, Chang W, Jenniges T, Garcia F, Haden K, Li J, Shaw CA, Belmont J, et

al. (2006) High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. Genome Res 16: 1136-1148

Platzer P, Upender MB, Wilson K, Willis J, Lutterbaugh J, Nosrati A, Willson JK, Mack D, Ried T, Markowitz S (2002) Silence of chromosomal amplifications in colon cancer. Cancer Res 62: 1134-1138

Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Pergamenschikov A, Williams CF, Jeffrey SS, Botstein D, Brown PO (1999) Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat Genet 23: 41-46

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38: 904-909

Purcell S, Cherny SS, Sham PC (2003) Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 19: 149-150

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559-575

Rauch A, Rüschendorf F, Huang J, Trautmann U, Becker C, Thiel C, Jones KW, Reis A, Nürnberg P (2004) Molecular karyotyping using an SNP array for genomewide genotyping. J Med Genet 41: 916-922

Read AP, Strachan T (1997) Human Molecular Genetics -1st edition, 466-468. BIOS Scientific Publishers

Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, et al. (2006) Global variation in copy number in the human genome. Nature 444: 444-454

Reich DE, Lander ES (2001) On the allelic spectrum of human disease. Trends Genet 17: 502–510

Rice WR (1989) Analyzing tables of statistical tests. Evolution 43: 223-225

Riley B, Williamson M, Collier D, Wilkie H, Makoff A (2002) A 3-Mb map of a large segmental duplication overlapping the alpha□7-nicotinic acetylcholine receptor gene (CHRNA7) at human 15q13-q14. Genomics 79: 197-209

Risch N (1990a) Linkage strategies for genetically complex traits. I. Multilocus models. Am J Hum Genet 46: 222-228

Risch N (1990b) Linkage strategies for genetically complex traits. II. The power of affected relative pairs. Am J Hum Genet 46: 229-241

Risch N (1990c) Linkage strategies for genetically complex traits. III. The effect of marker polymorphism on analysis of affected relative pairs. Am J Hum Genet 46: 242-253

Risch NJ (2000) Searching for genetic determinants in the new millennium. Nature 405: 847-856

Ritchie MD, Hahn LW, Roodi N, Bailey LR, Dupont WD, Parl FF, Moore JH (2001) Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. Am J Hum Genet 69: 138-147

Ritchie MD, White BC, Parker JS, Hahn LW, Moore JH (2003) Optimization of neural network architecture using genetic programming improves detection and modeling of gene-gene interactions in studies of human diseases. BMC Bioinformatics 4:28

Rockman MV, Hahn MW, Soranzo N, Goldstein DB, Wray GA (2003) Positive selection on a human-specific transcription factor binding site regulating *IL4* expression. Curr Biol 13: 2118-2123

Rockman MV, Hahn MW, Soranzo N, Loisel DA, Goldstein DB, Wray GA (2004) Positive selection on *MMP3* regulation has shaped heart disease risk. Curr Biol 14: 1531-1539

Rudan I, Smolej-Narancic N, Campbell H, Carothers A, Wright A, Janicijevic B, Rudan P (2003) Inbreeding and the genetic complexity of human hypertension. Genetics 163: 1011-1021

Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, Schaffner SF, Lander ES; International HapMap Consortium, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, Parkin M, Roy J, Stahl E, Winchester E, Ziaugra L, Altshuler D, Shen Y, Yao Z, Huang W, Chu X, He Y, Jin L, Liu Y, Shen Y, Sun W, Wang H, Wang Y, Wang Y, Xiong X, Xu L, Waye MM, Tsui SK, Xue H, Wong JT, Galver LM, Fan JB, Gunderson K, Murray SS, Oliphant AR, Chee MS, Montpetit A, Chagnon F, Ferretti V, Leboeuf M, Olivier JF, Phillips MS, Roumy S, Sallée C, Verner A, Hudson TJ, Kwok PY, Cai D, Koboldt DC, Miller RD, Pawlikowska L, Taillon-Miller P, Xiao M, Tsui LC, Mak W, Song YQ, Tam PK, Nakamura Y, Kawaguchi T, Kitamoto T, Morizono T, Nagashima A, Ohnishi Y, Sekine A, Tanaka T, Tsunoda T, Deloukas P, Bird CP, Delgado M, Dermitzakis ET, Gwilliam R, Hunt S, Morrison J, Powell D, Stranger BE, Whittaker P, Bentley DR, Daly MJ, de Bakker PI, Barrett J, Chretien YR, Maller J, McCarroll S, Patterson N, Pe'er I, Price A, Purcell S, Richter DJ, Sabeti P, Saxena R, Schaffner SF, Sham PC, Varilly P, Altshuler D, Stein LD, Krishnan L, Smith AV, Tello-Ruiz MK, Thorisson GA, Chakravarti A, Chen PE, Cutler DJ, Kashuk CS, Lin S, Abecasis GR, Guan W, Li Y, Munro HM, Qin ZS, Thomas DJ, McVean G, Auton A, Bottolo L, Cardin N, Eyheramendy S, Freeman C, Marchini J, Myers S, Spencer C, Stephens M, Donnelly P, Cardon LR, Clarke G, Evans DM, Morris AP, Weir BS, Tsunoda T, Johnson TA, Mullikin JC, Sherry ST, Feolo M, Skol A, Zhang H, Zeng C, Zhao H, Matsuda I, Fukushima Y, Macer DR, Suda E, Rotimi CN, Adebamowo CA, Ajayi I, Aniagwu T, Marshall PA, Nkwodimmah C, Royal CD, Leppert MF, Dixon M, Peiffer A, Qiu R, Kent A, Kato K, Niikawa N, Adewole IF, Knoppers BM, Foster MW, Clayton EW, Watkin J, Gibbs RA, Belmont JW, Muzny D, Nazareth L, Sodergren E, Weinstock GM, Wheeler DA, Yakub I, Gabriel SB, Onofrio RC, Richter DJ, Ziaugra L, Birren BW, Daly MJ, Altshuler D, Wilson RK, Fulton LL, Rogers J, Burton J, Carter NP, Clee CM, Griffiths M, Jones MC, McLay K, Plumb RW, Ross MT, Sims SK, Willey DL, Chen Z, Han H, Kang L, Godbout M, Wallenburg JC, L'Archevêque P, Bellemare G, Saeki K, Wang H, An D, Fu H, Li Q, Wang Z, Wang R, Holden AL, Brooks LD, McEwen JE, Guyer MS, Wang VO, Peterson JL, Shi M, Spiegel J, Sung LM, Zacharia LF, Collins FS, Kennedy K, Jamieson R, Stewart J (2007) Genome-wide detection and characterization of positive selection in human populations. Nature 449: 913-918

Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, Sherry S, Mullikin JC, Mortimore BJ, Willey DL, Hunt SE, Cole CG, Coggill PC, Rice CM, Ning Z, Rogers J, Bentley DR, Kwok PY, Mardis ER, Yeh RT, Schultz B, Cook L, Davenport R, Dante M, Fulton L, Hillier L, Waterston RH, McPherson JD, Gilman B, Schaffner S, Van Etten WJ, Reich D, Higgins J, Daly MJ, Blumenstiel B, Baldwin J, Stange-Thomann N, Zody MC, Linton L, Lander ES, Altshuler D; International SNP Map Working Group (2001) A map of human

genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 409: 928-933

Sakagami T, Witherspoon DJ, Nakajima T, Jinnai N, Wooding S, Jorde LB, Hasegawa T, Suzuki E, Gejyo F, Inoue I (2004) Local adaptation and population differentiation at the interleukin 13 and interleukin 4 loci. Genes Immun 5: 389-397

