**Publications (Total: 536)**

Including 2 books, 458 peer reviewed journal articles, 24 refereed proceeding articles, 3 letters, 21 journal review articles, and 28 book chapters.

**Books (2):**

[1] S. Lin, **H. Zhao** (eds) (2009) *Handbook of Statistical Genetics*. Springer-Verlag Berlin Heidelberg.

[2] H. H-S Lu, B. Schölkopf, **H. Zhao** (eds) (2011) *Handbook of Computational Statistics: Statistical Bioinformatics*, Springer.

**Peer-Reviewed Journal Articles (458):**

**1994 (1)**

[1] K. Doksum, S. Blyth, E. Bradlow, X. L. Meng, **H. Zhao**. (1994) Correlation curves as local measures of variance explained by regression. *Journal of American Statistical Association*, 89: 571-583.**1995 (3)**

[2] **H. Zhao**, M. S. McPeek, T. P. Speed. (1995) Statistical analysis of crossover interference using the chi-square model. *Genetics*, 139: 1045-1056.

[3] **H. Zhao**, M. S. McPeek, T. P. Speed. (1995) Statistical analysis of chromatid interference. *Genetics*, 139: 1057-1065.[4] D. R. Goldstein, **H. Zhao**, T. P. Speed. (1995) Relative efficiencies of several statistical models of recombination for exclusion mapping and gene ordering. *Genomics*, 27: 265-273.

**1996 (2)**

[5] **H. Zhao**, T. P. Speed. (1996) On genetic map functions. *Genetics*, 142: 1369-1377.

[6] J. D. Ohmen, H. Y. Yang, K. K. Yamamoto, **H. Zhao**, [Y. Ma](http://hmg.oxfordjournals.org/search?author1=Yuanhong+Ma&sortspec=date&submit=Submit), [L. G. Bentley](http://hmg.oxfordjournals.org/search?author1=L.+Gordon+Bentley&sortspec=date&submit=Submit), [Z. Huang](http://hmg.oxfordjournals.org/search?author1=Zhihan+Huang&sortspec=date&submit=Submit), [S. Gerwehr](http://hmg.oxfordjournals.org/search?author1=Scott+Gerwehr&sortspec=date&submit=Submit), [S. Pressman](http://hmg.oxfordjournals.org/search?author1=Sheila+Pressman&sortspec=date&submit=Submit), [C. McElree](http://hmg.oxfordjournals.org/search?author1=Colleen+McElree&sortspec=date&submit=Submit), [S. Targan](http://hmg.oxfordjournals.org/search?author1=Stephan+Targan&sortspec=date&submit=Submit), [J. I. Rotter,](http://hmg.oxfordjournals.org/search?author1=Jerome+I.+Rotter&sortspec=date&submit=Submit) [N. Fischel-Ghodsian](http://hmg.oxfordjournals.org/search?author1=Nathan+Fischel-Ghodsian&sortspec=date&submit=Submit) (1996) Susceptibility locus for inflammatory bowel disease on chromosome 16 has a role in Crohns disease, but not in Ulcerative Colotis. *Human Molecular Genetics*, 5: 1679-1683.

**1997 (3)**

[7] K. Lange, **H. Zhao**, T. P. Speed. (1997) Modeling chiasma interference using the Poisson-skip model. *Annals of Applied Probability*, 7: 299-313.

[8] D. R. Goldstein, **H. Zhao**, T. P. Speed. (1997) The effects of genotyping errors and interference on estimation of genetic distance. *Human Heredity*, 47: 86-100.

[9] **H. Zhao**, H. Zhang, J. I. Rotter. (1997) Cost-effective sib-pair designs in the mapping quantitative-trait loci. *American Journal of Human Genetics*, 60: 1211-1221.

**1998 (4)**

[10] K. K. Kidd, B. Morar, C. M. Castiglione, **H. Zhao**, A. J. Pakstis, W. C. Speed, B. Bonne-Tamir, R.-B. Lu, D. Goldman, C. Lee, Y. S. Nam, D. K. Grandy, T. Jenkins, J. R. Kidd. (1998) A global survey of haplotype frequencies and linkage disequilibrium at the DRD2 locus. *Human Genetics*, 103: 211-227.

[11] **H. Zhao**, T. P. Speed. (1998) Stochastic modeling of the crossover process during meiosis. *Communications in Statistics, Theory and Methods*, 27: 1557-1580.

[12] **H. Zhao**, T. P. Speed. (1998) Statistical analysis of ordered tetrads. *Genetics*, 150: 459-472.

[13] **H. Zhao**, T. P. Speed. (1998) Statistical analysis of half-tetrads. *Genetics*, 150: 473-485.

**1999 (2)**

[14] **H. Zhao**, A. J. Pakstis, J. R. Kidd, K. K. Kidd. (1999) Assessing linkage disequilibrium in a complex genetic system. *Annals of Human Genetics*, 63: 167-179.

[15] **H. Zhao**, K. R. Merikangas, K. K. Kidd. (1999) On a randomization test in linkage analysis. *American Journal of Human Genetics*, 65: 1449-1456.

**2000 (6)**

[16] J. R. Kidd, A. J. Pakstis, **H. Zhao**, R.-B. Lu, F. E. Okonofua, A. Odunsi, E. Grigorenko, B. Bonne-Tamir, J. Friedlaender, L. O. Schulz, J. Parnas, K. K. Kidd. (2000) Haplotypes and linkage disequilibrium at the phenylalanine hydroxylase locus (PAH) in a global representation of populations. *American Journal of Human Genetics*, 66: 1882-1899.

[17] S. Zhang, **H. Zhao**. (2000) Linkage disequilibrium mapping in populations of variable size using the decay of haplotype sharing and a stepwise-mutation model. *Genetic Epidemiology*, 19: S99-S105.

[18] K. Zhang, **H. Zhao**. (2000) Assessing reliability of gene clusters from gene expression data. *Functional and Integrative Genomics*, 1: 156-173.

[19] F. Sun, W. D. Flanders, Q. Yang, **H. Zhao**. (2000) Transmission/disequilibrium tests for quantitative traits. *Annals of Human Genetics*, 64: 555-565.

[20] **H. Zhao**, J. Li, W. P. Robinson. (2000) Multipoint genetic mapping with uniparental disomy data. *American Journal of Human Genetics*, 67: 851-861.

[21] **H. Zhao**, S. Zhang, K. R. Merikangas, M. Trixler, D. Wildenaur, F. Sun, K. K. Kidd. (2000) Transmission/disequilibrium tests using multiple tightly linked markers. *American Journal of Human Genetics*, 67: 936-946.

**2001 (8)**

[22] A. Kamina, R. W. Makuch, **H. Zhao**. (2001) Stochastic modeling of early HIV-1 population dynamics. *Mathematical Biosciences*, 170: 187-198.

[23] **H. Zhao**, F. Liang. (2001) On relationship inference using continuous gamete identity by descent data. *Journal of Computational Biology*, 8: 191-200.

[24] **H. Zhao**, J. Li, W. P. Robinson. (2001) Statistical analysis of uniparental disomy data using hidden Markov models. *Biometrics*, 57: 1074-1079.

[25] S. Zhang, **H. Zhao**. (2001) Quantitative similarity-based association test using population samples. *American Journal of Human Genetics*, 69: 601-614.

[26] S. Zhang, A. J. Pakstis, K. K. Kidd, **H. Zhao**. (2001) Comparisons of two methods for haplotype reconstruction and haplotype frequency estimates from population data. *American Journal of Human Genetics*, 69: 906-912.

[27] L. Ma, J. Li, J. Hager, L. Qu, Z. Chen, **H. Zhao**, X-W Deng. (2001) Light control of Arabidopsis development entails co-ordinated regulation of genome expression and cellular pathways. *Plant Cell*, 13: 2589-2607.

[28] H. Wang, L. Ma, J. Li, **H. Zhao**, X-W Deng. (2001) Direct interaction of Arabidopsis Cryptochromes with COP1 in light control development. *Science*, 294: 154-158.

[29] J. Li, S. Sherman, N. Lamb, **H. Zhao**. (2001) Multipoint genetic mapping with trisomy data. *American Journal of Human Genetics*, 69: 1255-1265.

**2002 (7)**

[30] S. Zhang, K. K. Kidd, **H. Zhao**. (2002) Detecting genetic association in case-control studies using similarity-based association tests. *Statistica Sinica*, 12: 337-359.

[31] S. Zhang, **H. Zhao**. (2002) Linkage disequilibrium mapping with genotype data. *Genetic Epidemiology*, 22: 66-77.

[32] X. Zhu, S. Zhang, **H. Zhao**, R. S. Cooper. (2002) Association mapping using a mixture model for complex traits. *Genetic Epidemiology*, 23: 181-196.

[33] A. Mani, S. M. Meraji, R. Houshyar, J. Radhakrishnan, A. Mani, M. Ahangar, T. M. Rezaie, M. A. Taghavinejad, B. Broumand, **H. Zhao**, C. Nelson-Williams, R. P. Lifton. (2002) Finding genetic contributions to sporadic disease: A recessive locus at 12q24 commonly contributes to patent ductus arteriosus. *Proceedings of the National Academy of Sciences*,99: 15054-15059.

[34] G. Wu, X. Tian, S. Nishimura, G. S. Markowitz, V. D'Agati, J. H. Park, L. Yao, L. Li, L. Geng, **H. Zhao**, W. Edelmann, S. Somlo. (2002) Trans-heterozygous Pkd1 and Pkd2 mutations modify expression of polycystic kidney disease. *Human Molecular Genetics*, 16: 1845-1854.

[35] L. Ma, Y. Gao, L. Qu, Z. Chen, J. Li, **H. Zhao**, X-W Deng. (2002) Genomic evidence for COP1 as a repressor of light-regulated gene expression and development in Arabidopsis. *Plant Cell*, 14: 2383-2398.

[36] H. Wang, L. Ma, J. Haboshi, J. Li, **H. Zhao**, X-W Deng. (2002) [Analysis of far-red light regulated genome expression profiles of phytochrome A pathway mutants in Arabidopsis](http://plantgenomics.biology.yale.edu/WM/WM.html). *Plant Journal*, 32: 723-733.

**2003 (18)**

[37] C. Dong, S. Wang, W-D Li, **H. Zhao**, R. A. Price. (2003) Interacting genetic loci in chromosome regions 20q and 10q influence extreme human obesity. *American Journal of Human Genetics*, 72: 115-124.

[38] S. Zhang, X. Zhu, **H. Zhao**. (2003) On a semi-parametric test to detect associations between quantitative traits and candidate genes using unrelated individuals. *Genetic Epidemiology*, 24: 44-56.

[39] S. Wang, K. K. Kidd, **H. Zhao**. (2003) On the use of DNA pooling to estimate haplotype frequencies. *Genetic Epidemiology*, 24: 74-82.

[40] L. Ma, **H. Zhao**, X-W Deng (2003) Analysis of the mutational effects of the COP/DET/FUS loci on genome expression profiles reveals their overlapping yet not identical roles in regulating Arabidopsis seedling development. *Development*, 130: 969-981.

[41] H-S Chen, X. Zhu, **H. Zhao**, S. Zhang. (2003) Qualitative semi-parametric test to detect genetic association in case-control design under structured population. *Annals of Human Genetics*, 67: 250-264.

[42] G. Zou, D. Pan, **H. Zhao**. (2003) Genotyping error detection through tightly linked markers. *Genetics*, 164: 1161-1173.

[43] B. Wu, T. Abbott, D. Fishman, W. McMurray, G. Mor, K. Stone, D. Ward, K. Williams, **H. Zhao**. (2003) Comparison of statistical methods for classification of ovarian cancer using a proteomics dataset. *Bioinformatics*, 19: 1636-1643.

[44] S. Wang, **H. Zhao**. (2003) Sample size to detect gene-gene interactions using association designs. *American Journal of Epidemiology*, 158: 899-914.

[45] W-D Li, D. Li, S. Wang, S. Zhang, **H. Zhao**, R. A. Price. (2003) Linkage and linkage disequilibrium mapping of genes influencing human obesity in chromosome region 7q22.1–7q35. *Diabetes*, 52: 1557-1561.

[46] N. M. Wikonkal, E. Remenyik, D. Knezevic, W. Zhang, M. Liu, **H. Zhao**, T. R. Berton, D. G. Johnson, D. E. Brash. (2003) Inactivating E2f1 reverts apoptosis resistance and cancer sensitivity in Trp53-deficient mice. *Nature Cell Biology* 5: 655-660.

[47] X. Luo, H. R. Kranzler, **H. Zhao**, J. Gelernter. (2003) Haplotypes at the OPRM1 locus are associated with susceptibility to substance dependence in European-Americans. *American Journal of Medical Genetics* 120B: 97-108.

[48] L. V. Sun, L. Chen, F. Greil, N. Negre, T.-R. Li, G. Cavalli, **H. Zhao**, B. Van Steensel, K. P. White (2003) DNA–protein interaction mapping using genomic tiling path microarrays in Drosophila. *Proceedings of the National Academy of Sciences*, 100: 9428-9433.

[49] G. Zou, **H. Zhao**. (2003) Haplotype frequency estimation in the presence of genotyping errors. *Human Heredity*, 56: 131-138.

[50] G. Garcia-Rostan, **H. Zhao**, M. Pollan, A. Herrero, R. L. Camp, J. Pardo, R. Wu, J. Costa, M. L. Carcangiu, G. Tallini. (2003) Ras mutations are associated with aggressive tumor phenotypes and poor prognosis in thyroid cancer. *Journal of Clinical Oncology*, 21: 3226-3235.

[51] Y. Jiao, H. Yang, L. Ma, N. Sun, H. Yu, T. Liu, Y. Gao, H. Gu, Z. Chen, M. Wada, M. Gerstein, **H. Zhao**, L. Qu, X. W. Deng. (2003) [A genome wide analysis of blue light regulation of Arabidopsis transcription factor gene expression during seedling development.](http://plantgenomics.biology.yale.edu/Jiao/LR.html) *Plant Physiology*, 133: 1480-1493.

[52] L. Zheng, S. Wang, P. Romans, **H. Zhao**, C. Luna, M. Q. Benedict (2003) Quantitative trait loci in Anopheles gambiae controlling the encapsulation response against Plasmodium cynomolgi Ceylon. *BMC Genetics* 2003, 4:16

[53] D. Pan, N. Sun, K-H Cheung, Z. Guan, L. Ma, M. Holford, X-W Deng, **H. Zhao** (2003) PathMAPA: a tool for displaying gene expression and performing statistical tests on metabolic pathways at multiple levels for Arabidopsis. *BMC Bioinformatics*, 4:56.

[54] N. Sun, L. Ma, D. Pan, **H. Zhao**, X-W Deng. (2003) Evaluation of regulatory potential of Calvin cycle pathway steps based on large-scale gene expression profiling data. *Plant Molecular Biology*, 53: 467-478.

**2004 (17)**

[55] G. Zou, **H. Zhao**. (2004) The impacts of errors in individual genotyping and DNA pooling on association studies. *Genetic Epidemiology*, 26: 1-10.

[56] W. Zhao, J. Wang, X. He, X. Huang, Y. Jiao, M. Dai, S. Wei, J. Fu, Y. Chen, X. Ren, Y. Zhang, P. Ni, J. Zhang, S. Li, J. Wang, **G. Wong**, **H. Zhao**, J. Yu, H. Yang, J. Wang. (2004) BGI-RIS: an integrated information resource and comparative analysis workbench for rice genomics. *Nucleic Acid Research*, 32: D377-382.

[57] T. N. Ferraro, G. T. Golden, G. G. Smith, J. F. Martin, F. W. Lohoff, T. A. Gieringer, D. Zamboni, C. L. Schwebel1, D. M. Press, S. O. Kratzer, **H. Zhao**, W. H. Berrettini, R. J. Buono. (2004) Fine mapping of a seizure susceptibility locus on mouse chromosome 1: Nomination of *Kcnj10* as a causative gene. *Mammalian Genome*, 15: 239-251.

[58] C.-S. Goh, N. Lan, S. Douglas, B. Wu, N. Echols, A. Smith, D. Milburn, G. T. Montelione, **H. Zhao**, M. Gerstein. (2004) Mining the structural genomics pipeline: Identification of protein properties that affect high-throughput experimental analysis. *Journal of Molecular Biology*, 336: 115-130.

[59] X. Sun, J. C. Stephens, **H. Zhao**. (2004) The impact of sample size and marker selection on the study of haplotype structures. *Human Genomics*, 1: 179-193.

[60] G. Zou, **H. Zhao**. (2004) The estimation of sibling genetic risk parameters revisited. *Genetic Epidemiology*, 26: 286-293.

[61] W. D. Li, C. Dong, D. Li, **H. Zhao**, R. A. Price. (2004) An obesity-related locus in chromosome region 12q23-24. *Diabetes* 53: 812-820.

[62] R. J. Buono , F. W. Lohoff, T. Sander, M. R. Sperling, M. J. O'Connor, D. J. Dlugos, S. G. Ryan, G. T. Golden, **H. Zhao**, T. M. Scattergood, W. H. Berrettini, T. N. Ferraro. (2004) Association between variation in the human KCNJ10 potassium ion channel gene and seizure susceptibility. *Epilepsy Research*, 58: 175-183*.*

[63] R. Marathe, Z. Guan, R. Anandalakshmi, **H. Zhao**, S. Dinesh-Kumar (2004) Study of Arabidopsis thaliana resistome in response to cucumber mosaic virus infection using whole genome microarray. *Plant Molecular Biology*, 55: 501-520*.*

[64] Hoek, K., Rimm, D.L., Williams, K.R., **Zhao, H**., Ariyan, S., Lin, A., Kluger, H.M., Berger, A.J., Cheng, E., Trombetta, E.S.,Wu, T., Niinobe, M., Yoshikawa, K., Hannigan, G.E., Halaban, R. (2004) Expression profiling reveals novel pathways in the transformation of melanocytes to melanomas. *Cancer Research*, 64: 5270-5282.

