RFA-HL-07-010: Methods of Analysis of Gene-Environment Interactions in Complex Diseases: The Genes and Environment Initiative (R01)

Abstracts and Publications

Last updated: July 30, 2010

Investigator: Chris Amos

Project: Statistical Methods for Gene Environment Interactions in Lung Cancer

1. Amos, C.I., Pinney, S.M., Li, Y., Kupert, E., Lee, J., de Andrade, M.A., Yang, P., Schwartz, A.G., Fain, P.R., Gazdar, A., Minna, J., Wiest, J.S., Zeng, D., Rothschild, H., Mandal, D., You, M., Coons, T., Gaba, C., Bailey-Wilson, J.E. and Anderson, M.W. (2010) A susceptibility locus on chromosome 6q greatly increases lung cancer risk among light and never smokers, *Cancer Res*, **70**, 2359-2367.

2. Landi, M.T., Chatterjee, N., Yu, K., Goldin, L.R., Goldstein, A.M., Rotunno, M., Mirabello, L., Jacobs, K., Wheeler, W., Yeager, M., Bergen, A.W., Li, Q., Consonni, D., Pesatori, A.C., Wacholder, S., Thun, M., Diver, R., Oken, M., Virtamo, J., Albanes, D., Wang, Z., Burdette, L., Doheny, K.F., Pugh, E.W., Laurie, C., Brennan, P., Hung, R., Gaborieau, V., McKay, J.D., Lathrop, M., McLaughlin, J., Wang, Y., Tsao, M.S., Spitz, M.R., Krokan, H., Vatten, L., Skorpen, F., Arnesen, E., Benhamou, S., Bouchard, C., Metsapalu, A., Vooder, T., Nelis, M., Valk, K., Field, J.K., Chen, C., Goodman, G., Sulem, P., Thorleifsson, G., Rafnar, T., Eisen, T., Sauter, W., Rosenberger, A., Bickeboller, H., Risch, A., Chang-Claude, J., Wichmann, H.E., Stefansson, K., Houlston, R., Amos, C.I., Fraumeni, J.F., Jr., Savage, S.A., Bertazzi, P.A., Tucker, M.A., Chanock, S. and Caporaso, N.E. (2009) A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma, *Am J Hum Genet*, **85**, 679-691.

3. Gu, X., Frankowski, R.F., Rosner, G.L., Relling, M., Peng, B. and Amos, C.I. (2009) A modified forward multiple regression in high-density genome-wide association studies for complex traits, *Genet Epidemiol*, **33**, 518-525.

4. Wang, Y., Broderick, P., Webb, E., Wu, X., Vijayakrishnan, J., Matakidou, A., Qureshi, M., Dong, Q., Gu, X., Chen, W.V., Spitz, M.R., Eisen, T., Amos, C.I. and Houlston, R.S. (2008) Common 5p15.33 and 6p21.33 variants influence lung cancer risk, *Nat Genet*, **40**, 1407-1409.

5. Spitz, M.R., Amos, C.I., Dong, Q., Lin, J. and Wu, X. (2008) The CHRNA5-A3 region on chromosome 15q24-25.1 is a risk factor both for nicotine dependence and for lung cancer, *J Natl Cancer Inst*, **100**, 1552-1556.

6. Liu, P., Vikis, H.G., Wang, D., Lu, Y., Wang, Y., Schwartz, A.G., Pinney, S.M., Yang, P., de Andrade, M., Petersen, G.M., Wiest, J.S., Fain, P.R., Gazdar, A., Gaba, C., Rothschild, H., Mandal, D., Coons, T., Lee, J., Kupert, E., Seminara, D., Minna, J., Bailey-Wilson, J.E., Wu, X., Spitz, M.R., Eisen, T., Houlston, R.S., Amos, C.I., Anderson, M.W. and You, M. (2008) Familial aggregation of common sequence variants on 15q24-25.1 in lung cancer, *J Natl Cancer Inst*, **100**, 1326-1330.

7. Peng, B. and Amos, C.I. (2008) Forward-time simulations of non-random mating populations using simuPOP, Bioinformatics, 24, 1408-1409.