Samonte RV, Eichler EE (2002) Segmental duplications and the evolution of the primate genome. Nat Rev Genet 3: 65-72

Sasieni PD (1997) From genotypes to genes: doubling the sample size. Biometrics 53: 1253-1261

Sawcer SJ, Maranian M, Singlehurst S, Yeo T, Compston A, Daly MJ, De Jager PL, Gabriel S, Hafler DA, Ivinson AJ, Lander ES, Rioux JD, Walsh E, Gregory SG, Schmidt S, Pericak-Vance MA, Barcellos L, Hauser SL, Oksenberg JR, Kenealy SJ, Haines JL (2004) Enhancing linkage analysis of complex disorders: an evaluation of high-density genotyping. Hum Mol Genet 13: 1943-1949

Schaid DJ, Guenther JC, Christensen GB, Hebbring S, Rosenow C, Hilker CA, McDonnell SK, Cunningham JM, Slager SL, Blute ML, Thibodeau SN (2004) Comparison of microsatellites versus single-nucleotide-polymorphisms in a genome linkage screen for prostate cancer-susceptibility loci. Am J Hum Genet 75: 948-965

Scott WK, Pericak-Vance MA, Haines JL (1997) Genetic analysis of complex diseases. Science 275: 1327; author reply 1329-1330

Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Månér S, Massa H, Walker M, Chi M, *et al.* (2004) Large-scale copy number polymorphism in the human genome. Science 305: 525-528

Sellick GS, Webb EL, Allinson R, Matutes E, Dyer MJ, Jonsson V, Langerak AW, Mauro FR, Fuller S, Wiley J, Lyttelton M, Callea V, Yuille M, Catovsky D, Houlston RS (2005) A high-density SNP genomewide linkage scan for chronic lymphocytic leukemia-susceptibility loci. Am J Hum Genet 77: 420-429

Servin B, Stephens M (2007) Imputation-based analysis of association studies: candidate regions and quantitative traits. PLoS Genet 3: e114

Sharp AJ, Cheng Z, Eichler EE (2006) Structural variation of the human genome. Annu Rev Genomics Hum Genet 7: 407-442

Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, Pertz LM, Clark RA, Schwartz S, Segraves R, et al. (2005) Segmental duplications and copy-number variation in the human genome. Am J Hum Genet 77: 78-88

She X, Horvath JE, Jiang Z, Liu G, Furey TS, Christ L, Clark R, Graves T, Gulden CL, Alkan C, et al. (2004) The structure and evolution of centromeric transition regions within the human genome. Nature 430: 857-864

Šidák Z (1967) Rectangular confidence regions for the means of multivariate normal distributions. J Am Stat Assoc 78: 626-633

Sillanpää MJ, Bhattacharjee M (2005) Bayesian association-based fine mapping in small chromosomal segments. Genetics 169: 427-439

Skol AD, Scott LJ, Abecasis GR, Boehnke M (2007) Optimal designs for two-stage genome-wide association studies. Genet Epidemiol 31:776-788

Spielman RS, Bastone LA, Burdick JT, Morley M, Ewens WJ, Cheung VG (2007) Common genetic

variants account for differences in gene expression among ethnic groups. Nat Genet 39: 226-231

Storey JD, Taylor JE, Siegmund D (2004) Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach. J Roy Stat Soc B 66: 187-205

Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. Proc Nat Acad Sci U S A 100: 9440-9445

Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, Redon R, Bird CP, de Grassi A, Lee C, et al. (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science 315: 848-853

Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, Redon R, Bird CP, de Grassi A, Lee C, *et al.* (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science 315: 848-853

Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, Beazley C, Ingle CE, Dunning M, Flicek P, Koller D, Montgomery S, Tavaré S, Deloukas P, Dermitzakis ET (2007) Population genomics of human gene expression. Nat Genet 39: 1217-1224

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585-595

Tanaka T, Ikari K, Furushima K, Okada A, Tanaka H, Furukawa K, Yoshida K, Ikeda T, Ikegawa S, Hunt SC, Takeda J, Toh S, Harata S, Nakajima T, Inoue I (2003) Genomewide linkage and linkage disequilibrium analyses identify *COL6A1*, on chromosome 21, as the locus for ossification of the posterior longitudinal ligament of the spine. Am J Hum Genet 73: 812-822

The International HapMap Consortium (2003) The International HapMap Project. Nature 426: 789-796

Thomas A, Camp NJ (2004) Graphical modeling of the joint distribution of alleles at associated loci. Am J Hum Genet 74: 1088-1101

Thompson EE, Kuttab-Boulos H, Witonsky D, Yang L, Roe BA, Di Rienzo A (2004) CYP3A variation and the evolution of salt-sensitivity variants. Am J Hum Genet 75: 1059-1069

Tomita Y, Tomida S, Hasegawa Y, Suzuki Y, Shirakawa T, Kobayashi T, Honda H (2004) Artificial neural network approach for selection of susceptible single nucleotide polymorphisms and construction of prediction model on childhood allergic asthma. BMC Bioinformatics 5:120

Tsai CT, Lai LP, Lin JL, Chiang FT, Hwang JJ, Ritchie MD, Moore JH, Hsu KL, Tseng CD, Liau CS, Tseng YZ (2004) Renin-angiotensin system gene polymorphisms and atrial fibrillation. Circulation 109:1640-1646

Tuzun E, Sharp AJ, Bailey JA, Kaul R, Morrison VA, Pertz LM, Haugen E, Hayden H, Albertson D, Pinkel D, et al. (2005) Fine-scale structural variation of the human genome. Nat Genet 37: 727-732

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M,

Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferriera S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigó R, Campbell MJ, Sjolander KV, Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A, Zhu X (2001) The sequence of the human genome. Science 291: 1304-1351

Verzilli CJ, Stallard N, Whittaker JC (2006) Bayesian graphical models for genomewide association studies. Am J Hum Genet 79: 100-112

Voight BF, Kudaravalli S, Wen X, Pritchard JK (2006) A map of recent positive selection in the human genome. PLoS Biol 4: e72

Webb EL, Sellick GS, Houlston RS (2005) SNPLINK: multipoint linkage analysis of densely distributed SNP data incorporating automated linkage disequilibrium removal. Bioinformatics 21: 3060-3061

Williamson SH, Hubisz MJ, Clark AG, Payseur BA, Bustamante CD, Nielsen R (2007) Localizing recent adaptive evolution in the human genome. PLoS Genet 3: e90

Wright A, Charlesworth B, Rudan I, Carothers A, Campbell H (2003) A polygenic basis for late-onset disease. Trends Genet 19: 97-106

Yoshiura K, Kinoshita A, Ishida T, Ninokata A, Ishikawa T, Kaname T, Bannai M, Tokunaga K, Sonoda S, Komaki R, Ihara M, Saenko VA, Alipov GK, Sekine I, Komatsu K, Takahashi H, Nakashima M, Sosonkina N, Mapendano CK, Ghadami M, Nomura M, Liang DS, Miwa N, Kim DK, Garidkhuu A, Natsume N, Ohta T, Tomita H, Kaneko A, Kikuchi M, Russomando G, Hirayama K, Ishibashi M, Takahashi A, Saitou N, Murray JC, Saito S, Nakamura Y, Niikawa N (2006) A SNP in the *ABCC11* gene is the determinant of human earwax type. Nat Genet 38: 324-330

Zaykin DV, Zhivotovsky LA, Westfall PH, Weir BS (2002) Truncated product method for combining P-values. Genet Epidemiol 22: 170-185

Zhou G, Zhai Y, Dong X, Zhang X, He F, Zhou K, Zhu Y, Wei H, Yao Z, Zhong S, Shen Y, Qiang B, He F (2004) Haplotype structure and evidence for positive selection at the human *IL13* locus. Mol Biol Evol 21: 29-35

Zipfel PF, Edey M, Heinen S, Józsi M, Richter H, Misselwitz J, Hoppe B, Routledge D, Strain L, Hughes AE, et al. (2007) Deletion of complement factor H-related genes CFHR1 and CFHR3 is associated with atypical hemolytic uremic syndrome. PLoS Genet 3: e41

堀内賢太郎、松田眞一 (2006) FDR を制御する多重比較法の性能評価. アカデミア（南山大学紀要）数理情報編 6: 17-30

□ 発行者：（株）ダイナコム
千葉県茂原市茂原 643
TEL：0475-25-8282　FAX：0475-25-8301
E-mail：info@dynacom.co.jp　URL：http://www.dynacom.co.jp/

□ 発行日：2008 年〇月〇日

Hum Genet (2008) 123:655–660
DOI 10.1007/s00439-008-0512-x

SHORT REPORT

Genome-wide association analysis with selective genotyping identifies candidate loci for adult height at 8q21.13 and 15q22.33-q23 in Mongolians

Tetsuaki Kimura · Terukazu Kobayashi · Batmunkh Munkhbat · Ganjuur Oyungerel · Tsolmon Bilegtsaikhan · Damdinsuren Anar · Jamiyansuren Jambaldorj · Sodnom Munkhsaikhan · Namid Munkhtuvshin · Hideki Hayashi · Akira Oka · Ituro Inoue · Hidetoshi Inoko

Received: 6 February 2008 / Accepted: 12 May 2008 / Published online: 20 May 2008

Abstract We performed a genome-wide association study with 23,465 microsatellite markers to identify genes related to adult height. Selective genotyping was applied to extremely tall and extremely short individuals from the Khalkh-Mongolian population. Two loci, 8q21.13 and 15q22.33, which showed the strongest association with microsatellites were subjected to further analyses of SNPs in 782 tall and 773 short individuals. The most significant association was observed with SNP rs2220456 at 8q21.13 ($P = 0.000016$). In the LD block at 15q22.32, SNP rs8038652 located in intron 1 of *IQCH* was strongly associated ($P = 0.0003$), especially the AA genotype of the SNP under a recessive model was strongly associated with adult height ($P = 0.000046$).

Introduction

Adult height is an explicit quantitative phenotype and stable once a person has grown and is easily measured. Adult height is largely controlled by genetic factors, with heritability ranging from 75 to 90% in various populations (Carmichael and McGue 1995; Silventoinen et al. 2003).

Adult height usually follows normal distribution in a given population and sex, the phenotype representing a typical polygenic model of a human quantitative trait influenced by multiple genes each with small effects. Numerous linkage studies have attempted to identify loci underlying adult height variation. Thompson et al. first reported a locus for adult height on chromosome 20 in Pima Indians (Thompson et al. 1995). Several other groups reported evidence of linkage with adult height in Europeans (Beck et al. 2003; Dempfle et al. 2006; Deng et al. 2002; Ellis et al. 2007; Geller et al. 2003; Hirschhorn et al. 2001; Liu et al. 2006; Mukhopadhyay et al. 2003; Mukhopadhyay and Weeks 2003; Perola et al. 2001; Perola et al. 2007; Sammalisto et al. 2005; Willemsen et al. 2004; Wiltshire et al. 2002; Xu et al. 2002). Wu et al. (2003) found evidence of linkage in four ethnic groups; White, Black, Mexican American, and Asian. Most recently, Visscher reported a large-scale linkage study with 11,214 sibling pairs showing that additive genetic variance is spread across multiple chromosomes, with no evidence of large between-chromosome epistatic effects (Visscher et al. 2007). While multiple evidence of linkage of adult height has been identified in several populations in these studies, common loci are not evident.

Since linkage study has limited power to detect genes of modest effect, especially where there is genetic heterogeneity, we applied association study with a sufficient number of subjects to identify genes with a small impact on the phenotype (Risch and Merikangas 1996). Recently, some groups reported genes associated with adult height variation using data from genome-wide association study (Gudbjartsson et al. 2008; Lettre et al. 2008; Sanna et al. 2008; Weedon et al. 2007, 2008). In the present study, we report results of a genome-wide association study of adult height with 1,555

T. Kimura · T. Kobayashi · B. Munkhbat · H. Hayashi · A. Oka · I. Inoue · H. Inoko (✉)
Department of Molecular Life Science and Molecular Medicine, Tokai University School of Medicine, 143 Shimokasuya, Isehara, Kanagawa 259-1193, Japan
e-mail: hinoko@is.icc.u-tokai.ac.jp

T. Kimura
e-mail: tekimura@is.icc.u-tokai.ac.jp

B. Munkhbat · G. Oyungerel · T. Bilegtsaikhan · D. Anar · J. Jambaldorj · S. Munkhsaikhan · N. Munkhtuvshin
Central Scientific Research Laboratory, National Institute of Medicine, Ulaanbaatar, Mongolia

individuals from the Khalkh population of Mongolia using 23,465 microsatellite markers. The Khalkh population has a relatively close genetic affinity to populations of the northern part of East Asia showing a relatively homogeneous genetic background, which provides an advantage to study complex phenotype (Katoh et al. 2002, 2005; Nakajima et al. 2004). We applied a selective genotyping strategy in which individuals with trait values deviating from the population mean were preferentially recruited to identify genetic variations underlying quantitative traits with improved power (Arking et al. 2006; Lander and Botstein 1989).

Material and methods

Study subject selection

Adult height for both male and female shows normal distribution with average height and standard deviation being 164.76 ± 5.74 cm for male and 153.76 ± 5.04 cm for female according to epidemiological and anthropometric surveys on adult height among Khalkh-Mongolians (Otgon et al. 2002; personal communication L. Namsrainaidan). A total of 1,555 unrelated individuals of Khalkh-Mongolian origin from the region of Ulaanbaatar, Mongolia participated in the current study. The selection of individuals from the general population was >95th percentile for the tall group corresponding to >173.9 cm and <5th percentile for the short group corresponding to <155.6 cm for male and 161.8 and 145.7 cm, respectively, for female. The subjects in the short group were over 18 years of age and those in the tall group were over 15 years of age at the time of examination. Individuals with medical conditions affecting adult height, such as dwarfism, gigantism, and acromegaly were excluded. The study was approved by the Institutional Review Board of Tokai University and the Medical Research Ethics Committee of the National Institute of Medicine and the Ethics Committee, Ministry of Health, Mongolia. The participants gave written, informed consent.