[65] Y. Lai, B. Wu, L. Chen, **H. Zhao** (2004) Statistical method for identifying differential gene-gene coexpression patterns.*Bioinformatics*, 20: 3146-3155*.*

[66] N. Liu, S. L. Sawyer, N. Mukherjee, A. J. Pakstis, J. R. Kidd, K. K. Kidd, A. J.Brookes, **H. Zhao** (2004) Haplotype block structures show significant variation among populations.*Genetic Epidemiology*, 27: 385-400*.*

[67] Y. Gao, J. Li, E. Strickland, S. Hua, **H. Zhao**, Z. Chen, L. Qu, X. W. Deng (2004) An Arabidopsis promoter microarray and its initial usage in the identification of HY5 binding targets in Vitro. *Plant Molecular Biology*,54: 683-699.

[68] M. V. Osier, **H. Zhao**, K-H Cheung (2004) Handling multiple testing while interpreting microarrays with the gene ontology database. *BMC Bioinformatics*, 5, 124.

[69] N. Lin, B. Wu, R. Jansen, M. Gerstein, **H. Zhao** (2004) Information assessment on predicting protein-protein interactions. *BMC Bioinformatics*, 5, 154.

[70] Y. Liu, **H. Zhao** (2004) A computational approach for ordering signal transduction pathway components from genomics and proteomics data. *BMC Bioinformatics*, 5, 158.

[71] F. H. Wilson, A. Hariri, A. Farhi, **H. Zhao**, K. F. Petersen, H. R. Toka, C. Nelson-Williams, K. M. Raja, M. Kashgarian, G. I. Shulman, S. J. Scheinman, R. P. Lifton RP (2004) A cluster of metabolic defects caused by mutation in a mitochondrial tRNA. *Science*, 306: 1190-1194.

**2005 (21)**

[72] N. Carriero, M. V. Osier, K-H Cheung, P. L. Miller, M. Gerstein, **H. Zhao**, B. Wu, S. Rifkin, J. Chang, H. Zhang, K. White, K. Williams, M. Schultz (2005) A 'High Productivity/Low Maintenance' approach to high performance computation for biomedicine: Four Case Studies. *Journal of American Medical Informatics Association*, 12: 90-98*.*

[73] K. Zhang, F. Sun, **H. Zhao**. (2005) HAPLORE: A program for haplotype reconstruction in general pedigrees without recombination. *Bioinformatics*, 21: 90-103.

[74] G. Zou, **H. Zhao** (2005) Family-based association tests for different family structures using pooled DNA. *Annals of Human Genetics,* 69: 429-442.

[75] Z. Guan, **H. Zhao** (2005) A semiparametric approach for marker gene selection based on gene expression data. *Bioinformatics,* 21: 529-536.

[76] L. Chen, **H. Zhao** (2005) Integrating mRNA decay information into co-regulation study. *Journal of Computer Science and Technology,* 20: 434-438.

[77] M. Holford, N. Li, P. Nadkarni, **H. Zhao** (2005) VitaPad: visualization tools for the analysis of pathway data. *Bioinformatics,* 21: 1596-1602.

[78] B. S. Kennedy, S. V. Kasl, J. Lichtman, **H. Zhao** (2005) Predicting readmission stroke type among blacks and whites in California. *Journal of Stroke and Cerebrovascular Disease*,14: 251-260.

[79] Y. Lai, **H. Zhao** (2005) A statistical method to detect chromosomal regions with DNA copy number alterations using SNP-array-based CGH data. *Computational Biology and Chemistry*, 29: 47-54*.*

[80] B. Z. Yang, **H. Zhao**, H. R. Kranzler, J. Gelernter (2005) Practical population group assignment with selected informative markers: Characteristics and properties of Bayesian clustering via STRUCTURE. *Genet Epidemiology,* 28: 302-312.

[81] S. S. Zhang, X. Xu, J. Li, M. G. Liu, **H. Zhao**, M. B. Soares, C. J. Barnstable, X. Y. Fu (2005) Comprehensive in silico functional specification of mouse retina transcripts. *BMC Genomics*, 6: 40.

[82] L. Chen, **H. Zhao** (2005) Gene expression analysis reveals that histone deacetylation sites may serve as partitions of chromatin gene expression domains.*BMC Genomics*, 6: 44.

[83] G. Mor, I. Visintin, Y. Lai , **H. Zhao**, P. Schwartz, T. Rutherford, L. Yue, P. Bray-Ward, D. Ward (2005) Serum protein markers for early detection of ovarian cancer. *Proceedings of the National Academy of Sciences,* 102: 7677-7682*.*

[84] Y. Jiao, P. Jia, X. Wang, N. Su, S. Yu, D. Zhang, L. Ma, Q. Feng, Z. Jin, L. Li, Y. Xue, Z. Cheng, **H. Zhao**, B. Han, X. W. Deng (2005) A tiling microarray expression analysis of rice chromosome 4 suggests a chromosomal level regulation of transcription. *Plant Cell,* 17: 1641-1657.

[85] N. Lin, **H. Zhao** (2005) Are scale-free networks robust to measurement errors? *BMC Bioinformatics*,6: 119.

[86] Y. Liu, N. Liu, **H. Zhao** (2005) Inferring protein-protein interactions through high throughput interaction data from diverse organisms. *Bioinformatics,* 21: 3279-3285.

[87] L. Ma, N. Sun, X. Liu, Y. Jiao, **H. Zhao**, X. W. Deng (2005) Organ-specific expression of Arabidopsis genome during development. *Plant Physiology,* 138: 80-91.

[88] F. W. Lohoff, T. N. Ferraro, T. Sander, **H. Zhao**, J. P. Dahl, W. H. Berrettini, R. J. Buono (2005) No association between common variations in the human alpha 2 subunit gene (ATP1A2) of the sodium-potassium-transporting ATPase and idiopathic generalized epilepsy. *Neuroscience Letters*,382:33-38.

[89] L. Chen, **H. Zhao** (2005) Negative correlation between compositional symmetries and local recombination rates. *Bioinformatics*, 21: 3951-1958*.*

[90] T. Huang, B. Wu, P. Lizardi, **H. Zhao** (2005) Detection of DNA copy number alterations using penalized least squares regression. *Bioinformatics,* 21: 3811-3817*.*

[91] B-Z Yang, **H. Zhao**, H. R. Kranzler, J. Gelernter (2005) Characterization of a likelihood based method and effects of markers informativeness in evaluation of admixture and population group assignment. *BMC Genetics,* 6:50*.*

[92] L. Ma, C. Chen, X. Liu, Y. Jiao, N. Su, L. Li, X. Wang, M. Cao, N. Sun, X. Zhang, J. Bao, J. Li, S. Pedersen, L. Bolund, **H. Zhao**, L. Yuan, G. K-S Wong, J. Wang, X. W. Deng, J. Wang (2005) A microarray analysis of the rice transcriptome and its comparison to Arabidopsis. *Genome Research,* 15: 1274-1283.

**2006 (18)**

[93] B. Wu, T. Abbott, D. Fishman, W. McMurray, G. Mor, K. Stone, D. Ward, K. Williams, **H. Zhao** (2006) Ovarian cancer classification based on mass spectrometry analysis of sera. *Cancer Informatics,* 2: 123-132*.*

[94] W. Yu, B. Wu, N. Lin, K. Stone, K. Williams, **H. Zhao** (2006) Detecting and aligning peaks in mass spectrometry data with applications to MALDI. *Computational Biology and Chemistry,* 30: 27-38.

[95] W. Yu, X. Li, J. Liu, B. Wu, K. Williams, **H. Zhao** (2006) Multiple peak alignment in sequential data analysis: A scale-space based approach. *IEEE/ACM Trans. Computational Biology and Bioinformatics* 3: 208-219*.*

[96] S. Wang, S. Huang, L. Zheng, **H. Zhao** (2006) Mapping quantitative trait loci in non-inbred mosquito crosses. *Genetics,* 172: 2293-2308*.*

[97] B. Wu, Z. Guan, **H. Zhao** (2006) Parametric and nonparametric FDR estimation revisited. *Biometrics,* 62: 735-744*.*

[98] N. Liu, I. Beerman, R. Lifton, **H. Zhao** (2006) Haplotype analysis in the presence of informatively missing genotype data. *Genetic Epidemiology,* 30: 290-300*.*

[99] J. Gelernter, C. Panhuysen, M. Wilcox, V. Hesselbrock, B. Rounsaville, J. Poling, R. Weiss, S. Sonne, **H. Zhao**, L. Farrer, H. R. Kranzler (2006) Genomewide linkage scan for opioid dependence and related traits. *American Journal of Human Genetics,* 78: 759-769*.*

[100] B. E. Enerson, A. Lin, B. Lu, **H. Zhao**, M. P. Lawton, E. Floyd (2006) Acute drug-induced vascular injury in beagle dogs: pathology and correlating genomic expression. *Toxicology Pathology,* 34: 27-32.

[101] K. Zhang, **H. Zhao** (2006) A comparison of several methods for haplotype frequency estimation and haplotype reconstruction for tightly linked markers from general pedigrees. *Genetic Epidemiology,* 30: 423-437*.*

[102] N. Sun, R. Carroll, **H. Zhao** (2006) Bayesian Error Analysis Model (BEAM) for reconstructing transcriptional regulatory networks. *Proceedings of the National Academy of Sciences* 103: 7988-7993*.*

[103] Y. Zuo, G. Zou, **H. Zhao** (2006) Two-stage designs in case-control association analysis. *Genetics,* 173: 1747-1760*.*

[104] T. Tsubouchi, **H. Zhao**, G. S. Roeder (2006) Crossover distribution by promoting synaptonemal complex formation together with Zip2. *Developmental Cell,* 10: 809-819*.*

[105] X. You, M. Pan, W. Gao, H. S. Shiah, J. Tao, D. Zhang, F. Koumpouras, S. Wang, **H. Zhao**, J. A. Madri, D. Baker, Y. C. Cheng, Z. Yin (2006) Effects of a novel tylophorine analog on collagen-induced arthritis through inhibition of the innate immune response. *Arthritis & Rheumatism*, 54: 877-886.

[106] N. Liu, **H. Zhao** (2006) A nonparametric approach for population structure inference using multilocus genotypes. *Human Genomics*, 2: 353-364.

[107] B. Wu, N. Liu, **H. Zhao** (2006) PSMIX: an R package for population stratification inference via maximum likelihood method. *BMC Bioinformatics,* 7: 317.

[108] H. Pang, A. Lin, M. Holford, B. E. Enerson, B. Lu, M. P. Lawton, E. Floyd, **H. Zhao** (2006) Pathway analysis using random forests classification and regression. *Bioinformatics,* 22: 2028-2036.

[109] S. S. Zhang, X. Xu, M. G. Liu, **H. Zhao**, M. B. Soares, C. J. Barnstable, X. Y. Fu (2006) A biphasic pattern of gene expression during mouse retina development. *BMC Developmental Biology*, 6: 48.

[110] S. P. Wu, L. Leng, Z. Feng, N. Liu, **H. Zhao**, C. McDonald, A. Lee, F. C. Arnett, P. K. Gregersen, M. D. Mayes, R. Bucala R (2006) Macrophage migration inhibitory factor promoter polymorphisms and the clinical expression of scleroderma. *Arthritis & Rheumatism*, 54: 3661-3669.

**2007 (8)**

[111] I. Kim, Y. Liu, **H. Zhao.** (2007) Bayesian methods for predicting interacting protein pairs using domain information. *Biometrics,* 63: 824-833*.*

[112] V. Dinu, P. Miller, **H. Zhao**. (2007) Evidence for association between multiple complement pathway genes and AMD. *Genetic Epidemiology,* 31: 224-237.

[113] T-S Lee, S. Mane, T. Eid, **H. Zhao**, A. Lin, Z. Guan, J. H. Kim, J. Schweitzer, D. King-Stevens, P. Weber, S. S. Spencer, D. D. Spencer, N. C. De Lanerolle. (2007) Gene expression in temporal lobe epilepsy is consistent with increased release of glutamate by astrocytes. *Molecular Medicine*, 13: 1-13.

[114] J. Lee, K. He, V. Stolc, H. Lee, P. Figueroa, Y. Gao, W. Tongprasit, **H. Zhao**, I. Lee, X. W. Deng (2007) Analysis of transcription factor HY5 genomic binding sites revealed its hierarchical role in light regulation of development. *Plant Cell,* 19: 731-749*.*

[115] S. Wang, **H. Zhao** (2007) Sample size needed to detect gene-gene interactions using linkage analysis. *Annals of Human Genetics,* 71: 828-842*.*

[116] W. Yu, J. Liu, C. Colangelo, E. Gulcicek, **H. Zhao** (2007) A new protocol of analyzing isotope coded affinity tag data from high resolution LC-MS spectrometry. *Computational Biology and Chemistry,* 31: 215-221*.*

[117] V. Dinu, **H. Zhao**, P. Miller (2007) Integrating domain knowledge with statistical and data mining methods for high density genomic SNP disease association analysis. *Journal of Biomedical Informatics,* 40: 750-760*.*

[118] B. Z. Yang, H. R. Kranzler, **H. Zhao**, J. R. Gruen, X. Luo, J. Gelernter (2007) Association of haplotypic variants in DRD2, ANKK1, TTC12 and NCAM1 to alcohol dependence in independent case-control and family samples. *Human Molecular Genetics*,16: 2844-2853*.*

**2008 (15)**

[119] Z. Guan, B. Wu, **H. Zhao** (2008) Nonparametric estimator of false discovery rate based on Bernstein polynomials. *Statistica Sinica,* 18: 905-923.

[120] D. Datta, **H. Zhao** (2008) Statistical methods to infer cooperative binding among transcription factors in Saccharomyces cerevisiae. *Bioinformatics*, 24: 545-552.

[121] Y. Zuo, G. Zou, J. Wang, **H. Zhao**, H. Liang (2008) Optimal two-stage design for case-control association analysis incorporating genotyping errors. *Annals of Human Genetics*,72: 375-387.

[122] S. Budel, S-O Shim, Z. Feng, **H. Zhao**, F. Hisama, S. M. Strittmatter (2008) No association between schizophrenia and polymorphisms of the PlexinA2 gene in Chinese Han Trios. *Schizophrenia Research*, 99: 365-366.

[123] W. Yu, Z. He, J. Liu, **H. Zhao** (2008) Improving mass spectrometry peak detection using multiple peak alignment results. *Journal of Proteome Research*, 7: 123-129.

[124] J. Liu, W. Yu, B. Wu, **H. Zhao** (2008) Bayesian mass spectra peak alignment from mass charge ratios. *Cancer Informatics*, 4: 217-241.

[125] T. Tong, **H. Zhao** (2008) Practical guidelines for assessing power and false discovery rate for a fixed sample size in microarray experiments. *Statistics in Medicine*, 27: 1960-1972.

[126] J. Liu, **H. Zhao**, J. Tan, D. Luo, W. Yu, E. J. Harner (2008) Is subcellular localization informative for modeling protein-protein interaction signal? *Research Letters in Signal Processing*, Volume 2008, Issue 1, Article 2.

[127] R. A. Price, W.-D. Li, **H. Zhao** (2008) FTO gene associated with extreme obesity in cases, controls and extremely discordant sister pairs. *BMC Medical Genetics*, 9, 4.

[128] H. Pang, **H. Zhao** (2008) Building pathway clusters from Random Forests classification using class votes. *BMC Bioinformatics*, 9, 87.

[129] W. Ji, J. N. Foo, B. J. O’Roak, **H. Zhao**, M. G. Larson, D. B. Simon, C. Newton-Cheh, M. W. State, D. Levy, R. P. Lifton (2008) Rare independent mutations in renal salt handling genes contribute to blood pressure variation. *Nature Genetics*, 40: 592-599.

[130] J. Cao, **H. Zhao** (2008) Estimating dynamic models for gene regulation networks. *Bioinformatics*, 24: 1619-1624.

[131] L. Chen, T. Tong, **H. Zhao** (2008) Considering dependence among genes and markers for false discovery control in eQTL mapping. *Bioinformatics*, 24: 2015-2022.

[132] B. Xiang, A. Li, D. Valentin, N. J. Nowak, **H. Zhao**, P. Li (2008) Analytical and clinical validity of whole-genome oligonucleotide array comparative genomic hybridization for pediatric patients with mental retardation and developmental delay. *American Journal of Medical Genetics A.* 146A: 1942-1954.

[133] S. Budel, T. Padukkavidana, B. Liu, Z. Feng, F. Hu, S. Johnson, J. Lauren, J. Park, A. McGee, J. Liao, A. Stillman, J-E Kim, S. Sodi, B-Z Yang, J. Gelernter, **H. Zhao**, F. Hisama, A. Arnsten, S. Strittmatter (2008) Genetic variants of Nogo-66 Receptor with possible association to schizophrenia block myelin inhibition of axon growth. *Journal of Neuroscience,* 28: 13161-13172.

**2009 (15)**

[134] Y Jiao, S. L. Tausta, N. Gandotra, N. Sun, T. Liu,N. Clay, T. Ceserani, M. Chen, L. Ma, M. Holford, **H. Zhao**, X. Deng, T. Nelson (2009) A transcriptome atlas of rice cell types reveals cellular, functional and developmental hierarchies. *Nature Genetics*, 41: 258-263.