8. Amos, C.I., Wu, X., Broderick, P., Gorlov, I.P., Gu, J., Eisen, T., Dong, Q., Zhang, Q., Gu, X., Vijayakrishnan, J., Sullivan, K., Matakidou, A., Wang, Y., Mills, G., Doheny, K., Tsai, Y.Y., Chen, W.V., Shete, S., Spitz, M.R. and Houlston, R.S. (2008) Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1, *Nat Genet*, **40**, 616-622.

Investigator: Nilanjan Chatterjee

Project: Integrative Analysis of Genetic and Environmental Data in Epidemiologic Studies

1. Park JH, Wacholder S, Gail MH, Peters U, Jacobs KB, Chanock SJ, et al. (2010) Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nat Genet*. **42**: 570-5.

2. Chatterjee N., Luo S., Chen Y.H., and Carroll R.J. Analysis of case-control association studies: SNPs, haplotypes and imputation, *Statistical Science*. In press.

3. Li H., Gail M., Brendt S., and Chatterjee N. Using cases to strengthen association between single nucleotide polymorphisms and a secondary phenotype in genome-wide association studies, *Genetic Epidemiology*. In press.

4. Menashe I., Meader D., Garcia-Closas M., Figueroa J., Bhattacharjee S., Rotunno M., Kraft P., Hunter D., Chanock S., Rosenberg P., and Chatterjee N. Pathway analysis of breast cancer genome wide association study highlights three pathways and one canonical signaling cascade, *Cancer Research*. Accepted.

5. Bhattacharjee, S., Wang, Z., Ciampa, J., Kraft, P., Chanock, S., Yu, K. and Chatterjee, N. (2010) Using principal components of genetic variation for robust and powerful detection of gene-gene interactions in case-control and case-only studies, *Am J Hum Genet*, **86**, 331-342.

6. Landi, M.T., Chatterjee, N., Yu, K., Goldin, L.R., Goldstein, A.M., Rotunno, M., Mirabello, L., Jacobs, K., Wheeler, W., Yeager, M., Bergen, A.W., Li, Q., Consonni, D., Pesatori, A.C., Wacholder, S., Thun, M., Diver, R., Oken, M., Virtamo, J., Albanes, D., Wang, Z., Burdette, L., Doheny, K.F., Pugh, E.W., Laurie, C., Brennan, P., Hung, R., Gaborieau, V., McKay, J.D., Lathrop, M., McLaughlin, J., Wang, Y., Tsao, M.S., Spitz, M.R., Krokan, H., Vatten, L., Skorpen, F., Arnesen, E., Benhamou, S., Bouchard, C., Metsapalu, A., Vooder, T., Nelis, M., Valk, K., Field, J.K., Chen, C., Goodman, G., Sulem, P., Thorleifsson, G., Rafnar, T., Eisen, T., Sauter, W., Rosenberger, A., Bickeboller, H., Risch, A., Chang-Claude, J., Wichmann, H.E., Stefansson, K., Houlston, R., Amos, C.I., Fraumeni, J.F., Jr., Savage, S.A., Bertazzi, P.A., Tucker, M.A., Chanock, S. and Caporaso, N.E. (2009) A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma, *Am J Hum Genet*, **85**, 679-691.

7. Maity, A., Carroll, R.J., Mammen, E. and Chatterjee, N. (2009) Testing in semiparametric models with interaction, with applications to gene-environment interactions, *J R Stat Soc Series B Stat Methodol*, **71**, 75-96.

8. Luo, S., Mukherjee, B., Chen, J. and Chatterjee, N. (2009) Shrinkage estimation for robust and efficient screening of single-SNP association from case-control genome-wide association studies, *Genet Epidemiol*, 33, 740-750.

9. Chen, Y.H., Chatterjee, N. and Carroll, R.J. (2009) Shrinkage Estimators for Robust and Efficient Inference in Haplotype-Based Case-Control Studies, *J Am Stat Assoc*, **104**, 220-233.

10. Mukherjee, B., Ahn, J., Gruber, S.B., Rennert, G., Moreno, V. and Chatterjee, N. (2008) Tests for gene-environment interaction from case-control data: a novel study of type I error, power and designs, *Genet Epidemiol*, 32, 615-626.

Investigator: David Herrington

Project: Machine Learning to Identify Complex Interactions in Genome-wide Association Data

1. Miller, D.J., Zhang, Y., Yu, G., Liu, Y., Chen, L., Langefeld, C.D., Herrington, D. and Wang, Y. (2009) An algorithm for learning maximum entropy probability models of disease risk that efficiently searches and sparingly encodes multilocus genomic interactions, *Bioinformatics*, **25**, 2478-2485.

2. Chen, L., Yu, G., Miller, D.J., Song, L., Langefeld, C.D., Herrington, D., Liu, Y. and Wang, Y. (2009) A Ground Truth Based Comparative Study on Detecting Epistatic SNPs, *Proc. IEEE Machine Learning for Signal Processing*, 26-31.

3. Yu, G., Herrington, D., Langefeld, C. and Wang, W. (2008) Detection of complex interactions of multi-locus SNPs, *Proc. IEEE Machine Learning for Signal Processing*, 85-90.

Investigator: Ed Iversen

Project: Bayesian Modeling and Optimal Design for Studies of Gene-Environment Association

1. Wilson, M.A., Iversen, Jr, E.S., Clyde, M.A., Schmidler, S.C., and Schildraut, J.M. Bayesian Model Search and Multilevel Inference for SNP Association Studies, *Annals of Applied Statistics*. In press.

2. Schildkraut, J.M., Iversen, E.S., Wilson, M.A., Clyde, M.A., Moorman, P.G., Palmieri, R.T., Whitaker, R., Bentley, R.C., Marks, J.R. and Berchuck, A. (2010) Association between DNA damage response and repair genes and risk of invasive serous ovarian cancer, *PLoS One*, **5**, e10061. Available online only.

3. Doherty, J.A., Rossing, M.A., Cushing-Haugen, K.L., Chen, C., Van Den Berg, D.J., Wu, A.H., Pike, M.C., Ness, R.B., Moysich, K., Chenevix-Trench, G., Beesley, J., Webb, P.M., Chang-Claude, J., Wang-Gohrke, S., Goodman, M.T., Lurie, G., Thompson, P.J., Carney, M.E., Hogdall, E., Kjaer, S.K., Hogdall, C., Goode, E.L., Cunningham, J.M., Fridley, B.L., Vierkant, R.A., Berchuck, A., Moorman, P.G., Schildkraut, J.M., Palmieri, R.T., Cramer, D.W., Terry, K.L., Yang, H.P., Garcia-Closas, M., Chanock, S., Lissowska, J., Song, H., Pharoah, P.D., Shah, M., Perkins, B., McGuire, V., Whittemore, A.S., Di Cioccio, R.A., Gentry-Maharaj, A., Menon, U., Gayther, S.A., Ramus, S.J., Ziogas, A., Brewster, W., Anton-Culver, H. and Pearce, C.L. (2010) ESR1/SYNE1 polymorphism and invasive epithelial ovarian cancer risk: an Ovarian Cancer Association Consortium study, *Cancer Epidemiol Biomarkers Prev*, **19**, 245-250.

4. 4. Schildkraut, J.M., Goode, E.L., Clyde, M.A., Iversen, E.S., Moorman, P.G., Berchuck, A., Marks, J.R., Lissowska, J., Brinton, L., Peplonska, B., Cunningham, J.M., Vierkant, R.A., Rider, D.N., Chenevix-Trench, G., Webb, P.M., Beesley, J., Chen, X., Phelan, C., Sutphen, R., Sellers, T.A., Pearce, L., Wu, A.H., Van Den Berg, D., Conti, D., Elund, C.K., Anderson, R., Goodman, M.T., Lurie, G., Carney, M.E., Thompson, P.J., Gayther, S.A., Ramus, S.J., Jacobs, I., Kruger Kjaer, S., Hogdall, E., Blaakaer, J., Hogdall, C., Easton, D.F., Song, H., Pharoah, P.D., Whittemore, A.S., McGuire, V., Quaye, L., Anton-Culver, H., Ziogas, A., Terry, K.L., Cramer, D.W., Hankinson, S.E., Tworoger, S.S., Calingaert, B., Chanock, S., Sherman, M. and Garcia-Closas, M. (2009) Single nucleotide polymorphisms in the TP53 region and susceptibility to invasive epithelial ovarian cancer, *Cancer Res*, **69**, 2349-2357.

Investigator: Kung-Yee Liang

Project: Novel Statistical Methods for Gene-Environment Interactions in Complex Diseases

1. Louis, T.A. and Ruczinski, I. (2010) Efficient evaluation of ranking procedures when the number of units is large, with application to SNP identification, *Biom J*, **52**, 34-49.

2. Mathias, R.A., Grant, A.V., Rafaels, N., Hand, T., Gao, L., Vergara, C., Tsai, Y.J., Yang, M., Campbell, M., Foster, C., Gao, P., Togias, A., Hansel, N.N., Diette, G., Adkinson, N.F., Liu, M.C., Faruque, M., Dunston, G.M., Watson, H.R., Bracken, M.B., Hoh, J., Maul, P., Maul, T., Jedlicka, A.E., Murray, T., Hetmanski, J.B., Ashworth, R., Ongaco, C.M., Hetrick, K.N., Doheny, K.F., Pugh, E.W., Rotimi, C.N., Ford, J., Eng, C., Burchard, E.G., Sleiman, P.M., Hakonarson, H., Forno, E., Raby, B.A., Weiss, S.T., Scott, A.F., Kabesch, M., Liang, L., Abecasis, G., Moffatt, M.F., Cookson, W.O., Ruczinski, I., Beaty, T.H. and Barnes, K.C. (2010) A genome-wide association study on African-ancestry populations for asthma, *J Allergy Clin Immunol*, **125**, 336-346 e334.

3. Scharpf, R.B. and Ruczinski, I. (2009) R classes and methods for SNP array data, *Methods Mol Biol*, **593**, 67-79.

4. Sleiman, P.M., Flory, J., Imielinski, M., Bradfield, J.P., Annaiah, K., Willis-Owen, S.A., Wang, K., Rafaels, N.M., Michel, S., Bonnelykke, K., Zhang, H., Kim, C.E., Frackelton, E.C., Glessner, J.T., Hou, C., Otieno, F.G., Santa, E., Thomas, K., Smith, R.M., Glaberson, W.R., Garris, M., Chiavacci, R.M., Beaty, T.H., Ruczinski, I., Orange, J.M., Allen, J., Spergel, J.M., Grundmeier, R., Mathias, R.A., Christie, J.D., von Mutius, E., Cookson, W.O., Kabesch, M., Moffatt, M.F., Grunstein, M.M., Barnes, K.C., Devoto, M., Magnusson, M., Li, H., Grant, S.F., Bisgaard, H. and Hakonarson, H. (2009) Variants of DENND1B associated with asthma in children, *N Engl J Med*, **362**, 36-44.

5. Gallicchio, L., Chang, H.H., Christo, D.K., Thuita, L., Huang, H.Y., Strickland, P., Ruczinski, I., Clipp, S. and Helzlsouer, K.J. (2009) Single nucleotide polymorphisms in obesity-related genes and all-cause and cause-specific mortality: a prospective cohort study, *BMC Med Genet*, 10, 103.

6. Jorgensen, T.J., Ruczinski, I., Kessing, B., Smith, M.W., Shugart, Y.Y. and Alberg, A.J. (2009) Hypothesis-driven candidate gene association studies: practical design and analytical considerations, *Am J Epidemiol*, **170**, 986-993.

7. Wheless, L., Ruczinski, I., Alani, R.M., Clipp, S., Hoffman-Bolton, J., Jorgensen, T.J., Liegeois, N.J., Strickland, P.T. and Alberg, A.J. (2009) The association between skin characteristics and skin cancer prevention behaviors, *Cancer Epidemiol Biomarkers Prev*, **18**, 2613-2619.

8. Scharpf, R.B., Parmigiani, G., Pevsner, J. and Ruczinski, I. (2009) Hidden Markov models for the assessment of chromosomal alterations using high-throughput SNP arrays, *Ann Appl Stat*, **2**, 687-713.

9. Price, A.L., Tandon, A., Patterson, N., Barnes, K.C., Rafaels, N., Ruczinski, I., Beaty, T.H., Mathias, R., Reich, D. and Myers, S. (2009) Sensitive detection of chromosomal segments of distinct ancestry in admixed populations, *PLoS Genet*, **5**, e1000519.

10. Miller, C.L., Murakami, P., Ruczinski, I., Ross, R.G., Sinkus, M., Sullivan, B. and Leonard, S. (2009) Two complex genotypes relevant to the kynurenine pathway and melanotropin function show association with schizophrenia and bipolar disorder, *Schizophr Res*, **113**, 259-267.

11. Kottgen, A., Glazer, N.L., Dehghan, A., Hwang, S.J., Katz, R., Li, M., Yang, Q., Gudnason, V., Launer, L.J., Harris, T.B., Smith, A.V., Arking, D.E., Astor, B.C., Boerwinkle, E., Ehret, G.B., Ruczinski, I., Scharpf, R.B., Ida Chen, Y.D., de Boer, I.H., Haritunians, T., Lumley, T., Sarnak, M., Siscovick, D., Benjamin, E.J., Levy, D., Upadhyay, A., Aulchenko, Y.S., Hofman, A., Rivadeneira, F., Uitterlinden, A.G., van Duijn, C.M., Chasman, D.I., Pare, G., Ridker, P.M., Kao, W.H., Witteman, J.C., Coresh, J., Shlipak, M.G. and Fox, C.S. (2009) Multiple loci associated with indices of renal function and chronic kidney disease, *Nat Genet*. Available online only.

12. Himes, B.E., Hunninghake, G.M., Baurley, J.W., Rafaels, N.M., Sleiman, P., Strachan, D.P., Wilk, J.B., Willis-Owen, S.A., Klanderman, B., Lasky-Su, J., Lazarus, R., Murphy, A.J., Soto-Quiros, M.E., Avila, L., Beaty, T., Mathias, R.A., Ruczinski, I., Barnes, K.C., Celedon, J.C., Cookson, W.O., Gauderman, W.J., Gilliland, F.D., Hakonarson, H., Lange, C., Moffatt, M.F., O'Connor, G.T., Raby, B.A., Silverman, E.K. and Weiss, S.T. (2009) Genome-wide association analysis identifies PDE4D as an asthma-susceptibility gene, *Am J Hum Genet*, **84**, 581-593.

Investigator: Sebastian Zoellner

Project: Computational Statistic Approaches to Gene-Environment Interaction

1. Zawistowski, M., Gopalakrishnan, S., Li, Y., and Zöllner, S. (2010) Extending rare variant testing strategies: analysis of non-coding sequence and imputed genotypes. Submitted.

2. Rosenberg, N.A., Huang, L., Jewett, E.M., Szpiech, Z.A., Jankovic, I., and Boehnke, M. (2010) Genome-wide association studies in diverse populations. *Nat Rev Genet*, **11**, 356-66.

3. Huang, L., Wang, C., and Rosenberg, N.A. (2009) The relationship between imputation error and statistical power in genetic association studies in diverse populations. *Am J Hum Genet*, **85**, 692-8.

4. Saunders, E.F., Zhang, P., Copeland, J.N., McLnnis, M.G., and Zöllner, S. (2009) Suggestive linkage at 9p22 in bipolar disorder weighted by alcohol abuse. *Am J Med Genet B Neuropsychiatr Genet*, **150B**, 1133-8.

5. Huang, L., Li, Y., Singleton, A.B., Hardy, J.A., Abecasis, G., Rosenberg, N.A., *et a*l. (2009) Genotype-imputation accuracy across worldwide human populations. *Am J Hum Genet*, **84**, 235-50.

6. Rosenberg, N.A., and Vanliere, J.M. (2009) Replication of genetic associations as pseudoreplication due to shared genealogy. *Genet Epidemio*, **33**, 479-87.

7. Jakobsson, M., Scholz, S.W., Scheet, P., Gibbs, J.R., VanLiere, J.M., Fung, H.C., *et al*. (2008) Genotype, haplotype and copy-number variation in worldwide human populations. *Nature*, **451**, 998-1003.