DNA pool construction and microsatellite genotyping

The pooled DNA method for microsatellite typing was performed according to the protocol of Collins et al. (2000) with a slight modification (Oka et al. 2003). DNA was extracted using QIAamp DNA blood kit (QIAGEN) under the standardized protocol to prevent variation of DNA quality. The DNA concentration was precisely measured using the PicoGreen fluorescence assay (Molecular Probes) as previously described (Tamiya et al. 2005; Kawashima et al. 2006). For the first round screening, four DNA pools were prepared. The first set for association study was DNA pools of 125 male-tall, 125 male-short, 125 female-tall, and 125 female-short samples, respectively. A second set was also grouped from another 125 male- and female-tall samples and 125 male- and female-short samples, respectively. In the first round screening, 23,465 microsatellite markers were used. Among them, showing statistical significance of $P < 0.05$ were subjected to the second round screening.

All microsatellite markers and methods for microsatellite genotyping used in this study are described by Tamiya et al. (2005). PCR on pooled DNAs was performed in a 20-μl reaction mixture containing 48 ng of pooled DNA, 0.5 U of AmpliTaq DNA polymerase, 1× reaction buffer with 1.5 mM $MgCl_2$ provided by the manufacturer (Applied Biosystems), 5 μM of each primer, and 0.25 mM of each deoxyriobonucleotide triphosphate (dNTP) in 96-well plates. The PCR amplification was performed on the GeneAmp PCR System 9700 (Applied Biosystems) with the following conditions: 96°C for 5 min (hot start), 57°C for 1 min, and 72°C for 1 min followed by 40 cycles of 96°C for 45 s, 57°C for 45 s and 72°C for 1 min. For the microsatellite genotyping of individual samples, PCR was performed in a 20 μl reaction containing 1 ng of genomic DNA. The amplification condition was the same as described above. The pooled and individual microsatellite genotyping procedures after PCR amplification were carried out according to standard protocols using ABI3730 DNA analyzer (Applied Biosystems). Peak positions and heights were automatically extracted by the PickPeak and MultiPeaks programs.

SNP genotyping

The SNPs in candidate regions were selected from the SNP database of Applied Biosystems (http://www2.appliedbiosystems.com/) using SNPbrowser software 3.5 (Applied Biosystems). The SNPs were genotyped by TaqMan assays. The TaqMan assays were carried out using the standard protocols for the ABI PRISM 7900HT Sequence Detection System using a 384-well block module and automation accessory (Applied Biosystems).

Statistical analysis

In pooled DNA typing, adult height associations with microsatellites were assessed by Fisher's exact test, with the use of 2×2 contingency tables for each allele. Allele frequencies in pooled DNA typing were estimated from the height of peaks: each allele frequency was determined by dividing the height of each allele by the summed height of all alleles. In individual genotyping, significance was evaluated by Fisher's exact test, with the use of 2×2 contingency tables for each allele.

For SNPs genotyping, adult height associations were assessed using chi-square test (Haploview 4.0 software [http://www.broad.mit.edu/mpg/haploview/]). Since multi-step

analysis was used, the nominal P values were corrected with 1,000,000 iterated permutations for all 82 SNPs. Significance level was set at .05 throughout the study.

To assess the extent of pair-wise linkage disequilibrium between SNPs, standard definition of D' and r^2 were calculated using Haploview software. D' and r^2 were calculated only for polymorphisms with minor-allele frequency (MAF) > 5%. LD blocks were then defined with pair-wise LD with $D' > 0.9$.

Results and discussion

Genome-wide association study

We performed a genome-wide association study with 23,465 microsatellite markers for detection of loci controlling adult height using the selective genotyping method. To reduce cost and technical burden of genome-wide genotyping, the pooled DNA method was applied, as previously described (Collins et al. 2000, Tamiya et al. 2005). Association results with the pooled DNA method and following re-genotyping of individual DNAs using the same set of 1,000 screened individuals, 23 markers showed significant differences by Fisher's exact test (Table 1). These markers were subjected to correction of multiple tests with the number of alleles, and nine microsatellites remained significant.

Visscher et al. reported that at least six chromosomes (3, 4, 8, 15, 17, and 18) were responsible for height variation in the European population (Visscher et al. 2007). We also detected significant association in those chromosomes, except chromosome 18. In addition, five regions overlapped at least partially with loci previously reported by

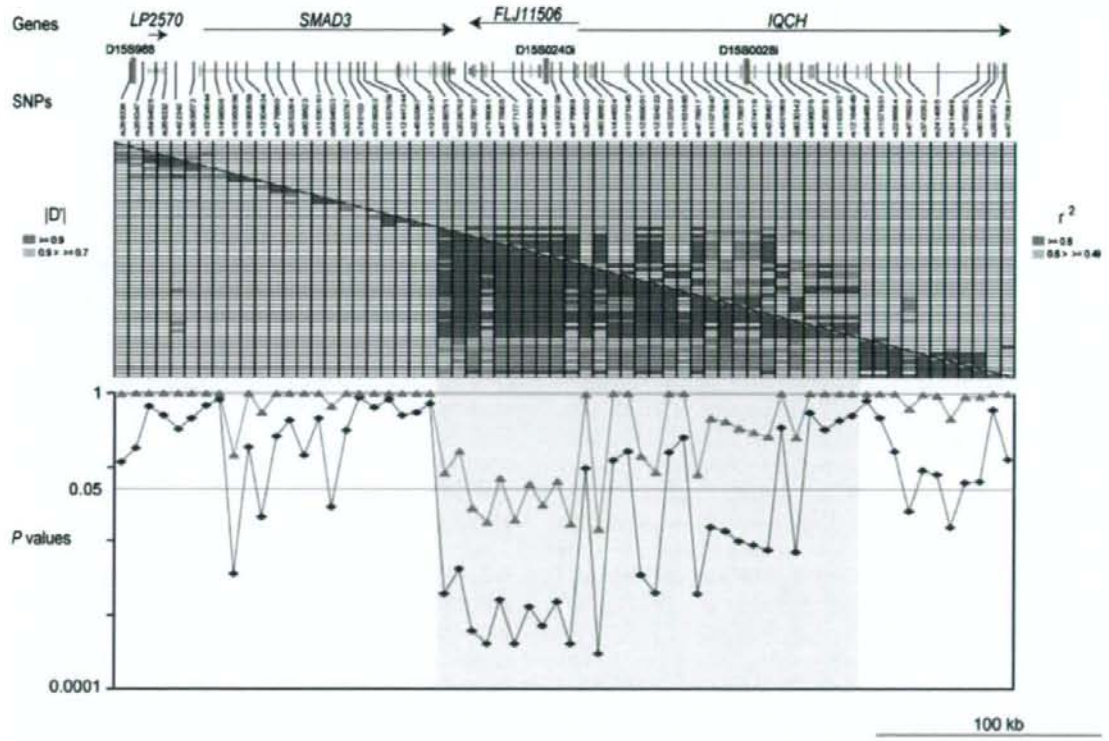

Fig. 1 SNP allelic association within 15q22.33-q23. SNP association analysis. The *blue line* shows P values calculated by chi square test. The red line shows P values generated after 1,000,000 iterated permutations. *Yellow background* indicates the 188 kb LDB. In the Mongolian population, we investigated the 188 kb LD block constructed by these significant markers spanning from intron 6 of *SMAD3* (rs2289791) to intron 10 of *IQCH* (rs12164949). The LD block contained the MH2 domain and the 3′ UTR of *SMAD3*, the entire coding sequence of *FLJ11506*, and the IQ domain of *IQCH*. SNP rs8038652 located in intron 1 of *IQCH* was most strongly associated ($P = 0.0003$, $Pc = 0.015$) with adult height. SNP rs227860 located in the 3′ UTR of *SMAD3* also was associated ($P = 0.0006$, $Pc = 0.028$). SNP rs7166081 ($P = 0.0004$, $Pc = 0.018$) was in an intergenic region between *SMAD3* and *FLJ11506*. Three remaining SNPs, rs4776908 ($P = 0.0004$, $Pc = 0.017$), rs877177 ($P = 0.0004$, $Pc = 0.020$), and rs4776906 ($P = 0.0007$, $Pc = 0.030$) located in intron 1, 5, and 5 of *FLJ11506*, respectively, were also associated

Table 1 Twenty-three positive microsatelllite markers from individual genotyping

No.	Markers	Locus	Number of allele	Allele frequencies		Fisher's exact *P* value		Odds ratio	95% CI
				Tall	Short	Exact2×2	Corrected*		
1.	D1S2660	1q36.32	12	0.093	0.061	0.0090	0.11	1.58	1.13–2.21
2.	D2S0257i	2q33.1	16	0.092	0.132	0.0050	0.080	0.67	0.50–0.88
3.	D3S0229i	3p24.1	16	0.122	0.169	0.0040	0.064	0.69	0.54–0.89
4.	D3S0085i	3q33.2	20	0.031	0.056	0.0080	0.16	0.54	0.35–0.85
5.	HUMUT880B	4q13.2	30	0.123	0.074	0.00029	0.0088	0.43	0.26–0.71
6.	D4S1126i	4q31.3	27	0.017	0.042	0.0010	0.027	0.19	0.07–0.57
7.	D5S1328i	5q21.3	10	0.045	0.081	0.0010	0.010	0.54	0.37–0.78
8.	D5S0703i	5q31.2	9	0.227	0.185	0.023	0.21	1.29	1.04–1.61
9.	HUMUT5011	6p12.3	19	0.075	0.110	0.0090	0.17	0.51	0.33–0.78
10.	D6S1146i	6p22.3	13	0.000	0.006	0.031	0.40	–	–
11.	HUMUT5779	6q25.1	11	0.287	0.217	0.00031	0.0034	1.45	1.19–1.78
12.	D7S0070i	7p21.1	10	0.426	0.351	0.00067	0.0067	1.37	1.15–1.65
13.	D7S0046i	7q11.22	11	0.105	0.165	0.000083	0.00091	0.59	0.46–0.77
14.	D8S0913i	8q21.3	13	0.026	0.010	0.011	0.14	2.64	1.27–5.50
15.	D8S0285i	8q21.13	6	0.208	0.150	0.000036	0.00022	0.68	0.57–0.82
16.	D9S0673i	9p13.3	14	0.109	0.076	0.011	0.15	1.50	1.10–2.04
17.	D11S1765	11q12.2	9	0.033	0.017	0.031	0.28	1.97	1.09–3.56
18.	D12S0914i	12q12	8	0.179	0.144	0.038	0.30	1.30	1.02–1.65
19.	D14S0504i	14q32.12	27	0.071	0.038	0.0020	0.054	1.93	1.29–2.89
20.	D15S988	15q22.33	15	0.100	0.062	0.0020	0.030	1.68	1.21–2.33
21.	D17S0234i	17p13.2	14	0.163	0.130	0.036	0.50	1.31	1.02–1.68
22.	HUMUT6385	19q13.2	9	0.071	0.116	0.00053	0.0048	0.58	0.43–0.79
23.	D21S0059i	21q21.1	16	0.158	0.124	0.034	0.54	1.32	1.03–1.70

P values calculated by Fisher's exact test, based on 2 × 2 contingency tables. The smallest *P* value was selected.

* *P* values were corrected by the number of alleles. The Fisher's exact test was carried out in the sex-pooled tall and short subjects ($n = 500$ each). CI, confidence interval

linkage analysis, 5q31 (Wu et al. 2003), 6q25 (Hirschhorn et al. 2001; Xu et al. 2002), 8q21.3 (Perola et al. 2007), 8q21.13 (Willemsen et al. 2004), and 21q21.1 (Hirschhorn et al. 2001), respectively. We also detected six strongly associated regions, 4q13.2, 4q31.3, 5q21.3, 7p21.1, 7q11.22, and 19q13.2, which have not been reported before. The inconsistent results of these studies may be due to population specificities and/or differences of technique.

Fine mapping by SNP

Among the nine most associated markers, we selected two: *D8S0285i* and *D15S988*. *D8S0285i* was the most strongly associated microsatellite, located at 8q21.13, and *D15S988* was flanked by a candidate gene, *SMAD3*, located at 15q22.33. 82 SNPs were surveyed and genotyped in a total of 1,555 samples (1,000 screened samples and additional 555 samples).

Ten SNPs at 8q21.13 showed nominal significance, among which SNP rs2220456 was the most strongly associated with height, showing empirical significance ($P = 0.000016$, $Pc = 0.0008$). These SNP associations might be reflected the reported evidence of linkage (Perola et al. 2007; Willemsen et al. 2004). Since an approximately 300 kb region in the vicinity of SNP rs2220456 and *D8S0285i* at 8q21.3 had no coding sequence according to NCBI build 36.2, we shifted our target to the locus at 15q22.33-q23. To cover a gene-containing region, we selected two additional microsatellites, *D15S0240i* and *D15S0028i*, and 64 SNPs at 15q22.33-q23 (Fig. 1). Among these, allele 230 of *D15S0240i* and six SNPs retained empirical significance ($Pc < 0.05$) as depicted in Figure1. SNP rs8038652, the most strongly associated SNP, is located in intron 1 of *IQCH*. The six SNPs maintained a strong LD index with each other ($D' > 0.9$ and $r^2 = 0.8$). Additionally, SNP rs8038652 and allele 230 of *D15S0240i* were in strong LD ($D' = 0.99$ and $r^2 = 0.77$).

Based on the SNP association results, SNP rs8038652 was further analyzed under different genetic models. Association analysis under a recessive model for SNP rs8038652 showed the lowest *P* value ($P = 0.000046$) with the AA genotype, indicating that the AA genotype of

rs8038652 has an adverse effect on adult height in Mongolians (odds ratio = 0.59, confidence interval, 0.46–0.76). Additionally, a deviation from HWE ($P = 0.04$) was observed in the tall height group with SNP rs8038652.

In conclusion, we have identified two candidate loci for adult height at 8q21.13 and 15q22.33-q23 in Mongolians. Although the causative polymorphisms were not determined in this study, we were able to locate genetic association with adult height to two regions. 15q22.33-q23 contains only three genes, so functional analyses should help to elucidate the causative polymorphisms. Analysis of the remaining seven highly associated microsatellite markers should lead to identification of new causative genes underlying adult height variation.

Acknowledgments This work was supported in part by the Research and Study Program of Tokai University Educational System General Research Organization, by the Center of Excellence 21st century program, and by a Grant-in-Aid for Scientific Research on Priority Areas "Applied Genomics" from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Arking DE, Pfeufer A, Post W, Kao WH, Newton-Cheh C, Ikeda M, West K, Kashuk C, Akyol M, Perz S, Jalilzadeh S, Illig T, Gieger C, Guo CY, Larson MG, Wichmann HE, Marbán E, O'Donnell CJ, Hirschhorn JN, Kääb S, Spooner PM, Meitinger T, Chakravarti A (2006) A common genetic variant in the NOS1 regulator NOS1AP modulates cardiac repolarization. Nat Genet 38:644–651

Beck SR, Brown WM, Williams AH, Pierce J, Rich SS, Langefeld CD (2003) Age-stratified QTL genome scan analyses for anthropometric measures. BMC Genet 4(Suppl 1):S31

Carmichael CM, McGue M (1995) A cross-sectional examination of height, weight, and body mass index in adult twins. J Gerontol A Biol Sci Med Sci 50:B237–B244

Cogan JD, Phillips JA 3rd, Sakati N, Frisch H, Schober E, Milner RD (1993) Heterogeneous growth hormone (GH) gene mutations in familial GH deficiency. J Clin Endocrinol Metab 76:1224–1228

Collins HE, Li H, Inda SE, Anderson J, Laiho K, Tuomilehto J, Seldin MF (2000) A simple and accurate method for determination of microsatellite total allele content differences between DNA pools. Hum Genet 106:218–226

Dempfle A, Wudy SA, Saar K, Hagemann S, Friedel S, Scherag A, Berthold LD, Alzen G, Gortner L, Blum WF, Hinney A, Nürnberg P, Schäfer H, Hebebrand J (2006) Evidence for involvement of the vitamin D receptor gene in idiopathic short stature via a genome-wide linkage study and subsequent association studies. Hum Mol Genet 15:2772–2783

Deng HW, Xu FH, Liu YZ, Shen H, Deng H, Huang QY, Liu YJ, Conway T, Li JL, Davies KM, Recker RR (2002) A whole-genome linkage scan suggests several genomic regions potentially containing QTLs underlying the variation of stature. Am J Med Genet 113:29–39

Ellis JA, Scurrah KJ, Duncan AE, Lamantia A, Byrnes GB, Harrap SB (2007) Comprehensive multi-stage linkage analyses identify a locus for adult height on chromosome 3p in a healthy Caucasian population. Hum Genet 121:213–222

Geller F, Dempfle A, Görg T (2003) Framingham Heart Study. Genome scan for body mass index and height in the Framingham Heart Study. BMC Genet 4(Suppl 1):S91

Gudbjartsson DF, Walters GB, Thorleifsson G, Stefansson H, Halldorsson BV, Zusmanovich P, Sulem P, Thorlacius S, Gylfason A, Steinberg S, Helgadottir A, Ingason A, Steinthorsdottir V, Olafsdottir EJ, Olafsdottir GH, Jonsson T, Borch-Johnsen K, Hansen T, Andersen G, Jorgensen T, Pedersen O, Aben KK, Witjes JA, Swinkels DW, den Heijer M, Franke B, Verbeek AL, Becker DM, Yanek LR, Becker LC, Tryggvadottir L, Rafnar T, Gulcher J, Kiemeney LA, Kong A, Thorsteinsdottir U, Stefansson K (2008) Many sequence variants affecting diversity of adult human height. Nat Genet 40:609–615

Hirschhorn JN, Lindgren CM, Daly MJ, Kirby A, Schaffner SF, Burtt NP, Altshuler D, Parker A, Rioux JD, Platko J, Gaudet D, Hudson TJ, Groop LC, Lander ES (2001) Genomewide linkage analysis of stature in multiple populations reveals several regions with evidence of linkage to adult height. Am J Hum Genet 69:106–116

Katoh T, Mano S, Ikuta T, Munkhbat B, Tounai K, Ando H, Munkhtuvshin N, Imanishi T, Inoko H, Tamiya G (2002) Genetic isolates in East Asia: a study of linkage disequilibrium in the X chromosome. Am J Hum Genet 71:395–400

Katoh T, Munkhbat B, Tounai K, Mano S, Ando H, Oyungerel G, Chae GT, Han H, Jia GJ, Tokunaga K, Munkhtuvshin N, Tamiya G, Inoko H (2005) Genetic features of Mongolian ethnic groups revealed by Y-chromosomal analysis. Gene 346:63–70

Kawashima M, Tamiya G, Oka A, Hohjoh H, Juji T, Ebisawa T, Honda Y, Inoko H, Tokunaga K (2006) Genomewide association analysis of human narcolepsy and a new resistance gene. Am J Hum Genet 79:252–263

Lander ES, Botstein D (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199

Lettre G, Jackson AU, Gieger C, Schumacher FR, Berndt SI, Sanna S, Eyheramendy S, Voight BF, Butler JL, Guiducci C, Illig T, Hackett R, Heid IM, Jacobs KB, Lyssenko V, Uda M; The Diabetes Genetics Initiative; FUSION; KORA; The Prostate, Lung Colorectal and Ovarian Cancer Screening Trial; The Nurses' Health Study; SardiNIA, Boehnke M, Chanock SJ, Groop LC, Hu FB, Isomaa B, Kraft P, Peltonen L, Salomaa V, Schlessinger D, Hunter DJ, Hayes RB, Abecasis GR, Wichmann HE, Mohlke KL, Hirschhorn JN (2008) Identification of ten loci associated with height highlights new biological pathways in human growth. Nat Genet. 40:584–591

Liu YZ, Xiao P, Guo YF, Xiong DH, Zhao LJ, Shen H, Liu YJ, Dvornyk V, Long JR, Deng HY, Li JL, Recker RR, Deng HW (2006) Genetic linkage of human height is confirmed to 9q22 and Xq24. Hum Genet 119:295–304

Mukhopadhyay N, Finegold DN, Larson MG, Cupples LA, Myers RH, Weeks DE (2003) A genome-wide scan for loci affecting normal adult height in the Framingham Heart Study. Hum Hered 55:191–201

Mukhopadhyay N, Weeks DE (2003) Linkage analysis of adult height with parent-of-origin effects in the Framingham Heart Study. BMC Genet. 41(Suppl 1):S76

Nakajima T, Wooding S, Sakagami T, Emi M, Tokunaga K, Tamiya G, Ishigami T, Umemura S, Munkhbat B, Jin F, Guan-Jun J, Hayasaka I, Ishida T, Saitou N, Pavelka K, Lalouel JM, Jorde LB, Inoue I (2004) Natural selection and population history in the human angiotensinogen gene (AGT): 736 complete AGT sequences in chromosomes from around the world. Am J Hum Genet 74:898–916

Oka A, Hayashi H, Tomizawa M, Okamoto K, Suyun L, Hui J, Kulski JK, Beilby J, Tamiya G, Inoko H (2003) Localization of a non-melanoma skin cancer susceptibility region within the major histocompatibility complex by association analysis using microsatellite markers. Tissue Antigens 61:203–210

Otgon G, Lkhagva L, Sarantuya G (2002) Normal values of physiological and laboratory parameters in human. Edmon Publishing Company, Ulaanbaatar, Mongolia

Perola M, Ohman M, Hiekkalinna T, Leppävuori J, Pajukanta P, Wessman M, Koskenvuo M, Palotie A, Lange K, Kaprio J, Peltonen L (2001) 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 69:117–123

Perola M, Sammalisto S, Hiekkalinna T, Martin NG, Visscher PM, Montgomery GW, Benyamin B, Harris JR, Boomsma D, Willemsen G, Hottenga JJ, Christensen K, Kyvik KO, Sørensen TI, Pedersen NL, Magnusson PK, Spector TD, Widen E, Silventoinen K, Kaprio J, Palotie A, Peltonen L, GenomEUtwin Project (2007) Combined genome scans for body stature in 6, 602 European twins: evidence for common Caucasian loci. PLoS Genet 8:1019–1028

Risch N, Merikangas K (1996) The future of genetic studies of complex human diseases. Science 273:1516–1517

Sammalisto S, Hiekkalinna T, Suviolahti E, Sood K, Metzidis A, Pajukanta P, Lilja HE, Soro-Paavonen A, Taskinen MR, Tuomi T, Almgren P, Orho-Melander M, Groop L, Peltonen L, Perola M (2005) A male-specific quantitative trait locus on 1p21 controlling human stature. J Med Genet 42:932–939

Sanna S, Jackson AU, Nagaraja R, Willer CJ, Chen WM, Bonnycastle LL, Shen H, Timpson N, Lettre G, Usala G, Chines PS, Stringham HM, Scott LJ, Dei M, Lai S, Albai G, Crisponi L, Naitza S, Doheny KF, Pugh EW, Ben-Shlomo Y, Ebrahim S, Lawlor DA, Bergman RN, Watanabe RM, Uda M, Tuomilehto J, Coresh J, Hirschhorn JN, Shuldiner AR, Schlessinger D, Collins FS, Davey Smith G, Boerwinkle E, Cao A, Boehnke M, Abecasis GR, Mohlke KL (2008) Common variants in the GDF5-UQCC region are associated with variation in human height. Nat Genet 40:198–203

Silventoinen K, Sammalisto S, Perola M, Boomsma DI, Cornes BK, Davis C, Dunkel L, De Lange M, Harris JR, Hjelmborg JV, Luciano M, Martin NG, Mortensen J, Nisticò L, Pedersen NL, Skytthe A, Spector TD, Stazi MA, Willemsen G, Kaprio J (2003) Heritability of adult body height: a comparative study of twin cohorts in eight countries. Twin Res 6:399–408

Tamiya G, Shinya M, Imanishi T, Ikuta T, Makino S, Okamoto K, Furugaki K, Matsumoto T, Mano S, Ando S, Nozaki Y, Yukawa W, Nakashige R, Yamaguchi D, Ishibashi H, Yonekura M, Nakami Y, Takayama S, Endo T, Saruwatari T, Yagura M, Yoshikawa Y, Fujimoto K, Oka A, Chiku S, Linsen SE, Giphart MJ, Kulski JK, Fukazawa T, Hashimoto H, Kimura M, Hoshina Y, Suzuki Y, Hotta T, Mochida J, Minezaki T, Komai K, Shiozawa S, Taniguchi A, Yamanaka H, Kamatani N, Gojobori T, Bahram S, Inoko H (2005) Whole genome association study of rheumatoid arthritis using 27 039 microsatellites. Hum Mol Genet 14:2305–2321

Thompson DB, Ossowski V, Janssen RC, Knowler WC, Bogardus C (1995) Linkage between stature and a region on chromosome 20 and analysis of a candidate gene, bone morphogenetic protein 2. Am J Med Genet 59:495–500

Visscher PM, Macgregor S, Benyamin B, Zhu G, Gordon S, Medland S, Hill WG, Hottenga JJ, Willemsen G, Boomsma DI, Liu YZ, Deng HW, Montgomery GW, Martin NG (2007) Genome partitioning of genetic variation for height from 11, 214 sibling pairs. Am J Hum Genet 81:1104–1110

Weedon MN, Lango H, Lindgren CM, Wallace C, Evans DM, Mangino M, Freathy RM, Perry JR, Stevens S, Hall AS, Samani NJ, Shields B, Prokopenko I, Farrall M, Dominiczak A; Diabetes Genetics Initiative; Wellcome Trust Case Control Consortium, Johnson T, Bergmann S, Beckmann JS, Vollenweider P, Waterworth DM, Mooser V, Palmer CN, Morris AD, Ouwehand WH; Cambridge GEM Consortium, Zhao JH, Li S, Loos RJ, Barroso I, Deloukas P, Sandhu MS, Wheeler E, Soranzo N, Inouye M, Wareham NJ, Caulfield M, Munroe PB, Hattersley AT, McCarthy MI, Frayling TM (2008) Genome-wide association analysis identifies 20 loci that influence adult height. Nat Genet. 40:575–583

Weedon MN, Lettre G, Freathy RM, Lindgren CM, Voight BF, Perry JR, Elliott KS, Hackett R, Guiducci C, Shields B, Zeggini E, Lango H, Lyssenko V, Timpson NJ, Burtt NP, Rayner NW, Saxena R, Ardlie K, Tobias JH, Ness AR, Ring SM, Palmer CN, Morris AD, Peltonen L, Salomaa V, The Diabetes Genetics Initiative, The Wellcome Trust Case Control Consortium, Smith GD, Groop LC, Hattersley AT, McCarthy MI, Hirschhorn JN, Frayling TM (2007) A common variant of HMGA2 is associated with adult and childhood height in the general population. Nat Genet 39:1245–1250

Willemsen G, Boomsma DI, Beem AL, Vink JM, Slagboom PE, Posthuma D (2004) QTLs for height: results of a full genome scan in Dutch sibling pairs. Eur J Hum Genet 12:820–828

Wiltshire S, Frayling TM, Hattersley AT, Hitman GA, Walker M, Levy JC, O'Rahilly S, Groves CJ, Menzel S, Cardon LR, McCarthy MI (2002) 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 70:543–546

Wu X, Cooper RS, Boerwinkle E, Turner ST, Hunt S, Myers R, Olshen RA, Curb D, Zhu X, Kan D, Luke A (2003) Combined analysis of genomewide scans for adult height: results from the NHLBI Family Blood Pressure Program. Eur J Hum Genet 11:271–274

Xu J, Bleecker ER, Jongepier H, Howard TD, Koppelman GH, Postma DS, Meyers DA (2002) 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 71:646–650

BMC Genomics

BioMed Central

Research article

Open Access

Major histocompatibility complex (Mhc) class Ib gene duplications, organization and expression patterns in mouse strain C57BL/6

Masato Ohtsuka*[1], Hidetoshi Inoko[1], Jerzy K Kulski[2] and Shinichi Yoshimura[1]

Address: [1]Division of Basic Molecular Science and Molecular Medicine, School of Medicine, Tokai University, Bohseidai, Isehara, Kanagawa 259-1193, Japan and [2]Centre for Forensic Science, The University of Western Australia, Nedlands, WA, Australia

Email: Masato Ohtsuka* - masato@is.icc.u-tokai.ac.jp; Hidetoshi Inoko - hinoko@is.icc.u-tokai.ac.jp; Jerzy K Kulski - kulski@mac.com; Shinichi Yoshimura - syoshimu@is.icc.u-tokai.ac.jp

* Corresponding author

Published: 17 April 2008

BMC Genomics 2008, **9**:178 doi:10.1186/1471-2164-9-178

Received: 7 November 2007
Accepted: 17 April 2008

This article is available from: http://www.biomedcentral.com/1471-2164/9/178

Abstract

Background: The mouse has more than 30 *Major histocompatibility complex* (*Mhc*) class Ib genes, most of which exist in the *H2* region of chromosome 17 in distinct gene clusters. Although recent progress in *Mhc* research has revealed the unique roles of several *Mhc* class Ib genes in the immune and non-immune systems, the functions of many class Ib genes have still to be elucidated. To better understand the roles of class Ib molecules, we have characterized their gene duplication, organization and expression patterns within the *H2* region of the mouse strain C57BL/6.

Results: The genomic organization of the *H2-Q*, *-T* and *-M* regions was analyzed and 21 transcribed *Mhc* class Ib genes were identified within these regions. Dot-plot and phylogenetic analyses implied that the genes were generated by monogenic and/or multigenic duplicated events. To investigate the adult tissue, embryonic and placental expressions of these genes, we performed RT-PCR gene expression profiling using gene-specific primers. Both tissue-wide and tissue-specific gene expression patterns were obtained that suggest that the variations in the gene expression may depend on the genomic location of the duplicated genes as well as locus specific mechanisms. The genes located in the *H2-T* region at the centromeric end of the cluster were expressed more widely than those at the telomeric end, which showed tissue-restricted expression in spite of nucleotide sequence similarities among gene paralogs.

Conclusion: Duplicated *Mhc* class Ib genes located in the *H2-Q*, *-T* and *-M* regions are differentially expressed in a variety of developing and adult tissues. Our findings form the basis for further functional validation studies of the *Mhc* class Ib gene expression profiles in specific tissues, such as the brain. The duplicated gene expression results in combination with the genome analysis suggest the possibility of long-range regulation of *H2-T* gene expression and/or important, but as yet unidentified nucleotide changes in the promoter or enhancer regions of the genes. Since the *Mhc* genomic region has diversified among mouse strains, it should be a useful model region for comparative analyses of the relationships between duplicated gene organization, evolution and the regulation of expression patterns.

Background

The *Major Histocompatibility Complex* (*MHC*) genomic region harbors duplicated genes that express protein molecules responsible for the rejection of transplanted tissue, restricted antigen presentation and the recognition of self and non-self [1,2]. The *Mhc* genomic region in the mouse, located on chromosome 17, is named *H2* and the genes within this region are usually classified into three distinct classes (I to III) based on their structure and function [3]. The class I molecules generally elicit immune responses by presenting peptide antigens derived from intracellular proteins to T lymphocytes and their genes can be classified into two groups, the classical *Mhc* class I (class Ia) genes and the non-classical *Mhc* class I (class Ib) genes. The classical *Mhc* class Ia genes, such as *H2-K* and *-D* in the mouse, are highly polymorphic, expressed widely and present antigens to CD8+ cytotoxic T cells. To date, most studies of the *MHC* class I genomic region have been focused on the immunological function of class Ia molecules [4-6].

The non-classical class Ib molecules are structurally similar to the classical class Ia proteins, but in contrast to the classical class Ia proteins, they have limited or no polymorphisms. They are more restricted in their tissue expression and some have functions other than antigen presentation to CD8+ T cells. The non-classical class Ib proteins have shorter cytoplasmic tails and some of them lack consensus residues associated with peptide binding [7]. The mouse is considered to have more than 30 *Mhc* class Ib genes in the genome [3]. Most *Mhc* class Ib genes are located at the telometric end of the 2 Mb-*H2* region within the *H2-Q*, *-T* and *-M* sub-regions, which were originally mapped and defined by recombination analysis. Although the non-classical class Ib genes are involved in immunological functions like the classical class Ia genes, they generally serve a more specialized role in the immune responses. The expression and function of some non-classical class Ib genes, including *H2-T23* (Qa-1), *-M3* and *-T3* (TL antigen), have been analyzed in detail. For example, Qa-1 is involved in the suppression of CD4+ T cell responses via CD94/NKG2A or CD94/NKG2C receptors [8,9]. The peptide presentation by the Qa-1 molecule may also have a role in CD8+ regulatory T cell activity [10]. H2-M3 molecules prime the rapid response of CD8+ T cells by presenting *N*-formylated bacterial peptides [11]. The TL antigen is involved in the formation of memory CD8+ T cells [12] and in the regulation of iIEL responses in the intestine by interaction with homodimeric CD8 alpha receptors [13].

The class Ib molecules are also involved in non-immune functions. For example, the *H2-M1* and *-M10* families of the class Ib genes specifically interact with the V2R class of pheromone receptors presented on the cell surfaces of the vomeronasal organ [14,15]. The Qa-2 proteins encoded by *H2-Q7* and *-Q9* class Ib genes influence the rate of preimplantation embryonic development and subsequent embryonic survival [16]. In addition, the class I molecules have recently been shown to contribute to the development and plasticity of the brain [17,18]. So far, there is little information about which of the non-classical class Ib genes are involved in this function.

The molecular functions of many of the other class Ib molecules are still far from being understood and even the expression patterns for many of the *Mhc* class Ib genes remain to be elucidated. The *Mhc* class Ib genes are members of gene clusters that have been generated by different rounds of duplication and deletion [19]. In the mouse, the telomeric 1 Mb of the *Mhc* including the *H2-M* region was well characterized using the 129/Sv inbred strain [20]. The possible evolutionary fates of duplicated genes are nonfunctionalization, neofunctionalization or subfunctionalization [21]. Genes recently duplicated may even have the same functions by having and using identical or similar expression domain sequences. In order to better understand the role of class Ib molecules expressed by duplicated genes in different tissues, we have undertaken to examine, identify and characterize the *Mhc* class Ib gene duplication, organization and expression patterns within the *H2* region of the mouse strain C57BL/6.

The whole genome of the laboratory mouse strain C57BL/6J has been almost fully sequenced [22]. However, the genomic organization of the *Mhc* class I region of mice varies markedly between different haplotypes and inbred strains [20]. In the present study, we selected *Mhc* class Ib DNA sequences from the mouse genome database (NCBI Entrez Genome Project ID 9559), and characterized the organization of the *Mhc* class Ib genomic region for the mouse C57BL/6 strain (haplotype b). Expression patterns of each of the *Mhc* class Ib genes were examined by RT-PCR using gene-specific primer sets, and we identified *Mhc* class Ib genes with either tissue-restricted expression or tissue-wide expression. We also identified monogenic and multigenic duplicated regions within the *H2-T* region of the mouse inbred-strain, C57BL/6. Based on the results of our comprehensive analysis of the *Mhc* class Ib gene duplication, organization and expression patterns, we discuss the possible relationships and regulatory outcomes between the genomic location and expression patterns of the mouse *Mhc* class Ib duplicated genes.

Results and Discussion

Identification and genomic organization of transcribed Mhc class Ib genes

As the aim of this study was to determine the tissue expression patterns for each of the duplicated *Mhc* class Ib genes, we first needed to identify the location and the number of transcribing *Mhc* class Ib genes in the mouse genomic