[135] N. Liu, D. Zhang, **H. Zhao** (2009) Genotyping error detection in samples of unrelated individuals without replicate genotyping. *Human Heredity*, 67: 154-162.

[136] B. Z. Yang, H. R. Kranzler, **H. Zhao**, J. R. Gruen, X. Luo, J. Gelernter (2008) Haplotypic variants in DRD2, ANKK1, TTC12, and NCAM1 are associated with comorbid alcohol and drug dependence. *Alcoholism: Clinical and Experimental Research,* 32: 2117-2127.

[137] M. E. Holford, H. Rajeevan, **H. Zhao**, K. K. Kidd, K-H Cheung (2009) Semantic web-based integration of cancer pathways and allele frequency data. *Cancer Informatics,* 8: 19-30*.*

[138] H. Pang, T. Tong, **H. Zhao** (2009) Shrinkage-based diagonal discriminant analysis and its applications in high-dimensional data. *Biometrics,* 65: 1021-1029*.*

[139] N. Liu, R. Bucala, **H. Zhao** (2009) Modeling informatively missing genotypes in haplotype analysis. *Communications in Statistics - Theory and Methods*, 38: 3445-3460.

[140] D. Datta, **H. Zhao** (2009) Effect of false positive and false negative rates on inference of binding target conservation across different conditions and species from ChIP-chip data. *BMC Bioinformatics,* 10, 23*.*

[141] H. R. Kranzler, J. Gelernter, R. F. Anton, A. J. Arias, A. Herman,**H. Zhao**, L. Burian, J. Covault (2009) Association of markers in the 3’ region of the GluR5 kainate receptor subunit gene (*GRIK1*) to alcohol dependence. *Alcoholism: Clinical and Experimental Research,* 33: 925-930*.*

[142] J. Sampson, **H. Zhao** (2009) Genotyping and inflated type I error rate in genome-wide association case/control studies. *BMC Bioinformatics,* 10, 68*.*

[143] X. Li, **H. Zhao** (2009) Weighted random subspace method for high dimensional data classification. *Statistics and Its Interface,* 2: 153-159*.*

[144] N. Papeta, K. T. Chan, S. Prakash, J. Martino, K. Kiryluk, D. Ballard, L. A. Bruggeman, R. Frankel, Z. Zheng, P. E. Klotman, **H. Zhao**, V. D. D'Agati, R. P. Lifton, A. G. Gharavi (2009) Susceptibility loci for murine HIV-associated nephropathy encode trans-regulators of podocyte gene expression. *Journal of Clinical Investigation,* 119: 1178-1188.

**[**145] K. He, X. Li, J. Zhou, XW Deng, **H. Zhao**, J. Luo J (2009) NTAP: For NimbleGen tiling array ChIP-chip data analysis. *Bioinformatics,* 25: 1838-1840.

[146] Z. Wu, **H. Zhao** (2009) Statistical power of model selection strategies for genome-wide association studies. *PLoS Genetics,* 5: e1000582.

[147] J. Sampson, **H. Zhao** (2009) Identifying individuals in complex mixture of DNA with unknown ancestry. *Statistical Applications in Genetics and Molecular Biology,* 8, Article 37.

[148] J. Liu, L. Chen, **H. Zhao**, D. F. Moore, Y. Lin, W. J. Shi (2009) On calculating the probability of a set of orthologous sequences. *Journal of Advances and Applications in Bioinformatics and Chemistry*,2: 37-48.

**2010 (16)**

[149] X. Qi, **H. Zhao** (2010) Asymptotic efficiency and finite-sample properties of the generalized profiling estimation of the parameters in ordinary differential equations. *Annals of Statistics,* 38: 435-481.

[150] D. Ballard, J. Cho, **H. Zhao** (2010) Comparisons of multi-marker association methods to detect association between a candidate region and disease. *Genetic Epidemiology,* 34: 201-212.

[151] J. Kang, J. Cho, **H. Zhao** (2010) Practical issues in building risk predicting models for complex diseases. *Journal of Biopharmaceutical Statistics*, 20: 415-440.

[152] R. Luo, C. M. Colangelo,W. C. Sessa,**H. Zhao** (2010) Bayesian analysis of iTRAQ data with nonrandom missingness: Identification of differentially expressed proteins. *Statistics in BioSciences*, 1: 228-245.

[153] S. Huang, T. Tong, **H. Zhao** (2010) Bias-corrected diagonal discriminant rules for high-dimensional classification. *Biometrics*, 66: 1096-1106.

[154] H. Pang, D. Datta, **H. Zhao** (2010) Pathway analysis using random forests with bivariate node-split for survival outcomes. *Bioinformatics*, 26: 250-258.

[155] W. Zheng, **H. Zhao**, E. Mancera, L. Steinmetz, M. Snyder (2010) Genetic analysis of variation in transcription factor binding in yeast. *Nature,* 464: 1187-1191.

[156] D. H. Ballard, J. Zhu, E. E. Schadt, **H. Zhao** (2010) Improving detection of genes associated with clinical traits using expression traits. *International Journal of Systems and Synthetic Biology, in press*.

[157] J. X. Hu, **H. Zhao**, H. H. Zhou (2010) False discovery rate control with groups. *Journal of American Statistical Association*, 105: 1215–1227.

[158] L. F. Mustavich, P. Miller, K. K. Kidd, **H. Zhao** (2010) Using a pharmacokinetic model to relate an individual's susceptibility to alcohol dependence to genotypes. *Human Heredity*, 70: 177-193.

[159] D. Ballard, C. Abraham, J. Cho, **H. Zhao** (2010) Pathway analysis comparison using Crohn's disease GWAS. *BMC Medical Genomics*, 3, 25.

[160] A. Davalos, C. Fernandez-Hernando, G. Sowa, B. Derakhshan, M. I. Lin, J. Y. Lee, **H. Zhao**, R. Luo, C. Colangelo, W. C. Sessa (2010) Quantitative proteomics of caveolin-1 regulated proteins: Characterization of PTRF/Cavin-1 in endothelial cells. *Mol Cell Proteomics*,9: 2109-2124.

[161] N. Santoro, R. Kursawe, E. D’Adamo, D. J. Dykas, C. K. Zhang, A. E. Bale, A. M. Cali, D. Narayan, M. M. Shaw, B. Pierpont, M. Savoye, D. Lartaud, S. Eldrich, S. W. Cushman, **H. Zhao**, G. I. Shulman, S. Caprio (2010) A Common variant in the Patatin-Like Phopholipase 3 Gene (*PNPLA3*) is associated with fatty liver disease in obese children and adolescents. *Hepatology*, 52: 1281-1290.

[162] C. Berger, K. Hoffmann1, J. G. Vasquez, S. Mane, J. Lewis, R. Filler, A. Lin, **H. Zhao**, T. Durazzo, A. Baird, W. Lin, F. Foss, I. Christensen, M. Girardi, R. Tigelaar, R. Edelson (2010) Rapid generation of maturationally synchronized human dendritic cells: Contribution to the clinical efficacy of extracorporeal photochemotherapy. *Blood,* 116: 4838-4847.

[163] P. K. Mistry, J. Liu, M. Yang, T. Nottoli, J. McGrath, D. Jain, K. Zhang, J. Keutzer, W-L Chuang, W. Z. Mehal, **H.**

**Zhao**, A. Lin, S. Mane, X. Liu, Y. Z. Peng, J. H. Li, M. Agrawal, L-L Zhu, H. C. Blair, L. J. Robinson, J. Iqbal, L. Sun,

M. Zaidi (2010) The GBA1 deficient mouse recapitulates Gaucher disease displaying system-wide cellular and molecular dysregulation beyond the macrophage. *Proceedings of the National Academy of Sciences,* 107: 19473-19478.

[164] R. D. Beech, L. Lowthert, J. J. Leffert, P. N. Mason, M. M. Taylor, S. Umlauf, A. Lin, J. Y. Lee, K. Maloney, A. Muralidharan, B. Lorberg, **H. Zhao**, S. S. Newton, S. Mane, C. N. Epperson, R. Sinha, H. Blumberg, Z. Bhagwagar. (2010) Increased peripheral blood expression of electron transport chain genes in bipolar depression. *Bipolar Disorders*, 12: 813-824.

**2011 (20)**

[165] R. Luo, **H. Zhao** (2011) Bayesian hierarchical modeling for signaling pathway inference from single cell interventional data. *Annals of Applied Statistics,* 5: 725–745.

[166] M. Chen, J. Cho, **H. Zhao** (2011) Detecting epistatic SNPs associated with complex diseases via a

Bayesian classification tree search method. *Annals of Human Genetics,* 75: 112-121.

[167] M. Chen, J. Cho, **H. Zhao** (2011) Incorporating biological pathways via a Markov random field model in genome-wide association studies. *PLoS Genetics*, 7: e1001353.

[168] X. Qi, **H. Zhao** (2011) [Some theoretical properties of Silverman's method for smoothed functional principal component analysis.](http://www.ncbi.nlm.nih.gov/pubmed/21516205) *Journal of Multivariate Analysis*, 102: 741-767.

[169] D. Li, **H. Zhao**, J. Gelernter (2011) [Strong association of the Alcohol Dehydrogenase 1B Gene (ADH1B) with alcohol dependence and alcohol-induced medical diseases.](http://www.ncbi.nlm.nih.gov/pubmed/21497796) *Biological Psychiatry*, 70: 504-512.

[170] H. Ma, E. E. Schadt, L. M. Kaplan, **H. Zhao** (2011) [COSINE: COndition-SpecIfic sub-NEtwork identification using a global optimization method.](http://www.ncbi.nlm.nih.gov/pubmed/21414987) *Bioinformatics*, 27: 1290-1298.

[171] X. Chang, F. Liu, X. Wang, A. Lin, **H. Zhao**, B. Su (2011) [The kinases MEKK2 and MEKK3 regulate transforming growth factor-β-mediated helper T cell differentiation.](http://www.ncbi.nlm.nih.gov/pubmed/21333552) *Immunity*, 34: 201-212.

[172] J. S. Lee, M. Choi, X. Yan, R. P. Lifton, **H. Zhao** (2011) [On optimal pooling designs to identify rare variants through massive resequencing.](http://www.ncbi.nlm.nih.gov/pubmed/21254222) *Genetic Epidemiology*, 35: 139-147.

[173] T. N. Ferraro, G. G. Smith, D. Ballard, **H. Zhao**, C. L. Schwebel, A. Gupta, E. F. Rappaport, S. E. Ruiz, F. W. Lohoff, G. A. Doyle, W. H. Berrettini, R. J. Buono (2011) [Quantitative trait loci for electrical seizure threshold mapped in C57BLKS/J and C57BL/10SnJ mice.](http://www.ncbi.nlm.nih.gov/pubmed/21129161) *Genes Brain Behavior*, 10: 309-315.

[174] H. Chun, D. Ballard, J. Cho, **H. Zhao** (2011) Identification of association between disease and multiple markers via sparse partial least squares regression. *Genetic Epidemiology*, 35: 479-486.

[175] J. Sampson, K. K. Kidd, J. R. Kidd, **H. Zhao** (2011) Select SNPs to identify ancestry. *Annals of Human Genetics*, 75: 539-553.

[176] N. Liu, **H. Zhao**, A. Patki, N. A. Limdi, D. B. Allison (2011) Controlling population structure in human genetic association studies with samples of unrelated individuals. *Statistics and Its Interface*, 4: 317-326.

[177] Y. Lai, B. Wu, **H. Zhao** (2011) A permutation test approach to the choice of size k for the nearest neighbors classifier. *Journal of Applied Statistics*, 8: 2289-2302.

[178] J. Kang, S. Kugathasan, M. Georges, **H. Zhao**, J. H. Cho; the NIDDK IBD Genetics Consortium (2011) Improved risk prediction for Crohn's disease with a multi-locus approach. *Human Molecular Genetics*, 20: 2435-2442.

[179] K. Wang, W-D Li, C. K. Zhang, Z. Wang, J. T. Glessner, S. F. A. Grant, **H. Zhao**, H. Hakonarson, R. A. Price (2011) A genome-wide association study on obesity and obesity-related traits. *PLoS One*, 6: e18939.

[180] F. Qian, X. Wang, L. Zhang, A. Lin, **H. Zhao**, E. Fikrig, R. R. Montgomery (2011) Impaired interferon signaling in dendritic cells from older donors infected in vitro with West Nile virus. *Journal of Infectious Disease* 203: 1415-1424.

[181] P. Xie, H. R. Kranzler, M. Krauthammer, K. P. Cosgrove, D. Oslin, R. F. Anton, L. A. Farrer, M. R. Picciotto, J. H. Krystal, **H. Zhao**, J. Gelernter (2011) Rare nonsynonymous variants in alpha-4 nicotinic acetylcholine receptor gene protect against nicotine dependence. *Biological Psychiatry,* 70: 528-536

[182] W. Zheng, L. Chung, **H. Zhao** (2011) Bias detection and correction in RNA-sequencing data. *BMC Bioinformatics,* 12: 290*.*

[183] M. Kato, X. Chen, S. Inukai, **H. Zhao**, F. J. Slack (2011) Age-associated changes in expression of small, noncoding RNAs, including microRNAs, in C. elegans. *RNA,* 17: 1804-1820.

[184] L. Zuo, C. K. Zhang, F. Wang, C-S R. Li, **H. Zhao**, L. Lu, X-Y Zhang, L. Lu, H. Zhang, F. Zhang, J. H. Krystal, X. Luo (2011) A novel, functional and replicable risk gene region for alcohol dependence identified by genome-wide

association study. *PLoS One, e26726*.

**2012 (28)**

[185] J. Ferguson, J. Cho, **H. Zhao** (2012) A new approach for the joint analysis of multiple Chip-Seq libraries with application to Histone Modification. *Statistical Applications in Genetics and Molecular Biology*, 11, 3 (1-19).

[186] B. Li, H. Chun, **H. Zhao** (2012) Sparse estimation of conditional graphical models with application to gene networks. *Journal of American Statistical Association*, 107: 152-167.

[187] L. Zuo, J. Gelernter, C. K. Zhang, **H.** **Zhao**, L. Lu, H. R. Kranzler, R. T. Malison, C. S. Li, F. Wang, X. Y. Zhang, H. W. Deng, J. H. Krystal, F. Zhang, X. Luo (2012) Genome-wide association study of alcohol dependence implicates KIAA0040 on chromosome 1q. *Neuropsychopharmacology*, 37: 557-566.

[188] I. Kim, H. Pang, **H. Zhao** (2012) Bayesian semiparametric regression models for evaluating pathway effects on clinical continuous and binary outcomes. *Statistics in Medicine*, 31: 1633-1651.

[189] N. Santoro,C. K. Zhang, **H. Zhao**, A. J. Pakstis, G. Kim, R. Kursawe, D. J. Dykas, A. E. Bale, C. Giannini, B. Pierpont, M. M. Shaw, L. Groop, S. Caprio (2012) A variant in the Glucokinase Regulatory Protein (*GCKR*) gene is associated with fatty liver in obese children and adolescents. *Hepatology,* 55: 781-789.

[190] Z. Wu, **H. Zhao** (2012) On model selection strategies to identify genes underlying binary traits using genome-wide association data. *Statistica Sinica,* 22: 1041-1074.

[191] D. Li, **H. Zhao**, J. Gelernter (2012) Strong protective effect of the aldehyde dehydrogenase gene (ALDH2) 504lys (\*2) allele against alcoholism and alcohol-induced medical diseases in Asians. *Hum Genetics*, 131: 725-737.

[192] G. Li, J. Gelernter, H. R. Kranzler, **H. Zhao** (2012) M3: An improved SNP calling algorithm for Illumina BeadArray Data. *Bioinformatics*, 28: 358-365.

[193] T. Tong, L. Chen, **H. Zhao** (2012) Improved mean estimation and its application to diagonal discriminant analysis. *Bioinformatics*, 28: 531-537.

[194] K. Morino, K. F. Petersen, S. Sono, C. S. Choi, V. T. Samuel, A. Lin, A. Gallo, **H. Zhao**, A. Kashiwagi, I. J. Goldberg, H. Wang, R. H. Eckel, H. Maegawa, G. I. Shulman (2012) Regulation of mitochondrial biogenesis by lipoprotein lipase in muscle of insulin resistant offspring of parents with type 2 diabetes. *Diabetes*, 61: 877-887.

[195] D. Li, **H. Zhao**, H. R. Kranzler, D. Oslin, R. F. Anton, L. A. Farrer, J. Gelernter (2012) Association of *COL25A1* with Comorbid Antisocial Personality Disorder (ASPD) and Substance Dependence. *Biological Psychiatry*, 71: 733-740.

[196] C. K. Zhang, P. B. Stein, J. Liu, G. M. Pastores, Z. Wang, R. Yang, J. H. Cho, P. K. Gregersen, J. M. F. G. Aerts, **H. Zhao**, P. K. Mistry (2012) Genome-wide association study of N370S homozygous gaucher disease reveals the candidacy of *CLN8* gene as a genetic modifier contributing to extreme phenotypic variation. *American Journal of Hematology*, 87: 377-383.

[197]D. Chen, W. Zheng, A. Lin, K. Uyhazi, **H. Zhao**, H. Lin (2012) Pumilio 1 suppresses multiple activators of p53 to safeguard spermatogenesis. *Current Biology*, 22: 420-425.

[198] C. Heffelfinger, Z. Ouyang, A. Engberg, D. J. Leffell, A. M. Hanlon, P. B. Gordon, W. Zheng, **H. Zhao**, M. P. Snyder, A. E. Bale AE (2012) Correlation of global microRNA expression with basal cell carcinoma subtype. *G3* 2: 279-286.

[199] R. D. Beech, J. Qu, J. J. Leffert, A. Lin, K. A. Hong, J. Hansen, S. Umlauf, S. Mane, **H. Zhao**, R. Sinha (2012)

Altered expression of cytokine signaling pathway genes in peripheral blood cells of alcohol dependent subjects: Preliminary Findings. *Alcohol Clin Exp Res*, 36: 1487-1496.

[200] D. Li, **H. Zhao**, J. Gelernter (2012) Further clarification of the contribution of the ADH1C gene to vulnerability of alcoholism and selected liver diseases. *Human Genetics*, 131: 1361-1374.

[201] X. Chen, J. B. Listman, F. Slack, J. Gelernter, **H. Zhao** (2012) Biases and errors on allele frequency estimation and disease association tests of next generation sequencing of pooled samples. *Genetic Epidemiology*, 36: 549-560.

[202] H. Ma, **H. Zhao** (2012) iFad: an integrative factor analysis model for drug-pathway association inference. *Bioinformatics*, 28: 1911-1918.

[203] T. Yuen, J. Iqbal, L.-L. Zhu, L. Sun, A. Lin, **H. Zhao**, J. Liu, P. K. Mistry (2012) Disease-drug pairs revealed by computational genomic connectivity mapping on GBA1 deficient, Gaucher disease mice. *Biochemical and Biophysical Research Communications*, 422: 573-577.

[204] S. Zhang, **H. Zhao** (2012) Community identification in networks with unbalanced structure. *Physical Review E*, 85: 066114.

[205] W. Zhang, J. Ferguson, S. M. Ng, K. Hui, G. Goh, A. Lin, E. Esplugues, R. A. Flavell, C. Abraham, **H. Zhao**, J. H. Cho (2012) Effector CD4+ T cell expression signatures and immune-mediated disease associated genes. *PLoS One* 7: e38510.

[206] M. J. Stulberg, A. Lin, **H. Zhao**, S. A. Holley (2012) Crosstalk between Fgf and Wnt signaling in the zebrafish tailbud. *Dev Biol*, 369: 298-307.

[207] X. Ge, Y. Ren, O. Bartulos, M. Y. Lee, Z. Yue, K.-Y. Kim, W. Li, P. J. Amos, E. C. Bozkulak, A. Iyer, W. Zheng, **H. Zhao**, K. A. Martin, D. N. Kotton, G. Tellides, I.-H. Park, L. Yue, Y. Qyang (2012) Modeling supravalvular aortic stenosis syndrome using human induced pluripotent stem cells. *Circulation*, 126: 1695-1704.

[208] H. Ma, **H. Zhao** (2012) FacPad: Bayesian sparse factor modeling for the inference of pathways responsive to drug treatment. *Bioinformatics*, 28: 2662-2670.

[209] L. Lowthert, J. J. Leffert, A. Lin, S. Umlauf, K. Maloney, A. Muralidharan, B. Lorberg, S. Mane, **H. Zhao**, R. Sinha, Z. Bhagwagar, R. D. Beech (2012) Increased ratio of anti-apoptotic to pro-apoptotic Bcl2 gene-family members in lithium-responders one month after treatment initiation. *Biology of Mood and Anxiety Disorders*, 2:15.

[210] C. Ramírez, Y. Díaz-Toro, J. Tellez, T. M. Castilho, R. Rojas, N. A. Ettinger, I. Tikhonova, N. D. Alexander, L. Valderrama, J. Hager, M. E. Wilson, A. Lin, **H. Zhao**, N. G. Saravia, D. McMahon-Pratt (2012) Human Macrophage Response to L. (Viannia) panamensis: Microarray Evidence for an Early Inflammatory Response. *PLoS Negl Trop Dis.* 6: e1866.

[211] T. Yuen, J. Iqbal, L.-L. Zhu, L. Sun, A. Lin,[**H. Zhao**](http://loop.frontiersin.org/people/226657/overview), J. Liu, P. K. Mistry, M. Zaidi (2012) [Disease-drug pairs revealed by computational genomic connectivity mapping on GBA1 deficient, Gaucher disease mice.](http://www.frontiersin.org/Profile/PublicationDetails.aspx?PubId=22588172)

*Biochemical and Biophysical Research Communications*, 422: 573-577.

[212] L. Jostins, S. Ripke, R. K. Weersma, R. H. Duerr, D. P. McGovern, K. Y. Hui, J. C. Lee, L. P. Schumm, Y. Sharma, C. A. Anderson, J. Essers, M. Mitrovic, K. Ning, I. Cleynen, E. Theatre, S. L. Spain, S. Raychaudhuri, P. Goyette, Z. Wei, C. Abraham, J. P. Achkar, T. Ahmad, L. Amininejad, A. N. Ananthakrishnan, V. Andersen, J. M. Andrews, L. Baidoo, T. Balschun, P. A. Bampton, A. Bitton, G. Boucher, S. Brand, C. Büning, A. Cohain, S. Cichon, M. D'Amato, D. De Jong, K. L. Devaney, M. Dubinsky, C. Edwards, D. Ellinghaus, L. R. Ferguson, D. Franchimont, K. Fransen, R. Gearry, M. Georges, C. Gieger, J. Glas, T. Haritunians, A. Hart, C. Hawkey, M. Hedl, X. Hu, T. H. Karlsen, L. Kupcinskas, S. Kugathasan, A. Latiano, D. Laukens, I. C. Lawrance, C. W. Lees, E. Louis, G. Mahy, J. Mansfield, A. R. Morgan, C. Mowat, W. Newman, O. Palmieri, C. Y. Ponsioen, U. Potocnik, N. J. Prescott, M. Regueiro, J. I. Rotter, R. K. Russell, J. D. Sanderson, M. Sans, J. Satsangi, S. Schreiber, L. A. Simms, J. Sventoraityte, S. R. Targan, K. D. Taylor, M. Tremelling, H. W. Verspaget, M. De Vos, C. Wijmenga, D. C. Wilson, J. Winkelmann, R. J. Xavier, S. Zeissig, B. Zhang, C. K. Zhang, **H. Zhao**; International IBD Genetics Consortium (IIBDGC), M. S. Silverberg, V. Annese, H. Hakonarson, S. R. Brant, G. Radford-Smith, C. G. Mathew, J. D. Rioux, E. E. Schadt, M. J. Daly, A. Franke, M. Parkes, S. Vermeire, J. C. Barrett, J. H. Cho (2012) Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature*, 491: 119-124.

**2013 (35)**

[213] H. Zhang, A. I. Herman, H. R. Kranzler, R. F. Anton, **H. Zhao**, W. Zheng, J. Gelernter (2013) Array-based profiling of DNA methylation changes associated with alcohol dependence. *Alcoholism: Clinical and Experimental Research*, 37 Suppl 1:E108-15.

[214] X. Qi, **H. Zhao** (2013) Sparse principal component analysis by choice of norm. *Journal of Multivariate Analysis*, [114](http://www.sciencedirect.com/science/journal/0047259X/114/supp/C): 127-160.

[215] J. Ferguson, C. Yang, J. Cho, **H. Zhao** (2013) Empirical Bayes correction for the winner's curse in genetic association studies. *Genetic Epidemiology*, 37: 60-68.

[216] I. Kim, H. Pang, **H. Zhao** (2013) Statistical properties on semiparametric regression for evaluating pathway effects. *Journal of Statistical Planning and Inference* 143: 745-763.

[217] J. Ferguson, W. Wheeler, Y-P. Fu, L. Prokunina-Olsson, **H. Zhao**, J. Sampson (2013) Statistical tests for detecting associations with groups of genetic variants: generalization, evaluation, and implementation. *European Journal of Human Genetics*, 21: 680-686.

[218] X. Zhou, C. Yang, X. Wan, **H. Zhao**, W. Yu (2013) Multi-sample aCGH data analysis via total variation and spectral regularization. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 10: 230-235.

[219] F. Qian, C. Bolen, C. Jing, X. Wang, W. Zheng, **H. Zhao**, E. Fikrig, R. Bruce, S. Kleinstein, R. Montgomery (2013) Impaired TLR3-mediated immune responses from macrophages of patients chronically infected with Hepatitis C virus. *Clinical and Vaccine Immunology*, 20: 146-155.

[220] X. Wan, C. Yang, Q. Yang, **H. Zhao**, W. Yu (2013) HapBoost: A fast approach to boosting haplotype association analyses in genome-wide association studies. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 10: 207-212.

[221] L. Wang, Y. Jiao, Y. Huang, X. Liu, G. Gibson, B. Bennett, K.M. Hamre, D. Li, **H. Zhao**, J. Gelernter, H. Kranzler, L. Farrer, L. Lu, R. Williams, Y. Wang, W. Gu (2013) A critical evaluation of the transcription factor, Atf2, as a candidate modulator of alcohol preference in mouse and human populations. *Genetics and Molecular Research*, 12: 5992-6005.

[222] X. Wan, C. Yang, Q. Yang, **H. Zhao**, W. Yu (2013) The complete compositional epistasis detection in genome-wide association studies. *BMC Genetics*, 14:7.

[223] C. Yang, L. Wang, S. Zhang, **H. Zhao** (2013) Accounting for non-genetic factors by low-rank representation and sparse regression for eQTL mapping. *Bioinformatics*, 29: 1026-1034.

[224] L. Wang, W. Zheng, **H. Zhao**, M. Deng (2013) Statistical analysis reveals co-expression patterns of many pairs of genes in yeast are jointly regulated by interacting loci. *PLoS Genetics*, 9: e1003414.

[225] J. S. Lee, **H. Zhao** (2013) [On estimation of allele frequencies via next-generation DNA resequencing with barcoding.](http://www.ncbi.nlm.nih.gov/pubmed/23730349) *Statistics in BioSciences*, 5: 26-53.

[226] T. Tong, Z. Feng, J. S. Hilton, **H. Zhao** (2013) Estimating the proportion of true null hypotheses using the pattern of observed *p*-values. *Journal of Applied Statistics*, 40: 1949-1964.

[227] W. Zheng, **H. Zhao** (2013) Studying the evolution of transcription factor binding events using multi-species ChIP-Seq data. *Stat Appl Genet Mol Biol*, 2013 12: 1-15.

[228] L. M. Chung, J. P. Ferguson, W. Zheng, F. Qian, V. Bruno, R. R. Montgomery, **H. Zhao** (2013) Differential expression analysis for paired RNA-Seq data. *BMC Bioinformatics*,14: 110.

[229] W. Zhang, K. Y. Hui, A. Gusev, N. Warner, S. M. Ng, J. Ferguson, M. Choi, A. Burberry, C. Abraham, L. Mayer, R. J. Desnick, C. J. Cardinale, H. Hakonarson, M. Waterman, Y. Chowers, A. Karban, S. R. Brant, M. S. Silverberg, P. K. Gregersen, S. Katz, R. P. Lifton, **H. Zhao**, G. Nuñez, I. Pe'er, I. Peter, J. H. Cho (2013) Extended haplotype association study in Crohn's disease identifies a novel, Ashkenazi Jewish-specific missense mutation in the NF-κB pathway gene, HEATR3. *Genes Immun* 14: 310-316.

[230] S. Zaidi, M. Choi, H. Wakimoto, L. Ma, J. Jiang, J. D. Overton, A. Romano-Adesman, R. D. Bjornson, R. E. Breitbart, K. K. Brown, N. J. Carriero, Y. H. Cheung, J. Deanfield, S. DePalma, K. A. Fakhro, J. Glessner, H. Hakonarson, M. J. Italia, J. R. Kaltman, J. Kaski, R. Kim, J. K. Kline, T. Lee, J. Leipzig, A. Lopez, S. M. Mane, L. E. Mitchell, J. W. Newburger, M. Parfenov, I. Pe'er, G. Porter, A. E. Roberts, R. Sachidanandam, S. J. Sanders, H. S. Seiden, M. W. State, S. Subramanian, I. R. Tikhonova, W. Wang, D. Warburton, P. S. White, I. A. Williams, **H. Zhao**, J. G. Seidman, M. Brueckner, W. K. Chung, B. D. Gelb, E. Goldmuntz, C. E. Seidman, R. P. Lifton (2013) De novo mutations in histone-modifying genes in congenital heart disease. *Nature*, 498: 220-223.

[231] C. K. Zhang, J. Hewett, J. Hemming, T. Grant, **H. Zhao**, C. Abraham, I. Oikonomou, M. Kanakia, J. H. Cho, D. D. Proctor (2013) The influence of depression on quality of life in patients with inflammatory bowel disease. *Inflamm Bowel Dis*., 19: 1732-1739.

[232] J. P. Saxe, M. Chen, **H. Zhao**, H. Lin (2013) Tdrkh is essential for spermatogenesis and participates in primary piRNA biogenesis in the germline. *EMBO J.*, 32: 1869-1885.

[233] P. Xie, H. R. Kranzler, C. Yang, **H. Zhao**, L. A. Farrer, J. Gelernter (2013) Genome-wide association study identifies new susceptibility loci for posttraumatic stress disorder. *Biol Psychiatry*, 74: 656-663.

[234] G. A. Ryslik, Y. Cheng, K. H. Cheung, Y. Modis, **H. Zhao** (2013) Utilizing protein structure to identify non-random somatic mutations. *BMC Bioinformatics*, 14: 190.

[235] X. Chen, F. J. Slack, **H. Zhao** (2013) Joint analysis of expression profiles from multiple cancers improves the identification of microRNA-gene interactions. *Bioinformatics*, 29: 2137-2145.

[236] H. Zhang, F. Wang, H. R. Kranzler, **H. Zhao**, J. Gelernter J (2013) Profiling of childhood adversity-associated DNA methylation changes in alcoholic patients and healthy controls. *PLoS One*, 8: e65648.

[237] F. Qian, L. Chung, W. Zheng, V. Bruno, R. P. Alexander, Z. Wang, X. Wang, S. Kurscheid, **H. Zhao**, E. Fikrig, M. Gerstein, M. Snyder, R. R. Montgomery (2013) Identification of genes critical for resistance to infection by West Nile virus using RNA-Seq analysis. *Viruses*, 5: 1664-1681.

[238] Y. Gao, S. A. Nish, R. Jiang, L. Hou, P. Licona-Limón, J. S. Weinstein, **H. Zhao**, R. Medzhitov (2013) Control of T helper 2 responses by transcription factor IRF4-dependent dendritic cells. *Immunity*, 39: 722-732.

[239] J. Gelernter, H. R. Kranzler, R. Sherva, L. Almasy, R. Koesterer, A. H. Smith, R. Anton, U. W. Preuss, M. Ridinger, D. Rujescu, N. Wodarz, P. Zill, **H. Zhao**, L. A. Farrer (2013) Genome-wide association study of alcohol dependence: significant findings in African- and European-Americans including novel risk loci. *Mol Psychiatry*, 19: 41-49.

[240] H. Zhu, F. Hu, **H. Zhao** (2013) Adaptive clinical trial designs to detect interaction between treatment and a dichotomous biomarker. *Canadian Journal of Statistics*, 41: 525-539.

[241] W. Chen, Y. Cheng, C. Zhang, S. Zhang, **H. Zhao** (2013) MSClust: A Multi-Seeds based Clustering algorithm for microbiome profiling using 16S rRNA sequence. *Journal of Microbiological Methods*, 94: 347-355.

[242] W. Chen, C. K. Zhang, Y. Cheng, S. Zhang, **H. Zhao** (2013) A comparison of methods for clustering 16S rRNA sequences into OTUs. *PLoS One*, 8: e70837.

[243] S. Zhang, **H. Zhao** (2013) Normalized modularity optimization method for community identification with degree adjustment. *Physical Review E*, 88, 052802

[244] W. D. Li, H. Jiao, K. Wang, C. K. Zhang, J. T. Glessner, S. F. Grant, **H. Zhao**, H. Hakonarson, A. R. Price (2013)

A genome wide association study of plasma uric acid levels in obese cases and never-overweight controls. *Obesity*, 21: E490-494.

[245] S. Han, B-Z Yang, H. R. Kranzler, X. Liu, **H. Zhao**, L. A. Farrer, E. Boerwinkle,J. Potash, J. Gelernter (2013)

Integrating GWAS and human protein interaction network identifies a gene sub-network underlying alcohol dependence. *Am J Hum Genet*, 93: 1027-1034.

[246] M. Chen, M. Gunel, **H. Zhao** (2013) SomatiCA: Identifying, Characterizing and Quantifying Somatic Copy Number Aberrations from Cancer Genome Sequencing Data. *PLoS One*, 8: e78143.

[247] S. Oh, S. Song, G. Grabowski, **H. Zhao**, J. P. Noonan (2013) Time series expression analyses using RNA-seq: A statistical approach. *BioMed Research International*, 2013: 203681.

**2014 (34)**

[248] R. D. Beech, J. J. Leffert, A. Lin, L. G. Sylvia, S. Umlauf, S. Mane, **H. Zhao**, C. Bowden, J. R. Calabrese, E. S. Friedman, T. A. Ketter, D. V. Iosifescu, N. A. Reilly-Harrington, M. Ostacher, M. E. Thase, A. Nierenberg (2014) Gene-expression differences in peripheral blood between lithium responders and non-responders in the Lithium Treatment-Moderate dose Use Study (LiTMUS). *Pharmacogenomics J*., 14: 182-191.

[249] C. Yang, C. Li, H. R. Kranzler, L. A. Farrer, **H. Zhao**, J. Gelernter (2014) Exploring the genetic architecture of alcohol dependence in African-Americans via analysis of a genomewide set of common variants. *Human Genetics*, 133: 617-624.

[250] C. Li, C. Yang, J. Gelernter, **H. Zhao** (2014) Improving genetic risk prediction by leveraging pleiotropy. *Human Genetics*, 133: 639-650.

[251]L. Wang, T. Rundek, A. Beecham, B. Hudson, S. H. Blanton, **H. Zhao**, R. L. Sacco, C. Dong (2014) Genome-wide interaction study identifies RCBTB1 as a modifier for smoking effect on carotid intima-media thickness. *Arterioscler Thromb Vasc Biol*, 34: 219-225.

[252] J. E. Brown, B. R. Evans, W. Zheng, V. Obas, L. Barrera-Martinez, A. Egizi, **H. Zhao**, A. Caccone, J. R. Powell (2014) Human impacts have shaped historical and recent evolution in *aedes aegypti*, the Dengue and Yellow Fever mosquito. *Evolution*, 68: 514-525.

[253] D. Li, A. Sulovari, C. Cheng, **H. Zhao**, H. R. Kranzler, J. Gelernter (2014) Association of gamma-aminobutyric acid A receptor α2 gene (GABRA2) with alcohol use disorder. *Neuropsychopharmacology*, 39: 907-918.

[254] L. Hou, M. Chen, C. K. Zhang, J. Cho, **H. Zhao** (2014) Guilt by Rewiring: Gene prioritization through network rewiring in genome wide association studies. *Human Molecular Genetics*, 23: 2780-2790.

[255] B. Li, H. Chun, **H. Zhao** (2014) On an additive semi-graphoid model for statistical networks with application to pathway analysis. *Journal of American Statistical Association*, 109: 1188-1204.

[256] F. Xu, L. Li, P. V. Schulz, G. P. Gallagher, B. Xiang, **H. Zhao**, P. Li (2014) Cytogenomic mapping and bioinformatic mining reveal interacting brain expressed genes for intellectual disability. *Molecular Cytogenetics*, 7:4.

[257] M. Kim, X. Chen, L. Chin, T. Paranjape, W. Speed, K. K. Kidd, **H. Zhao**, J. B. Weidhaas, F. J. Slack FJ (2014) Extensive sequence variation in the 3′ untranslated region of the KRAS gene in lung and ovarian cancer cases. *Cell Cycle*, 13: 1030-1040.

[258] D. Schenten, S. A. Nish, S. Yu, X. Yan, H. K. Lee, I. Brodsky, L. Pasman, B. Yordy, F. T. Wunderlich, J. C. Brüning, **H. Zhao**, R. Medzhitov (2014) Signaling through the Adaptor Molecule MyD88 in CD4(+) T Cells Is Required to Overcome Suppression by Regulatory T Cells. *Immunity* 40: 78-90.

[259] Z. Wu, Y. Sun, S. He, J. Cho, **H. Zhao**, J. Jin (2014) Detection boundary and higher criticism approach for rare and weak genetic effects. *Annals of Applied Statistics*, 8: 824-851.

[260] R. Zhu, **H. Zhao**, S. Ma (2014) Identifying gene-environment and gene-gene interactions using a progressive penalization approach. *Genetic Epidemiology*, 38:353-368.

[261] Brownstein et al. (2014) An international effort towards developing standards for best practices in analysis,

interpretation and reporting of clinical genome sequencing results in the CLARITY Challenge. *Genome Biology*, 15: R53.

[262] G. A. Ryslik, Y. Cheng, K.-H. Cheung, Y. Modis, **H. Zhao** (2014) A graph theoretic approach to utilizing protein structure to identify non-random somatic mutations. *BMC Bioinformatics*, 15: 86.

[263] A. R. Gupta, M. Pirruccello, F. Cheng, H. J. Kang, T. V. Fernandez, J. M. Baskin, M. Choi, L. Liu, A. G. Ercan-Sencicek, J. D. Murdoch, L. Klei, B. M. Neale, D. Franjic, M. J. Daly, R. P. Lifton, P. De Camilli, **H. Zhao**, N. Sestan, M. W. State (2014) Rare deleterious mutations of the gene EFR3A in autism spectrum disorders. *Molecular Autism*, 5:31.

[264] P. Xie, H. R. Kranzler, J. H. Krystal, L. A. Farrer, **H. Zhao**, J. Gelernter (2014) Deep resequencing of 17 glutamate system genes identifies rare variants in DISC1 and GRIN2B affecting risk of opioid dependence. *Addict Biol*, 19: 955-964.

[265] D. N. Assis, L. Leng, X. Du, C. K. Zhang, G. Grieb, M. Merk, A. B. Garcia, C. McCrann, J. Chapiro, A. Meinhardt, Y. Mizue, D. J. Nikolic-Paterson, J. Bernhagen, M. M. Kaplan, **H. Zhao**, J. L. Boyer, R. Bucala (2014) The role of macrophage migration inhibitory factor (MIF) in autoimmune liver disease. *Hepatology*, 59: 580-591.

[266] J. Gelernter, R. Sherva, R. Koesterer, L. Almasy, **H. Zhao**, H. R. Kranzler, L. Farrer (2014) Genome-wide association study of cocaine dependence and related traits: FAM53B identified as a risk gene. *Mol Psychiatry*, 19: 41-49.

[267] C. Giannini, C. D. Man, L. Groop, C. Cobelli, **H. Zhao**, M. M. Shaw, E. Duran, B. Pierpont, A. E. Bale, S. Caprio, N. Santoro (2014) The co-occurrence of risk alleles in or near genes modulating insulin secretion predisposes obese youth to prediabetes. *Diabetes Care*, 37: 475-482.

[268] J. Gelernter, H. R. Kranzler, R. Sherva, R. Koesterer, L. Almasy, **H. Zhao**, L. A. Farrer LA (2014) Genome-wide association study of opioid dependence: multiple associations mapped to calcium and potassium pathways. *Biol Psychiatry*, 76: 66-74.

[269] L. Chung, C. Colangelo, **H. Zhao** (2014) Data pre-processing for label-free multiple reaction monitoring (MRM) experiments. *Biology*, 3: 383-402.

[270] S. A. Nish, D. Schenten, T. Wunderlich, S. D. Pope, Y. Gao, N. Hoshi, S. Yu, X. Yan, H. K. Lee, L. Pasman, I. Brodsky, B. Yordy, **H.** **Zhao**, J. Brüning, R. Medzhitov (2014) T cell-intrinsic role of IL-6 signaling in primary and memory responses. *Elife*. 2014 May 19:e01949.

[271] H. Zhang, F. Wang, C. Yang, H. Xu, Z. Wang, **H. Zhao**, J. Gelernter(2014) Identification of methylation quantitative trait loci (mQTLs) influencing DNA methylation in the promoter regions of alcohol dependence risk genes. *Human Genetics*, 133: 1093-1104.

[272] G. A. Ryslik, Y. Cheng, K.-H. Cheung, R. D. Bjornson, D. Zelterman, Y. Modis, **H. Zhao** (2014) A spatial simulation approach to account for protein structure when identifying non-random somatic mutations. *BMC Bioinformatics*, 15: 231.

[273] P. D. Harvey, L. J. Siever, G. D. Huang, S. Muralidhar, **H. Zhao**, P. Miller, M. Aslan, S. Mane, M. McNamara, T. Gleason, M. Brophy, R. Przygodszki, T. J. O’Leary, M. Gaziano, J. Concato (2014) The genetics of functional disability in schizophrenia and bipolar illness: Methods and initial results for VA cooperative study #572. *Am J Med Genet B Neuropsychiatr Genet* 165: 381-389.

[274] D. Della-Morte, L. Wang, A. Beecham, S. H. Blanton, **H. Zhao**, R. L. Sacco, T. Rundek, C. Dong (2014) Novel genetic variants modify the effect of smoking on carotid plaque burden in Hispanics. *J Neurol Sci*, 344: 27-31.

[275] H. Zhang, F. Wang, J. Liu, **H. Zhao**, J. Gelernter (2014) Differentially co-expressed genes in postmortem prefrontal cortex tissues of individuals with alcohol use disorders. *Human Genetics*, 133: 1383-1394.

[276] R. D. Beech, J. J. Leffert, A. Lin, K. A. Hong, J. Hansen, S. Umlauf, S. Mane, **H. Zhao**, R. Sinha(2014) Stress-related alcohol consumption in heavy drinkers correlates with expression of miR-10a, miR-21 and components of the TAR-RNA binding protein (TRBP)-associated complex. *Alcoholism: Clinical and Experimental Research*, 38: 2743-2753.

[277] L. Hou, T. Ma, **H. Zhao** (2014) Incorporating functional annotation information in prioritizing disease associated SNPs from genome wide association studies. *Science China Life Sciences*, 57: 1072-1079.

[278] H. Pang, H. Zhao (2014) Stratified pathway analysis for heterogeneity among the samples. *Cancer Informatics*, 13(Suppl 4): 73-78.

[279] J. E. Gallagher, W. Zheng, X. Rong, N. Miranda, Z. Lin, B. Dunn, **H. Zhao**, M. P. Snyder (2014) Divergence in a master variator generates distinct phenotypes and transcriptional responses. *Genes Dev* 28: 409-421.

[280] D. Chung, C. Yang, C. Li, J. Gelernter, **H. Zhao** (2014) GPA: A statistical approach to prioritizing GWAS results by integrating pleiotropy and annotation. *PLOS Genetics*, 10: e1004787.

[281] C. Heffelfinger, C. A. Fragoso, M. A. Moreno, J. D. Overton, J. P. Mottinger, **H. Zhao**, J. Tohme, S. L. Dellaporta (2014) Flexible and scalable genotyping-by-sequencing strategies for population studies. *BMC Genomics* 15: 979.

**2015 (27)**

[282] D. Li, **H. Zhao**, H. R. Kranzler, M. D. Li, K. P. Jensen, T. Zayats, L. A. Farrer, J. Gelernter (2015) Genome-wide association study of copy number variations (CNVs) with opioid dependence. *Neuropsychopharmacology*, 40: 1016-1026.

[283] J. Gelernter, H. R. Kranzler,R. Sherva, L. Almasy, A. I. Herman, R. Koesterer, **H. Zhao**, L. A. Farrer (2015) Genomewide association study of nicotine dependence in American populations: identification of novel risk loci in both African- and European-Americans. *Biological Psychiatry*, 77: 493-503.

[284] X. Chen, D. Chung, G. Stefani, F. J. Slack, **H. Zhao** (2015) Statistical issues in binding site identification through CLIP-seq. *Statistics and Its Interface*, 8: 419–436.

[285] X. Chen, T. Paranjape, C. Stahlhut, T. McVeigh, F. Keane, S. Nallur, N. Miller, M. Kerin, Y. Deng, X. Yao, **H. Zhao**, J. Weidhaas, F. Slack (2015) Targeted resequencing of the microRNAome and 3'UTRome reveals functional germline DNA variants with altered prevalence in epithelial ovarian cancer. *Oncogene*, 34: 2125-2137.

[286] H. Pang, I. Kim. **H. Zhao** (2015) Random effects model for multiple pathway analysis with applications to type II diabetes microarray data. *Statistics in Biosciences*, 7: 167–186.

[287] X. Qi, R. Luo, R. Carroll, **H. Zhao** (2015) Sparse regression by projection and sparse discriminant analysis. *Journal of Computational and Graphical Statistics*, 24: 416-438.

[288] H. Chun, X. Zhang, **H. Zhao** (2015) Gene regulation network inference with joint sparse Gaussian graphical models. *Journal of Computational and Graphical Statistics*, 24: 954–974*.*

[289] J. Sun, **H. Zhao** (2015) The application of sparse estimation of covariance matrix to quadratic discriminant analysis. *BMC Bioinformatics*, 16:48.

[290] E. F. Foxman, J. A. Storer, K. Vanaja, B. R. Wasik, L. Hou, **H. Zhao**, P. E. Turner, A. Levchenko, A. Iwasaki (2015) Temperature-dependent innate defense against the common cold virus limits viral replication at warm temperature in mouse airway cells. *Proc Natl Acad Sci USA*, 112: 827-32.

[291] C. Zheng,C. D. Man, C. Cobelli, L. Groop, **H. Zhao**,A. E. Bale, M. Shaw, E. Duran, B. Pierpont, S. Caprio, N. Santoro (2015) A common variant in the MTNR1B gene is associated with increased risk of impaired fasting glucose (IFG) in obese youth. *Obesity*, 23: 1022-1029.

[292] S. Zhang, **H. Zhao**, M. K. Ng (2015) Functional module analysis for gene coexpression networks with network integration. *IEEE/ACM Transactions on Computational Biology and Bioinformatics,* 12: 1146 – 1160.

[293] Z. Lin, S. Sanders, M. Li, N. Sestan, M. State, **H. Zhao** (2015) A Markov random field-based approach to characterizing human brain development using spatial-temporal transcriptome data. *Annals of Applied Statistics*, 9: 429–451.

[294] A. M. B. Tadeu, S. Lin, L. Hou, L. Chung, M. Zhong, **H. Zhao**, V. Horsley (2015)Transcriptional profiling of ectoderm specification to keratinocyte fate in human embryonic stem cells. *PLOS One*, 10: e0122493.

[295] W. Lam, Z. Jiang, F. Guan, X. Huang, R. Hu, J. Wang, S. Bussom, S.-H. Liu, **H. Zhao**, Y. Yen, Y.-C. Cheng (2015) PHY906(KD018), an adjuvant based on a 1800-year-old Chinese medicine, enhanced the anti-tumor activity of Sorafenib by changing the tumor microenvironment. *Scientific Reports*, 5: 9384.

[296] B. R. Evans, A. Gloria-Soria, L. Hou, C. McBride, M. Bonizzoni, **H. Zhao**, J. R. Powell (2015) A multipurpose high throughput SNP chip for the dengue and yellow fever mosquito, Aedes aegypti. *G3* 5: 711-718.

[297] X. Yan, J.-H. Chu, J. Gomez, M. Koenigs, C. Holm, X. He, M. F. Perez, **H. Zhao**, S. Mane, F. D. Martinez, C. Ober, D. L. Nicolae, K. C. Barnes, S. J. London, F. Gilliland, S. T. Weiss, B. A. Raby, L. Cohn, G. L. Chupp (2015) Non-invasive analysis of the sputum transcriptome discriminates clinical phenotypes of asthma. *American Journal of Respiratory and Critical Care Medicine*, 191(10):1116-1125.

[298] G. Stefani, X. Chen, **H.** **Zhao**, F. J. Slack (2015) A novel mechanism of LIN-28 regulation of let-7 microRNA expression revealed by in vivo HITS-CLIP in C. elegans. *RNA*, 21: 985-996.

[299] K. Xu, H. R. Kranzler, R. Sherva, C. E. Sartor, L. Almasy, R. Koesterer, **H. Zhao**, L. A. Farrer, J. Gelernter(2015) Genome-wide association study for maximum number of alcohol drinks in European Americans and African Americans. *Alcoholism: Clinical and Experimental Research*, 39: 1137-1147.

[300] P. Goyette, G. Boucher, D. Mallon, E. Ellinghaus, L. Jostins, H. Huang, S. Ripke, E. S. Gusareva, V. Annese, S. L. Hauser, J. R. Oksenberg, I. Thomsen, S. Leslie; **International Inflammatory Bowel Disease Genetics Consortium**; Australia and New Zealand IBDGC; Belgium IBD Genetics Consortium; Italian Group for IBD Genetics Consortium; NIDDK Inflammatory Bowel Disease Genetics Consortium; United Kingdom IBDGC; Wellcome Trust Case Control Consortium; Quebec IBD Genetics Consortium, M. J. Daly, K. Van Steen, R. H. Duerr, J. C. Barrett, D. P. McGovern, L. P. Schumm, J. A. Traherne, M. N. Carrington, V. Kosmoliaptsis, T. H. Karlsen, A. Franke, J. D. Rioux (2015) High-density mapping of the MHC identifies a shared role for HLA-DRB1\*01:03 in inflammatory bowel diseases and heterozygous advantage in ulcerative colitis. *Nature Genetics*, 47: 172-179.

[301] Q. Lu, Y. Hu, J. Sun, Y. Cheng, K.-H. Cheung, **H. Zhao** (2015) A statistical framework to predict functional non-coding regions in the human genome through integrated analysis of annotation data. *Scientific Reports*, 5: 10576.

[302] Y. Xie, Y. Jin, B. L. Merenick, M. Ding, K. M. Fetalvero, R. J. Wagner, A. Mai, S. Gleim, D. Tucker, M. J. Birnbaum, B. A. Ballif, A. K. Luciano, W. C. Sessa, E. M. Rzucidlo, R. J. Powell, L. Hou, H. Zhao, J. Yu, J. Hwa, K. A. Martin (2015) Phosphorylation of GATA-6 is required for vascular smooth muscle cell differentiation after mTORC1 inhibition. *Science Signaling*, 8(376): ra44.

[303] L. Zuo, C. K Zhang, F. G. Sayward, K.-H. Cheung, K. Wang, J. H Krystal, H. Zhao, X. Luo (2015) Gene-based and pathway-based genome-wide association study of alcohol dependence. *Shanghai Archives of Psychiatry*, 27: 111-118.

[304] Q. Wang, C. Yang, J. Gelernter, **H. Zhao** (2015) Pervasive pleiotropy between psychiatric disorders and immune disorders revealed by integrative analysis of multiple GWAS. *Human Genetics*, 134: 1195-1209.

[305] R. Polimanti, C. Yang, **H. Zhao**, J. Gelernter (2015) Dissecting ancestry genomic background in substance dependence genome-wide association studies. *Pharmacogenomics*, 16: 1487-1498.

[306] H. Fang, C. Huang, **H. Zhao**, M. Deng (2015) CCLasso: Correlation inference for compositional data through Lasso. *Bioinformatics*, 31: 3172-3180.

[307] G. Li, Y. Cui, **H. Zhao** (2015) An Empirical Bayes risk prediction model using multiple traits for sequencing data. *Statistical Applications in Genetics and Molecular Biology*, 14: 551-573.

[308] G. Li, **H. Zhao** (2015) M3-S: a genotype calling method incorporating information from samples with known genotypes. *BMC Bioinformatics*, 16:403.

**2016 (29)**

[309] M. Chen, Z. Ren, **H. Zhao**, H. Zhou (2016) Asymptotically normal and efficient estimation of covariate-adjusted Gaussian graphical model. *Journal of American Statistical Association*, 111: 394-406.

[310] C. Li, C. Yang, G. Hather, R. Liu, **H. Zhao** (2016) Efficient drug-pathway association analysis via integrative penalized matrix decomposition. *IEEE/ACM Transactions on Computational Biology and Bioinformatics,* 13: 531-540.

[311] M. Chen, C. Gao, **H. Zhao** (2016) Posterior contraction rates of the phylogenetic Indian buffet processes. *Bayesian Analysis*, 11: 477-497.

[312] T. Wang, M. Chen, **H. Zhao** (2016) Estimating DNA methylation levels by joint modeling of multiple methylation profiles from microarray data. *Biometrics*, 72: 354-363.

[313] Q. Lu, X. Yao, Y. Hu, **H. Zhao** (2016) GenoWAP: GWAS signal prioritization through integrated analysis of genomic functional annotation. *Bioinformatics*, 32: 542-548.

[314] J. Jiang, C. Li, D. Paul, C. Yang, **H. Zhao** (2016) On high-dimensional misspecified mixed model analysis in genome-wide association study. *Annals of Statistics*, 44: 2127–2160.

[315] F. Wang, H. Xu, **H. Zhao**, J. Gelernter, H. Zhang (2016) DNA co-methylation modules in postmortem prefrontal cortex tissues of European Australians with alcohol use disorders. *Scientific Reports*, 6: 19430.

[316] C. A. Fragoso, C. Heffelfinger, **H. Zhao**, S. L. Dellaporta (2016) Imputing genotypes in biallelic populations from low-coverage sequence data. *Genetics*, 202: 487-495.

[317] P. Harvey, M. Aslan, M. Du, **H. Zhao**, L. Siever, A. Pulver, J. M. Gaziano, J. Concato (2016) Factor structure of cognition and functional capacity in two studies of schizophrenia and bipolar disorder: Implications for genomic studies. *Neuropsychology*, 30: 28-39.

[318] M. Goffredo, S. Caprio, A. Feldstein, E. D’Adamo, M. M. Shaw, B. Pierpont, M. Savoye, **H. Zhao**, A. E. Bale, N. Santoro (2016) Role of the *TM6SF2* rs58542926 in the pathogenesis of pediatric fatty liver disease (NAFLD): a multiethnic study. *Hepatology*, 63: 117-125.

[319] X. Huang, D. Stern, **H. Zhao** (2016) Transcriptional profiles from paired normal samples offer complementary information on cancer patient survival -- Evidence from TCGA Pan-Cancer Data. *Scientific Reports*, 6: 20567.

[320] R. Sherva, Q. Wang, H. Kranzler, **H. Zhao**, R. Koesterer, A. Herman, L. Farrer, J. Gelernter (2016)

Genome wide association study of cannabis dependence severity reveals novel risk variants, genes previously implicated in schizophrenia risk, and shared risk with major depressive disorder. *JAMA Psychiatry*, 73: 472-480.

[321] M. Chen. H. Lin, **H. Zhao** (2016) Change point analysis of histone modifications reveals epidenteic blocks with distinct regulatory activity and biological functions. *Annals of Applied Statistics*, 10: 506-526.

[322] Z. Lin, M. Li, N. Sestan, **H. Zhao** (2016) A Markov Random Field-based approach for joint estimation of differentially expressed genes in mouse transcriptome data. *Statistical Applications in Molecular Biology and Genetics*, 15: 139-150.

[323] G. Ryslik, Y. Cheng, Y. Modis. **H. Zhao** (2016) Leveraging protein quaternary structure to identify oncogenic driver mutations. *BMC Bioinformatics*, 17: 137.

[324] Q. Lu, R. Powles, Q. Wang, J. He, **H. Zhao** (2016) Integrative tissue-specific functional annotations in the human genome provide novel insights on many complex traits and improve signal prioritization in genome wide association studies. *PLOS Genetics*, 12: e1005947.

[325] J. Liu, C. Yang, X. Shi, C. Li, J. Huang, **H. Zhao,** S. Ma (2016)Analyzing association mapping in pedigree-based GWAS using apenalized multi-trait mixed model. *Genetic Epidemiology*, 40: 382-393.

[326] Z. Xie, D. Zhang, D. Chung, Z. Tang, H. Huang, L. Dai, S. Qi, J. Li, G. Colak, Y. Chen, C. Xia, C. Peng, H. Ruan, M. Kirkey, D. Wang, L. M. Jensen, S. D. Pletcher, M. Tan, D. B. Lombard, K. P. White, **H. Zhao**, J. Li, R. G. Roeder, X. Yang, Y. Zhao(2016) Metabolic regulation of gene expression by histone lysine -hydroxybutyrylation. *Molecular Cell*, 62: 194-206.

[327] N. Li, L. Subrahmanyan, E. Smith E, X. Yu, S. Zaidi, M. Choi, S. Mane, C. Nelson-Williams, M. Bahjati, M. Kazemi, M. Hashemi, M. Fathzadeh, A. Narayanan, L. Tian, F. Montazeri, M. Mani, M. L. Begleiter, B. G. Coon, H. T. Lynch, E. N. Olson, **H. Zhao**, J. Ruland, R. P. Lifton, A. Mani (2016) Mutations in the Histone Modifier PRDM6 Are Associated with Isolated Nonsyndromic Patent Ductus Arteriosus. *Am J Hum Genet*, 98: 1082-1091.

[328] L. S. Chuang, N. Villaverde, K. Y. Hui, A. Mortha, A. Rahman, A. P. Levine, T. Haritunians, S. M. Ng, W. Zhang, N. Y. Hsu, J. A. Facey, T. Luong, H. Fernandez-Hernandez, D. Li, M. Rivas, E. R. Schiff, A. Gusev, L. P. Schumm, B. M. Bowen, Y. Sharma, K. Ning, R. Remark, S. Gnjatic, P. Legnani, J. George, B. E. Sands, J. M. Stempak, L. W. Datta, S. Lipka, S. Katz, A. S. Cheifetz, N. Barzilai, N. Pontikos, C. Abraham, M. J. Dubinsky, S. Targan, K. Taylor, J. I. Rotter, E. J. Scherl, R. J. Desnick, M. T. Abreu, **H.** **Zhao**, G. Atzmon, I. Pe'er, S. Kugathasan, H. Hakonarson, J. L. McCauley, T. Lencz, A. Darvasi, V. Plagnol, M. S. Silverberg, A. M. Muise, S. R. Brant, M. J. Daly, A. W. Segal, R. H. Duerr, M. Merad, D. P. McGovern, I. Peter, J. H. Cho (2016) A Frameshift in CSF2RB predominant among Ashkenazi Jews increases risk for Crohn's disease and reduces monocyte signaling via GMCSF. *Gastroenterology*. pii: S0016-5085(16): 34717-34725.

[329] R. Polimanti, B. Z. Yang, **H. Zhao**, J. Gelernter (2016) Evidence of polygenic adaptation in the systems genetics of anthropometric traits. *PLOS One*, 11: e0160654.

[330] K. Lee, B. Li, **H. Zhao** (2016) Variable selection via additive conditional independence. *Journal of the Royal Statistical Society - Series B*, 78: 1037-1055.

[331] K. Dong, **H. Zhao**, T. Tong, X. Wan (2016) NBLDA: Negative Binomial Linear Discriminant Analysis for RNA-Seq Data. *BMC Bioinformatics*, 17: 369.

[332] A. T. Timberlake, J. Choi, S. Zaidi, Q. Lu, C. Nelson-Williams, E. D. Brooks, K. Bilguvar, I. Tikhonova, S. Mane, J. F. Yang, R. Sawh-Martinez, S. Persing, E. G. Zellner, E. Loring, C. Chuang, A. Galm, P. W. Hashim, D. M. Steinbacher, M. L. DiLuna, C. C. Duncan, K. A. Pelphrey, **H. Zhao**, J. A. Persing, R. P. Lifton (2016) Digenic inheritance of non-syndromic midline craniosynostosis via rare SMAD6 and common BMP2 alleles. *eLife*, 5: e20125.

[333] Y. Liu, **H. Zhao** (2016) Predicting synergistic effects between compounds through their structural similarity and effects on transcriptomes. *Bioinformatics*, 32: 3782-3789.

[334] Y. Hu, **H. Zhao** (2016) CCor: a whole genome network-based similarity measure between two genes. *Biometrics*, 72: 1216-1225.

[335] R. Zhu, Q. Zhao, **H. Zhao**, S. Ma (2016) Integrating multidimensional omics data for cancer outcome. *Biostatistics*, 17: 605-618.

[336] Z. Lin, C. Yang, Y. Zhu, J. C. Duchi, Y. Fu, Y. Wang, B. Jiang, M. Zamanighomi, X. Xu, M. Li, N. Sestan, **H. Zhao**, W. H. Wong (2016) AC-PCA: simultaneous dimension reduction and adjustment for confounding variation. *Proceedings of the National Academy of Sciences*, 113: 14662-14667.

[337] K. Lee, B. Li, **H. Zhao** (2016) On an additive partial correlation operator and nonparametric estimation of graphical models. *Biometrika*, 103: 513-530.

**2017 (48)**

[338] L. Hou, N. Sun, S. Mane, F. Sayward, N. Rajeevan, K.-H. Cheung, K. Cho, S. Pyarajan, M. Aslan, P. Miller, P. D. Harvey, J. M. Gaziano, J. Concato, **H. Zhao** (2017) Impacts of genotyping errors on statistical power of association tests in genomic analyses: a case study. *Genetic Epidemiology*, 41: 152-162.

[339] F. Hong, S. M. Rachidi, D. Lundgren, D. Han, X. Huang, **H. Zhao**, Y. Kimura, H. Hirano, O. Ohara, H. Udono, S. Meng, B. Liu, Z. Li (2017) Mapping the client network of a major mammalian endoplasmic reticulum heat shock protein 90 reveals its roles in innate immunity, Wnt signaling and organ development. *PLOS One*, 12: e0169260

[340] C. Heit, S. Marshall, S. Singh, X. Yu, G. Charkoftaki, **H.** **Zhao**, D. J. Orlicky, K. S. Fritz, D. C. Thompson, V. Vasiliou (2017) Catalase deletion promotes prediabetic phenotype in mice. *Free Radic Biol Med* 103: 48-56.

[341] Q. Wang, R. Polimanti, H. R. Kranzler, L. A. Farrer, **H. Zhao**, J. Gelernter (2017) Genetic factor common to schizophrenia and HIV infection is associated with risky sexual behavior. *Human Genetics*, 136: 75-83.

[342] R. Polimanti, Q. Wang, S. Meda, K. Patel, G. Pearlson, **H. Zhao**, L. Farrer, H. Kranzler, J. Gelernter (2017) The interplay between risky sexual behaviors and alcohol dependence: genome-wide association and neuroimaging support for LHPP as a risk gene. *Neuropsychopharmacology*, 42: 598-605.

[343] B. Zhao, Q. Lu, Y. Cheng, J. M. Belcher, E. D. Siew, D. E. Leaf, S. C. Body, A. A. Fox, S. W. Waikar, C. D. Collard, H. Thiessen-Philbrook, T. Alp Ikizler, L. B. Ware, C. Edelstein, A. X. Garg, M. Choi, J. Schaub, **H. Zhao**, R. P. Lifton, C. R. Parikhfor the TRIBE-AKI Consortium (2017) A genome-wide association study to identify single nucleotide polymorphisms for acute kidney injury. *American Journal of Respiratory and Critical Care Medicine*, 195: 482-490.

[344] R. Polimanti, **H. Zhao**, H. Kranzler, J. Gelernter (2017) S100A10 identified in a genome-wide gene-by-cannabis dependence interaction analysis of risky sexual behaviors. *Journal of Psychiatry and Neuroscience*, 42: 160189.

[345] D. Chung, H. J. Kim, **H. Zhao** (2017) graph-GPA: A graphical model for prioritizing GWAS results and investigating pleiotropic architecture. *PLOS Computational Biology*,13: e1005388.

[346] A. Adkins, L. Hack, T. Bigdeli, V. Williamson, G. McMichael, M. Mamdani, A. Edwards, F. Aliev, R. Chan, P. Bhandari, R. Rabbe, J. Alaimo, G. Blackwell, A. Moscati, R. Poland, B. Rood, D. Patterson, D. Walsh, of Alcohol Dependence, Collaborative Study of the Genetics; J. Whitfield, G. Zhu, G. Montgomery, A. Henders, N. Martin, A. Heath, P. Madden, J. Frank, M. Ridinger, N. Wodarz, M. Soyka, P. Zill, M. Ising, M. Nöthen, F. Kiefer, M. Rietschel, of Addiction Consortium, The German Study of the Genetics; J. Gelernter, R. Sherva, R. Koesterer, L. Almasy, **H. Zhao**, H. Kranzler, L. Farrer, B. Maher, C. Prescott, D. Dick, S.-A. Bacanu, L. Mathies, A. Davies, V. Vladimirov, M. Grotewiel, M. Bowers, J. Bettinger, B. Webb, M. Miles, K. Kendler, B. Riley (2017) Genomewide Association Study of Alcohol Dependence Identifies Risk Loci Altering Ethanol-response Behaviors in Model Organisms. *Alcoholism: Clinical and Experimental Research*, 41: 911-928.

[347] A. C. Justice, K. A. McGinnis, J. Tate, K. Xu, W. C. Becker, **H. Zhao**, J. Gelernter, H. R. Kranzler (2017) Validating harmful alcohol use as a phenotype for genetic discovery using phosphatidylethanol and a polymorphism in ADH1B. *Alcoholism: Clinical and Experimental Research*,41: 998-1003.

[348] R. Polimanti, H. Zhang, A. H. Smith, **H. Zhao**, L. A. Farrer, H. R. Kranzler, J. Gelernter (2017) Genome-wide association study of body mass index in subjects with alcohol dependence. *Addiction Biology,* 22: 535-549.

[349] R. Zhu, Y.-Q. Zhao, G. Chen, S. Ma, **H. Zhao** (2017) Greedy outcome weighted tree learning of optimal personalized treatment rules. *Biometrics*, 73: 391-400.

[350] Y. Zhang, Z. Ouyang, **H. Zhao** (2017) A statistical framework for data integration through graphical models with application to cancer genomics. *Annals of Applied Statistics,* 11: 161-184.

[351] J. Sun, J. Warren, **H. Zhao** (2017) A Bayesian semiparametric factor analysis model for subtype identification. *Statistical Application in Genetics and Molecular Biology*, 16: 145-158*.*

[352] C. A. Fragoso, M. Moreno, Z. Wang, C. Heffelfinger, L. J. Arbelaez, J. A. Aguirre, N. Franco, L. E. Romero, K. Labadie, **H. Zhao**, S. L. Dellaporta, M. Lorieux (2017) Genetic architecture of a rice nested association mapping population. *G3: Genes, Genomes, Genetics,* 7: 1913-1926.

[353] Y. Hu, Q. Lu, R. Powles, X. Yao, C. Yang, F. Fang, X. Xu, **H. Zhao** (2017) Leveraging functional annotations in genetic risk prediction for human complex diseases. *PLOS Computational Biology*, 13: e1005589.

[354] Y. Hu, Q. Lu, W. Liu, Y. Zhang, M. Li, **H. Zhao** (2017) Joint modeling of genetically correlated diseases and functional annotations increases accuracy of polygenic risk prediction. *PLOS Genetics,* 13: e1006836.

[355] M. Goffredo, N. Santoro, D. Tricò, C. Giannini, E. D'Adamo, **H. Zhao**, G. Peng, X. Yu, T. T. Lam, B. Pierpont, S. Caprio, R. I. Herzog (2017) A branched-chain amino acid-related metabolic signature characterizes obese adolescents with non-alcoholic fatty liver disease. *Nutrients, 9*, 642.

[356] X. Yan, A. Liang, J. Gomez, L. Cohn, **H. Zhao**, G. L. Chupp (2017) A novel pathway-based distance score enhances assessment of disease heterogeneity in gene expression. *BMC Bioinformatics*, 18: 309.

[357] M. I. Klein, D. F. Stern, **H. Zhao** (2017) GRAPE: A pathway template method to characterize tissue-specific functionality from gene expression profiles. *BMC Bioinformatics*, 18: 317.

[358] T. Wang, **H. Zhao** (2017) Structured subcomposition selection in regression and its application to microbiome data

analysis. *Annals of Applied Statistics,* 11: 771-791.

[359] Q. Lu, R. L. Powles, S. Abdallah, D. Ou, Q. Wang, Y. Hu, Y. Lu, W. Liu, B. Li, S. Mukherjee, P. K. Crane, **H. Zhao** (2017)Systematic tissue-specific functional annotation of the human genome highlights immune-related DNA elements for late-onset Alzheimer’s disease. *PLOS Genetics*, 13: e1006933.

[360] Z. Lin, T. Wang, C. Yang, **H. Zhao** (2017) On joint estimation of Gaussian graphical models for spatial and temporal data. *Biometrics*, 73: 769-779.

[361] T. Wang, **H. Zhao** (2017) A Dirichlet-tree multinomial regression model for associating dietary nutrients with gut microorganisms. *Biometrics*, 73: 792-801.

[362] H. Fang, C. Huang, **H. Zhao**, M. Deng (2017) gCoda: Conditional dependence network inference for compositional data. *Journal of Computational Biology*, 24: 699-708.

[363] J. Sun, J. D. Herazo-Mayab, N. Kaminski, **H. Zhao**, J. L. Warren (2017) A Dirichlet process mixture model for clustering longitudinal gene expression data. *Statistics in Medicine,* 36: 3495-3506.

[364] S. A. Streicher, A. P. Klein, S. H. Olson, L. T. Amundadottir, A. T. DeWan, **H. Zhao**, H. A. Risch (2017)Impact of sixteen established pancreatic cancer susceptibility loci in American Jews. *Cancer Epidemiology, Biomarkers & Prevention*, 26: 1540-1548.

[365] K. R. Williams, C. M. Colangelo, L. Hou, L. Chung, J. M. Belcher, T. Abbott, I. E. Hall, **H. Zhao**, L. G. Cantley, C. R. Parikh (2017) Use of a Targeted Urine Proteome Assay (TUPA) to identify protein biomarkers of delayed recovery after kidney transplant. *Proteomics Applications*,11: doi: 10.1002/prca.201600132.

[366] J. D Herazo-Maya, J. Sun, P. L. Molyneaux, Q Li , J. Villalba-Nunez, A. Tzouvelekis, H. Lynn, B. M. Juan-Guardela, C. Risquez, J. C. Osorio, X. Yan, G. Michel, N. Aurelien, K. O. Lindell, M. J. Klesen, M. F. Moffatt, W. O. Cookson, Y. Zhang, J. G. Garcia, I. Noth, A. Prasse, Z. Bar-Joseph, K. F. Gibson, **H. Zhao**, E. L. Herzog, I. O. Rosas, T. M. Maher, N. Kaminski (2017) Validation of a 52-gene risk *profile* for outcome prediction in patients with idiopathic pulmonary fibrosis: an international, multicentre, cohort study. *Lancet Respiratory Medicine,* 5: 857–868.

[367] C. Wu, **H. Zhao**, H. Fang, M. Deng (2017) Graphical model selection with latent variables. *Electronic Journal of Statistics*, 11: 3485-3521.

[368] J. Richardson, K.-Y. Lee, P. Mireji, J. Enyaru, M. Sistrom, S. Aksoy, **H. Zhao**, A. Caccone (2017) Genomic analyses of African *Trypanosoma* strains to assess evolutionary relationships and identify markers for strain identification. *PLOS Neglected Tropical Diseases*, 11: e0005949.

[369] S. C. Jin, J. Homsy, S. Zaidi, Q. Lu, S. Morton, S. R. DePalma, X. Zeng, H. Qi, W. Chang, W.-C. Hung, M. C. Sierant, S. Haider, J. Zhang, J. Knight, R. D. Bjornson, C. Castaldi, I. R. Tikhonoa, K. Bilguvar, S. M. Mane, S. J. Sanders, S. Mital, M. Russell, W. Gaynor, J. Deanfield, A. Giardini, G. A. Porter Jr., D. Srivastava, C. W. Lo, Y. Shen, W. S. Watkins, M. Yandell, H. J. Yost, M. Tristani-Firouzi, J. W. Newburger, A. E. Roberts, R. Kim, **H. Zhao**, J. R. Kaltman, E. Goldmuntz, W. K. Chung, J. G. Seidman, B. D. Gelb, C. E. Seidman2, R. P. Lifton, M. Brueckner (2017)

Contribution of rare inherited and de novo variants among 2,871 congenital heart disease probands. *Nature Genetics*, 49: 1593-1601.

[370] L. E. Mechanic, S. Lindstrom, K. M. Daily, S. K. Sieberts, C. I. Amos, H. S. Chen, N. J. Cox, M. Dathe, E. J. Feuer, M. J. Guertin, J. Hoffman, Y. Liu, J. H. Moore, C. L. Myers, M. D. Ritchie, J. Schildkraut, F. Schumacher, J. S. Witte, W. Wang, S. M. Williams, U4C Challenge Participants; [U4C Challenge Data Contributors](https://www.ncbi.nlm.nih.gov/pubmed/?term=U4C%20Challenge%20Data%20Contributors%5BCorporate%20Author%5D), E. M. Gillanders (2017) Up For A Challenge (U4C): Stimulating innovation in breast cancer genetic epidemiology. *PLOS Genetics*, 13: e1006945.

[371] G. Charkoftaki, Y. Chen, M. Han, M. Sandoval, X. Yu, **H.** **Zhao**, D. J. Orlicky, D. C. Thompson, V. Vasiliou (2017) Transcriptomic analysis and plasma metabolomics in Aldh16a1-null mice reveals a potential role of ALDH16A1 in renal function. *Chem Biol Interact* 276: 15-22.

[372] Y. Zhang, D. Zhang, X. Li, P. Ip, F. Ho, Y. Jiang, W. Sun, Q. Zhu, W. Zhu, J. Zhang, **H. Zhao**, G. Wang, X. Shen, F. Jiang (2017) Daily time-use patterns and obesity and mental health among primary school students in Shanghai: A population-based cross-sectional study. *Scientific Reports*, 7: 16200.

[373] R. Polimanti, A. Amstadter, M. Stein, L. Almli, D. Baker, L. Bierut, B. Bradley, L. Farrer, E. Johnson, A. King, H. Kranzler, A. Maihofer, J. Rice, A. Roberts, N. Saccone, **H. Zhao**, I. Liberzon, K. Ressler, C. Nievergelt, K. Koenen, J. Gelernter (2017) A putative causal relationship between genetically-determined female body shape and post-traumatic stress disorder. *Genome Medicine*, 9: 99.

[374] R. [Polimanti](https://www.ncbi.nlm.nih.gov/pubmed/?term=Polimanti%20R%5BAuthor%5D&cauthor=true&cauthor_uid=28990359), **H. Zhao**, L. A. Farr, H. R. Kranzler, J. Gelernter (2017) Ancestry-specific and sex-specific risk alleles identified in a genome-wide gene-by- alcohol dependence interaction study of risky sexual behaviors. *Neuropsychiatric Genetics*, 174: 846-853.

[375] Q. Lu, C. Jin, J. Sun, R. Bowler, K. Kechris, N. Kaminski, **H. Zhao** (2017) Post-GWAS prioritization through data integration provides novel insights on chronic obstructive pulmonary disease. *Statistics in Biosciences*, 9: 605–621.

[376] K. Yan, **H. Zhao**, H. Pang (2017) Comparison of graph- and kernel-based -omics data integration algorithms for classifying complex traits. *BMC Bioinformatics*, 18: 539.

[377] B. Zhu, N. Song, R. Shen, A. Arora, M. Machiela, L. Song, M. Landi, D. Ghosh, N. Chatterjee, V. Baladandayuthapani, **H. Zhao** (2017) Integrating clinical and multiple omics data for prognostic assessment across human cancers. *Scientific Reports*, 7: 16954.

[378] Q. Lu, B. Li, D. Ou, M. Erlendsdottir, R. L. Powles, T. Jiang, Y. Hu, D. Chang, C. Jin, W. Dai, Q. He, Z. Liu, S. Mukherjee, P. K. Crane, **H. Zhao** (2017) A powerful approach to estimating annotation-stratified genetic covariance using GWAS summary statistics. *American Journal of Human Genetics*, 101: 939-964.

[379] Y. Liu, **H. Zhao** (2017) Variable importance-weighted random forests. *Quantitative Biology*,5: 338-351.

[380] T. Wang, **H. Zhao** (2017) Constructing predictive microbial signatures at multiple taxonomic levels. *Journal of the American Statistical Association*, 112: 1022-1031.

[381] H. Zhou, R. Polimanti, B.-Z. Yang, Q. Wang, S. Han, R. Sherva, Y. Z. Nunez, **H. Zhao**, L. Farrer, H. R. Kranzler, J. Gelernter (2017) Genetic risk variants associated with comorbid alcohol dependence and major depression. *JAMA Psychiatry*, 74: 1234-1241.

[382] J. Lu, Y. Lu, X. Wang, X. Li, G. C. Linderman, C. Wu, X. Cheng, L. Mu, H. Zhang, J. Liu, M. Su, **H. Zhao**, E. S. Spatz, J. A. Spertus, F. A. Masoudi, H. M. Krumholz, L. Jiang (2017) Prevalence, awareness, treatment, and control of hypertension in China: data from 1.7 million adults in a population-based screening study (China PEACE Million Persons Project). *Lancet*, 390: 2549-2558.

[383] X. Zhang, Y. Hu, A. Justice, B. Li, Z. Wang, **H. Zhao**, J. Krystal, K. Xu (2017) DNA methylation signatures of illicit drug injection and hepatitis C are associated with HIV frailty. *Nature Communications,* 8: 2243.

[384] R. Polimanti, J. Kaufman, **H. Zhao**, H. R. Kranzler, R. J. Ursano, R. C. Kessler, M. B. Stein, J. [Gelernter](https://www.ncbi.nlm.nih.gov/pubmed/?term=Gelernter%20J%5BAuthor%5D&cauthor=true&cauthor_uid=29230810) (2017) Trauma exposure interacts with bipolar-disorder genetic risk in alcohol misuse of US soldiers. *Acta Psychiatrica Scandinavica*, 137: 148-156.

[385] N. Rauniyar, G. Peng, T. T. Lam, **H. Zhao**, G. Mor, K. R. Williams (2017) Data-Independent Acquisition (DIA) and Parallel Reaction Monitoring (PRM) mass spectrometry identification of serum biomarkers for ovarian cancer. *Biomarker Insights*, 12: 1-12.

**2018 (26)**

[386] T. Wang, **H. Zhao**, M. Chen, L. Zhu (2018) Estimating a sparse reduction for general regression in high dimensions. *Statistics and Computing*, 28: 33-46.

[387] T. Canli, L. Yu, X. Yu, **H. Zhao**, D. Fleischman, R. Wilson, P. De Jager, D. Bennett (2018) Loneliness five years ante-mortem is associated with disease-related differential gene expression in postmortem dorsolateral prefrontal cortex. *Translational Psychiatry*, 8, 2.

[388] R. Polimanti, J. Kaufman, **H. Zhao**, H. Kranzler, R. Ursano, R. Kessler, J. Gelernter, M. Stein (2018) A genome-wide gene-by-trauma interaction study of alcohol misuse in two independent cohorts identifies PRKG1 as a risk locus. *Molecular Psychiatry*, 23: 154-160.

[389] L. Zhu, J. Zhao, **H. Zhao** (2018) Pivotal variable detection of the covariance matrix and its application to high dimensional factor models. *Statistics and Computing*, 28: 775–793.

[390] K. Koenen, L. Duncan, A. Ratanatharathorn, A. Aiello, L. Almli, A. Amstadter, A. Koch, D. Baker, J. Beckham, L. Bierut, J. Bisson, B. Bradley-Davino, C.-Y. Chen, S. Dalvie, L. Farrer, S. Galea, M. Garrett, J. Gelernter, G. Guffanti, M. Hauser, E. Johnson, R. Kessler, N. Kimbrel, A. King, N. Koen, H. Kranzler, M. Logue, A. Maihofer, A. Martin, M. Miller, R. Morey, N. Nugent, J. Rice, S. Ripke, A. Roberts, N. Saccone, J. Smoller, D. Stein, M. Stein, J. Sumner, M. Uddin, R. Ursano, D. Wildman, R. Yehuda, **H. Zhao**, M. Daly, I. Liberzon, K. Ressler, C. Nievergelt (2018) Largest GWAS of PTSD (N=20,070) yields genetic overlap with schizophrenia and sex differences in heritability. *Molecular Psychiatry*, 23: 666-673.

[391] A. Gloria-Soria, W. Dunn, X. Yu, A. Vigneron, K.-Y. Lee, M. Li, B. Weiss, **H. Zhao**, S. Aksoy, A. Caccone (2018) Uncovering genomic regions associated with *Trypanosoma* infections in wild populations of the tsetse fly *Glossina fuscipes*. *G3*, 8: 887-897.

[392] Y. Zhang, M. H. Linder, A. Shojaie, Z. Ouyang, R. Shen, K. A. Baggerly, V. Baladandayuthapani, **H. Zhao** (2018) Dissecting pathway disturbances using network topology and multi-platform genomics data. *Statistics in Biosciences*, 10: 86-106.

[393] F. Villarroel-Espindola, X. Yu, I. Datar, N. Mani, M. Sanmamed, V. Velcheti, K. Syrigos, M. Toki, **H. Zhao**,L. Chen,R. S. Herbst, K. A. Schalper(2018) Spatially resolved and quantitative analysis of VISTA/PD-1H as a novel immunotherapy target in human non-small cell lung cancer (NSCLC). *Clinical Cancer Research*, 24: 1562-1573.

[394] P. Cui, T. Zhong, Z. Wang, T. Wang, **H. Zhao**, C. Liu, H. Lu (2018) Identification of human circadian genes based on time course gene expression profiles by using a deep learning method. *BBA - Molecular Basis of Disease*, 864(6 Pt B): 2274-2283.*.*

[395] J. Sun, J. D. Herazo-Maya, X. Huang, N. Kaminski, **H. Zhao** (2018) Distance-correlation based gene set analysis in longitudinal studies. *Statistical Applications in Genetics and Molecular Biology*, 17: 20170053.

[396] S. Park, **H. Zhao** (2018) Spectral clustering based on learning similarity matrix. *Bioinformatics*, 34: 2069-2076.

[397] H. J. Kim, Z. Yu, A. Lawson, **H. Zhao**, D. Chung (2018)Improving SNP prioritization and pleiotropic architecture estimation by incorporating prior knowledge using graph-GPA. *Bioinformatics*, 34: 2139-2141.

[398] J. Zhang, X. Li, N. Hawley, Z. Zheng, Z. Zou, L. Tan, Q. Chen, H. Shi, **H. Zhao**, Z. Zhang (2018) Trends in the Prevalence of overweight and obesity among Chinese school-age children and adolescents from 2010 to 2015. *Childhood Obesity*, 14: 182-188.

[399] M. Liu, Jian Peng, N. Tai, J. A. Pearson, C. Hu, J. Guo, L. Hou, **H. Zhao**, F. S. Wong, L. Wen (2018) Toll like receptor 9 Negatively Regulates Pancreatic Islet beta cell Growth and Function in a mouse model of type 1 diabetes. *Diabetologia*, 61: 2333-2343.

[400] S. Gettinger, J. Choi, N. Mani, M. Sanmamed, I. Datar, R. Sowell, V. Du, E. Kaftan, S. Goldberg, W. Dong, D. Zelterman, K. Politi, P. Kavathas, S. Kaech, X. Yu, **H. Zhao**, J. Schlessinger, R. Lifton, D. Rimm, L. Chen, R. Herbst, K. Schalper (2018) A dormant TIL phenotype defines non-small cell lung carcinomas sensitive to immune checkpoint blockers. *Nature Communications*, 9, 3196.

[401] L. A. Santos, H. Adhikarla, X. Yan, Z. Wang, D. E. Fouts, J. M. Vinetz, L. C. Jr Alcantara, R. A. Hartskeerl, M. G. A. Goris, M. Picardeau, M. G. D. Reis, J. P. Townsend, **H. Zhao**, A. I. Ko, E. A. Wunder Jr (2018) Genomic comparison among global isolates of *L. interrogans* Serovars Copenhageni and Icterohaemorrhagiae identified natural genetic variation caused by an indel. *Front Cell Infect Microbiol* 8: 193.

[402] Minică CC, Verweij KJH, van der Most PJ, Mbarek H, Bernard M, van Eijk KR, Lind PA, Liu MZ, Maciejewski DF, Palviainen T, Sánchez-Mora C, Sherva R, Taylor M, Walters RK, Abdellaoui A, Bigdeli TB, Branje SJT, Brown SA, Casas M, Corley RP, Smith GD, Davies GE, Ehli EA, Farrer L, Fedko IO, Garcia-Martínez I, Gordon SD, Hartman CA, Heath AC, Hickie IB, Hickman M, Hopfer CJ, Hottenga JJ, Kahn RS, Kaprio J, Korhonen T, Kranzler HR, Krauter K, van Lier PAC, Madden PAF, Medland SE, Neale MC, Meeus WHJ, Montgomery GW, Nolte IM, Oldehinkel AJ, Pausova Z, Ramos-Quiroga JA, Richarte V, Rose RJ, Shin J, Stallings MC, Wall TL, Ware JJ, Wright MJ, **Zhao H**, Koot HM, Paus T, Hewitt JK, Ribasés M, Loukola A, Boks MP, Snieder H, Munafò MR, Gelernter J, Boomsma DI, Martin NG, Gillespie NA, Vink JM, Derks EM (2018) Genome-wide association meta-analysis of age at first cannabis use. *Addiction*, 113: 2073-2086.

[403] J. Montalvo-Ortiz, H. Holbrook, K. O'Loughlin, C. Orr, C. Kearney, B-Z Yang, T. Wang, **H. Zhao**, R. Althoff, H. Garavan, J. Gelernter, J. Hudziak, J. Kaufman (2018) Adverse childhood experiences, epigenetic measures, and obesity in youth. *The Journal of Pediatrics*, 202: 150-156.

[404] J. Kaufman, N. F. Wymbs, J. L. Montalvo-Ortiz,C. Orr,M. D. Albaugh,R. Althoff, K. O’Loughlin, H. Holbrook,H. Garavan, C. Kearney,B.-Z. Yang, **H. Zhao**, C. Peña, E. J. Nestler, R. S. Lee, S. Mostofsky, J. Gelernter, J. Hudziak (2018)Methylation in *OTX2* and related genes, maltreatment, and depression in children. *Neuropsychopharmacology*, 43: 2204-2211.

[405] Justice AC, Smith RV, Tate JP, McGinnis K, Xu K, Becker WC, Lee KY, Lynch K, Sun N, Concato J, Fiellin DA, **Zhao H**, Gelernter J, Kranzler HR; VA Million Veteran Program (2018) AUDIT-C and ICD codes as phenotypes for harmful alcohol use: association with ADH1B polymorphisms in two U.S. populations. *Addiction*, 113: 2214-2224.

[406] J. Montalvo-Ortiz, H. Zhou, I. D'Andrea, L. Maroteaux, A. Lori, A. Smith, K. Ressler, Y. Nunez, L. Farrer, **H. Zhao**, H. Kranzler, J. Gelernter (2018) Translational studies support a role for serotonin 2B receptor gene (HTR2B) in aggression-related cannabis response. *Molecular Psychiatry*, 23(12):2277-2286.

[407] Y Xie, A. C. Ostriker, Y. Jin, H. Hu, A. J. Sizer, G. Peng, A. H. Morris. C. Ryu, E. L. Herzog, T. Kyriakides, **H. Zhao,** A. Dardik, J. Yu, J. Hwa, K. A. Martin (2018) LMO7 is a Negative Feedback Regulator of TGF-β Signaling and Fibrosis. *Circulation*, 139:679–693.

[408] Walters RK, Polimanti R, Johnson EC, McClintick JN, Adams MJ, Adkins AE, Aliev F, Bacanu SA, Batzler A, Bertelsen S, Biernacka JM, Bigdeli TB, Chen LS, Clarke TK, Chou YL, Degenhardt F, Docherty AR, Edwards AC, Fontanillas P, Foo JC, Fox L, Frank J, Giegling I, Gordon S, Hack LM, Hartmann AM, Hartz SM, Heilmann-Heimbach S, Herms S, Hodgkinson C, Hoffmann P, Jan Hottenga J, Kennedy MA, Alanne-Kinnunen M, Konte B, Lahti J, Lahti-Pulkkinen M, Lai D, Ligthart L, Loukola A, Maher BS, Mbarek H, McIntosh AM, McQueen MB, Meyers JL, Milaneschi Y, Palviainen T, Pearson JF, Peterson RE, Ripatti S, Ryu E, Saccone NL, Salvatore JE, Sanchez-Roige S, Schwandt M, Sherva R, Streit F, Strohmaier J, Thomas N, Wang JC, Webb BT, Wedow R, Wetherill L, Wills AG; 23andMe Research Team, Boardman JD, Chen D, Choi DS, Copeland WE, Culverhouse RC, Dahmen N, Degenhardt L, Domingue BW, Elson SL, Frye MA, Gäbel W, Hayward C, Ising M, Keyes M, Kiefer F, Kramer J, Kuperman S, Lucae S, Lynskey MT, Maier W, Mann K, Männistö S, Müller-Myhsok B, Murray AD, Nurnberger JI, Palotie A, Preuss U, Räikkönen K, Reynolds MD, Ridinger M, Scherbaum N, Schuckit MA, Soyka M, Treutlein J, Witt S, Wodarz N, Zill P, Adkins DE, Boden JM, Boomsma DI, Bierut LJ, Brown SA, Bucholz KK, Cichon S, Costello EJ, de Wit H, Diazgranados N, Dick DM, Eriksson JG, Farrer LA, Foroud TM, Gillespie NA, Goate AM, Goldman D, Grucza RA, Hancock DB, Harris KM, Heath AC, Hesselbrock V, Hewitt JK, Hopfer CJ, Horwood J, Iacono W, Johnson EO, Kaprio JA, Karpyak VM, Kendler KS, Kranzler HR, Krauter K, Lichtenstein P, Lind PA, McGue M, MacKillop J, Madden PAF, Maes HH, Magnusson P, Martin NG, Medland SE, Montgomery GW, Nelson EC, Nöthen MM, Palmer AA, Pedersen NL, Penninx BWJH, Porjesz B, Rice JP, Rietschel M, Riley BP, Rose R, Rujescu D, Shen PH, Silberg J, Stallings MC, Tarter RE, Vanyukov MM, Vrieze S, Wall TL, Whitfield JB, **Zhao H**, Neale BM, Gelernter J, Edenberg HJ, Agrawal A (2018) Trans-ancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nature Neuroscience*, **21**, pages1656–1669.

[409] Y. Zhu, A. M. M. Sousa, T. Gao, M. Skarica, M. Li, G. Santpere, P. Esteller-Cucala, D. Juan, L. Ferrández-Peral, F. O. Gulden, M. Yang, D. J. Miller, T. Marques-Bonet, Y. Imamura Kawasawa, H. Zhao, N. Sestan (2018) Spatiotemporal transcriptomic divergence across human and macaque brain development. *Science* 362(6420).

[410] Li M, Santpere G, Imamura Kawasawa Y, Evgrafov OV, Gulden FO, Pochareddy S, Sunkin SM, Li Z, Shin Y, Zhu Y, Sousa AMM, Werling DM, Kitchen RR, Kang HJ, Pletikos M, Choi J, Muchnik S, Xu X, Wang D, Lorente-Galdos B, Liu S, Giusti-Rodríguez P, Won H, de Leeuw CA, Pardiñas AF; BrainSpan Consortium; PsychENCODE Consortium; PsychENCODE Developmental Subgroup, Hu M, Jin F, Li Y, Owen MJ, O'Donovan MC, Walters JTR, Posthuma D, Levitt P, Weinberger DR, Hyde TM, Kleinman JE, Geschwind DH, Hawrylycz MJ, State MW, Sanders SJ, Sullivan PF, Gerstein MB, Lein ES, Knowles JA, Sestan N (2018) Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* 362(6420).

[411] X. Zhang, Y. Hu, B. E. Aouizerat, G. Peng, V. C. Marconi, M. Corley, H. Todd, B. Kendall, **H. Zhao**, J. Krystal, A. Justice, K. Xu (2018) Machine learning selected smoking-associated DNA methylation signatures that predict HIV prognosis and mortality. *Clinical Epigenetics*, 10: 155.

**2019 (31)**

[412] G Peng, C. A. de Fontnouvelle, G. M. Enns, T. M. Cowan, **H. Zhao**, C. Scharfe (2019) Elevated methylmalonic acidemia (MMA) screening markers in Hispanic and preterm newborns. *Molecular Genetics and Metabolism,* 126: 39-42.

[413] K. Radhakrishnan, Mihaela Aslan, K. M. Harrington, R. H. Pietrzak, G. Huang, S. Muralidhar, K. Cho, R. Quaden, D. Gagnon, S. Pyarajan, N. Sun, **H. Zhao**, M. Gaziano, J. Concato, M. B. Stein, J. Gelernter (2019) Genomics of Posttraumatic Stress Disorder in Veterans: Methods and Rationale for Veterans Affairs Cooperative

Study #575B. *International Journal of Methods in Psychiatric Research*, 2019: e1767.

[414] Y. Hu, M. Li, Q. Lu, H. Weng, J. Wang, S. M. Zekavat, Z. Yu, B. Li, J. Gu, S. Muchnik, Y. Shi, B. W. Kunkle, S. Mukherjee, P. Natarajan, A. Naj, A. Kuzma, Y. Zhao, P. K. Crane, Alzheimer's Disease Genetics Consortium, H. Lu, **H. Zhao** (2019) A statistical framework for cross-tissue transcriptome-wide association analysis. *Nature Genetics*, 51: 568-576.

[415] D. Zhao, **H. Zhao**, P. D. Cleary (2019) Understanding the determinants of public trust in the health care system in China: an analysis of cross sectional survey*. Journal of Health Services Research & Policy*, 24: 37-43.

[416] D. Zhao, **H. Zhao**, P. Cleary (2019) International variations in trust in health care systems. *The International Journal of Health Planning and Management*, 34: 130-139.

[417] G. Peng, P. Shen, N. Gandotra, A. Le, E. Fung, L. Jelliffe-Pawlowski, R. W. Davis, G. M. Enns, **H. Zhao**, T. M. Cowan, C. Scharfe (2019) Combining newborn metabolic and DNA analysis for second-tier testing of methylmalonic academia. *Genetics in Medicine*, 21: 896-903.

[418] S. Park, **H. Zhao** (2019) Sparse principal component analysis with missing observations. *Annals of Applied Statistics*, 13: 1016–1042.

[419] L. Zeng, J. L. Warren, **H. Zhao** (2019) Phylogeny-based tumor subclone identification using a Bayesian feature allocation model. *Annals of Applied Statistics*, 13: 1212–1241.

[420] M. Scott, K. Quinn, Q. Li, R. Carroll, H. Warsinske, F. Vallania, S. Chen, M. Carns, K. Aren, J. Sun, K. Koloms, J. Lee, J. Baral, J. Kropski, **H. Zhao**, E. Herzog, F. Martinez, B. B. Moore, M. Hinchcliff, J. Denny, N. Kaminski, J. Herazo-Maya, N. Shah, P. Khatri (2019) Identification and validation of a cellular biomarker for poor outcomes in fibrosis: a multicenter cohort study. *The Lancet Respiratory Medicine*, 7: 497-508.

[421] K. M. Harrington, R. Quaden, M. B. Stein, J. P. Honerlaw, S. Cissell, R. H. Pietrzak, **H. Zhao**, K. Radhakrishnan, M. Aslan, M. Gaziano, J. Concato, D. Gagnon, J. Gelernter, K. Cho (2019) Validation of an electronic medical record-based algorithm for identifying posttraumatic stress disorder in US veterans. *Journal of Traumatic Stress*, 32: 226-237.

[422] H. Kranzler, H. Zhou, R. Kember, R. V. Smith, A. Justice, S. Damrauder, P. Tsao, D. Klarin, D. Rader, N. Katsanis, Z. Cheng, J. Tate, W. Becker, J. Concato, K. Xu, R. Polimanti, **H. Zhao**, J. Gelernter (2019) Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations. *Nature Communications*, 10, 1499.

[423] S. M. Zekavat, K. Aragam, C. Emdin, A. V. Khera, D. Klarin, **H. Zhao**, P. Natarajan (2019) Genetic association of finger photoplethysmography-derived arterial stiffness index with blood pressure and coronary artery disease. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 39: 1253-1261.

[424] S. M. Zekavat, C. Roselli, G. Hindy, S. A. Lubitz, P. T. Ellinor, **H. Zhao**, P. Natarajan (2019) Genetic link between arterial stiffness and atrial fibrillation. *Circ Genom Precis Med.* 12(6): e002453

[425] M. P. Menden, D. Wang, M. J. Mason, B. Szalai, K. C. Bulusu, Y. Guan, T. Yu, J. Kang, M. Jeon, R. Wolfinger, T. Nguyen, M. Zaslavskiy; **AstraZeneca-Sanger Drug Combination DREAM Consortium**, I. S. Jang, Z. Ghazoui, M. E. Ahsen, R. Vogel, E. C. Neto, T. Norman, E. K. Y. Tang, M. J. Garnett, G. Y. D. Veroli, S. Fawell, G. Stolovitzky, J. Guinney, J. R. Dry, J. Saez-Rodriguez (2019) Community assessment to advance computational prediction of cancer drug combinations in a pharmacogenomic screen.*Nature Communications* 10: 2674.

[426] Z. Duan, Y. Qiao, J. Lu, H. Lu, W. Zhang, F. Yan, C. Sun, Z. Hu, Z. Zhang, G. Li, H. Chen, Z. Xiang, Z. Zhu, **H.** **Zhao**, Y. Yu, C. Wei (2019) HUPAN: a pan-genome analysis pipeline for human genomes. *Genome Biology* 20:149.

[427] B. Li, Q, Lu, **H. Zhao** (2019) An evaluation of non-coding genome annotation tools through enrichment analysis of fifteen genome-wide association studies. *Briefings in Bioinformatics* 20: 995-1003.

[428] J. Sun, J. D. Herazo-Maya, J-L Wang, N. Kaminski, **H. Zhao** (2019) LCox: A tool for selecting genes related to survival outcomes using longitudinal gene expression data. *Statistical Applications in Genetics and Molecular Biology*, 18, 2.

[429] Y. Wang, **H. Zhao**, M. Deng, H. Fang, D. Yang (2019) Network clustering analysis using mixture exponential-family random graph models and its application in genetic interaction data. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 16: 1743-1752.

[430] K. A. McGinnis, A. C. Justice, J. P. Tate, H. R. Kranzler, H. A. Tindle, W. C. Becker, J. Concato, J. Gelernter, B. Li, X. Zhang, **H. Zhao**, K. Crothers, K. Xu, For the VACS Project Group (2019) Using DNA methylation to validate an electronic medical record phenotype for smoking. *Addiction Biology*, 24: 1056-1065.

[431] G. Peng, R. Wilson, Y. Tang, T. T. Lam, A. C. Nairn, K. Williams, **H. Zhao** (2019) ProteomicsBrowser: MS/Proteomics Data Visualization and Investigation. *Bioinformatics*, 35: 2313-2314.

[432] Y. Liu, J. Warren, **H. Zhao** (2019) A hierarchical Bayesian model for single-cell clustering using RNA-sequencing data. *Annals of Applied Statistics*, Volume 13: 1733-1752.

[433] T. Wang, C. Yang, **H. Zhao** (2019) Prediction analysis for microbiome sequencing data. *Biometrics*, 75: 875-884.

[434] J. Gelernter, N. Sun, R. Polimanti, R. Pietrzak, D. F. Levey, Q. Lu, Y. Hu, B. Li, K. Radhakrishnan, M. Aslan, K.-H. Cheung, Y. Li, N. Rajeevan, F. Sayward, K. Harrington, Q. Chen, K. Cho, J. Honerlaw, S. Pyarajan, T. Lencz, R. Quaden, Y. Shi, H. Hunter-Zinck, J. M. Gaziano, H. R. Kranzler, J. Concato, **H. Zhao**, M. B. Stein, on behalf of the Department of Veterans Affairs Cooperative Studies Program (#575B) and Million Veteran Program (2019) Genomewide association study of maximum habitual alcohol intake in >140,000 US European- and African-American veterans yields novel risk loci. *Biological Psychiatry*, 86: 365-376.

[435] J. Gelernter, N. Sun,R. Polimanti, R. Pietrzak, J. Bryois, Q. Lu, Y Hu, B. Li, K. Radhakrishnan, M. Aslan, K.H. Cheung, Y. Li, N. Rajeevan, F. Sayward, K. Harrington, Q. Chen, K. Cho, S. Pyarajan, P.F. Sullivan, R. Quaden, Y. Shi, H. Hunter-Zinck, J. M. Gaziano, J. Concato, **H. Zhao**, M.B. Stein, on behalf of the Department of Veterans Affairs Cooperative Studies Program (#575B) and Million Veteran Program (2019) GWAS of PTSD re-experiencing symptoms in >165,000 US veterans yields new biological knowledge. *Nature Neuroscience*, 22: 1394-1401.

[436] M. T. Guerra, R. M. Florentino, A. Franca, A. C. Lima Filho, M. L. Dos Santos, R. C. Fonseca, F. O. Lemos, M. C. Fonseca, E. Kruglov, A. Mennone, B. Njei, J. Gibson, F. Guan, Y. C. Cheng, M. Ananthanaravanam, J. Gu, J. Jiang, **H. Zhao**, C. X. Lima, P. T. Vidigal, A. G. Oliverira, M. H. Nathanson, M. F. Leite (2019) Expression of the type 3 InsP3 receptor is a final common event in the development of hepatocellular carcinoma. *Gut*, 68: 1676-1687.

[437] L. Zeng, Z. Yu, **H. Zhao** (2019) A pathway-based kernel boosting method for sample classification using genomic data. *Genes*, 10: 670.

[438] M. Liu, J. Sun, J. D. Herazo-Maya, N. Kaminski, **H. Zhao** (2019) Joint Models for Time-to-Event Data and Longitudinal Biomarkers of High Dimension. *Statistics in Biosciences*, 11: 614–629.

[439] X. Wu, C. Heffelfinger, H. Zhao, S. L. Dellaporta (2019) Benchmarking variant identification tools for plant diversity discovery. *BMC Genomics* 20: 701.

[440] C. Nievergelt, ….., **H. Zhao**, L. Zoellner, I. Liberzon, K. Ressler, M. Haas, K. Koenen (2019) International meta-analysis of PTSD genome-wide association studies identifies sex- and ancestry-specific genetic risk loci. *Nature Communications*, 10: 4558.

[441] J. Jiang, J. Huang, J. Gu, X. Cai, **H. Zhao**, H. Lu (2019) Genomic analysis of a spinal muscular atrophy (SMA) discordant family identifies a novel mutation in TLL2, an activator of growth differentiation factor 8 (myostatin): a case report. *BMC Med Genet* 20: 204.

[442] W. Wu, Z. Wang, K. