

Supplementary Information for

New Models of Collaboration in Genome-Wide Association Studies: The Genetic Association Information Network (GAIN)

The GAIN Collaborative Research Group: Teri A. Manolio*, Laura Lyman Rodriguez, Lisa Brooks, Gonçalo Abecasis for the Collaborative Association Study of Psoriasis, Dennis Ballinger, Mark Daly, Peter Donnelly, Stephen V. Faraone for the International Multi-Center ADHD Genetics Project, Kelly Frazer, Stacey Gabriel, Pablo Gejman for the Molecular Genetics of Schizophrenia Collaboration, Alan Guttmacher, Emily L. Harris, Thomas Insel, John R. Kelsoe for the Bipolar Genome Study, Eric Lander, Norma McCowin, Matthew D. Mailman, Elizabeth Nabel, James Ostell, Elizabeth Pugh, Stephen Sherry, Patrick F. Sullivan for the Major Depression: Stage 1 Genomewide Association in Population-Based Samples, John F. Thompson, James Warram for the Genetics of Kidneys in Diabetes (GoKinD) Study, David Wholley, Patrice M. Milos, Francis S. Collins

*To whom correspondence should be addressed. E-mail: manolio@nih.gov

Submitted : 8 August 2007

This PDF file includes:

Supplementary Note

Supplementary Table 1

Supplementary Note: Other Groups Participating in GAIN

Genome-Wide Association Study of Schizophrenia: Pablo V. Gejman, Douglas F. Levinson, Farooq Amin, Donald W. Black, Nancy G. Buccola, William F. Byerley, C. Robert Cloninger, Jubao Duan, Frank Dudbridge, Ayman H. Fanous, Robert Freedman, Peter A. Holmans, Kenneth S. Kendler, Jun Li, Maria Martinez, Paul McKeigue, Bryan J. Mowry, Itsik Pe'er, John P. Rice, Alan R. Sanders, Jianxin Shi, Mark D. Shriver, Jeremy M. Silverman, Alice S. Whittemore.

Bipolar Genome Study: John R. Kelsoe, John I. Nurnberger, Jr., Eric T. Meyer, Thomas B. Barrett, William F. Byerley, William Coryell, William Scheftner, Dean MacKinnon, J. Raymond DePaulo, James B. Potash, Peter Zandi, David Cutler, Elliot S. Gershon, Judith Badner, Chun-yu Liu, Francis McMahon, Howard Edenberg, John Rice, Melvin McInnis, Wade Berrettini, William Lawson, Eleazar Eskin, David Craig, Nicholas Schork,

Major Depression: Stage 1 Genomewide Association in Population-Based Samples: Patrick Sullivan, Dorret I. Boomsma, Brenda W.J.H. Pennix, Eco J.C. de Geus, Gonneke Willemsen, Witte Hoogendijk, Danyu Lin, Jung-Ying Tzeng

International Multi-Center ADHD Genetics Project: Stephen V. Faraone, Philip Asherson, Tobias Banaschewski, Jan Buitelaar, Richard P. Ebstein, Barbara Franke, Michael Gill, Christoph Lange, Nan Laird, Jessica Lasky-Su, Ana Miranda, Fernando Mulas, Benjamin Neale, Robert D. Oades, Herbert Roeyers, Aribert Rothenberger, Joseph Sergeant, Edmund Sonuga-Barke, H.-C. Steinhausen, Margaret Thompson

Collaborative Association Study of Psoriasis: Gonçalo R. Abecasis, Kristina Callis, Jun Ding, Jennifer Gardner, David Goldgar, Cynthia Helms, Ravi Hiremagalore, Alan Menter, Rajan Nair, Tyler Nelson, Amy Perlmutter, Phil Stuart, Trilok Tejasvi, John Voorhees, Carol Wise, Gerald G. Krueger, Anne M. Bowcock, James T Elder

Search for Susceptibility Genes for Diabetic Nephropathy in Type 1 Diabetes: James Warram, Andrzej S. Krolewski, David Poznik, John J. Rogus

GAIN Steering Committee: James F. Battey, Michael Brown, Gregory L. Burke, Francis Collins, Peter Donnelly, Gary H. Gibbons, Jack M. Guralnik, Katrina Gwinn-Hardy, Robert Hoover, Bartha M. Knoppers, Richard P. Lifton, Yvonne Maddox, Patrice M. Milos, Elizabeth G. Nabel, Ross Prentice, Griffin Rodgers, Ralph Snyderman, Sharon F. Terry, Tom Willis

GAIN Analysis Methodology Group: Peter Donnelly, Nancy Cox, Gonçalo Abecasis, Christoph Lange, Doug Levinson, Danyu Lin, Josyf Mychaleckyj, Nik Schork

GAIN Data Access Committee: Emily Harris, Christine Grady, Thomas Lehner, Catherine McKeon, Joel Moss, Brad Ozenberger, William Sharrock, Vaurice Stark

GAIN Genotyping Group: Gonçalo Abecasis, Dennis Ballinger, James Battey, Erica Beilharz, Lisa Brooks, David Craig, Mark Daly, Peter Donnelly, Howard Edenberg, Steve Faraone, Mike Feolo, Kelly Frazer, Stacey Gabriel, Pablo Gejman, Emily Harris, John Kelsoe, Steve Lincoln, Norma McCowin, Daniel Mirel, Justin Paschall, David Poznik, Elizabeth Pugh, Alan Sanders,

Stephen Sherry, Jeffery Struewing, Patrick Sullivan, John Thompson, James Warram, David Wholley

GAIN Technical Advisory Group: Hugh Auchincloss, Lisa Brooks, Mark Daly, Ana Diez-Roux, Kelly Frazer, Emily Harris, Matthew D. Mailman, Teri A. Manolio, Anne Newman, Pilar Ossorio, Laura Lyman Rodriguez, Stephen Sherry, John Thompson

dbGaP Project Team: Jeff Beck, Michael Feolo, Al Graeff, Matthew Mailman, James Ostell, Justin Paschall, Donald Preuss, Stephen Sherry, Karl Sirotkin, Bart Trawick, Eugene Yaschenko

FNIH Board: John Edward Porter, Deeda Blair, Kathy Bloomgarden, Buffy Cafritz, Mark D. Ein, Caroline A. Kovac, Sherry Lansing, Paul M. Montrone, Martin J. Murphy, Steven M. Paul, Philip A. Pizzo, Lily Safra, Charles A. Sanders, Ellen V. Sigal, Solomon H. Snyder, Stephen P. Spielberg, Patrick C. Walsh

FNIH Staff: David Wholley, Gary Barbarash, Daniel Carucci, Michael Gottlieb, Stephanie James, Norma Harris, Norma McCowin, Amy McGuire, Charles Pucie, Susan Weiner, Julie Wolf-Rodda

NIH Staff: Francis Collins, Solome Abebe, Catherine Bennett, Lisa Brooks, Phyliss Frosst, Al Graeff, Mark Guyer, Emily Harris, David Lipman, Matthew Mailman, Teri Manolio, Barbara Mittleman, James Ostell, Laura Lyman Rodriguez, Stephen Sherry, Jeffery Struewing, Larry Thompson, Susan Vasquez

Supplementary Table 1. Coverage, call rates, and concordance of genotype data for the Perlegen platform and the Affymetrix SNP Array 5.0 platform on the HapMap phase II samples.

	Perlegen		Affymetrix/Broad	
Number of good SNPs ^a	480,744		439,249	
Coverage in each HapMap analysis panel, Release 21				
	Single-Marker	Multi-Marker	Single-Marker	Multi-Marker
CEU	0.89	0.96	0.76	0.84
CHB+JPT	0.84	0.90	0.75	0.81
YRI	0.59	0.78	0.59	0.71
Average call rate per SNP	98.9%		99.3%	
Concordance with HapMap SNPs, Release 22				
For genotypes homozygous in the HapMap samples	99.8%		99.9%	
For genotypes heterozygous in the HapMap samples	99.8%		99.8%	

^a Definitions of good samples, SNPs, and genotype calls (performed in this order):

1. Good samples have calls at the Quality Score (QS) cutoff or higher for at least 80% of the SNPs. This cutoff is 10 for Perlegen and 4000 for Broad; the QS values are not comparable for the two platforms.
2. Good SNPs have calls at the QS cutoff or higher in at least 90% of the good samples, show no extreme Hardy-Weinberg deviations ($p \geq 0.00033$ in each HapMap analysis panel), have no more than 1 Mendelian error in the 60 HapMap trios, and have a minor allele frequency of at least 1%.
3. For the good samples and good SNPs, any genotype calls with QS less than the cutoff are set to "no call."

^b The mean maximum r^2 value between the platform SNPs and HapMap phase II SNPs ($MAF \geq 0.05$ in each HapMap analysis panel) as assessed by single marker and multimarker methods [Pe'er I, de Bakker PI, Maller J, Yelensky R, Altshuler D, Daly MJ. Evaluating and improving power in whole-genome association studies using fixed marker sets. *Nat. Genet.* 38, 663-7 (2006)].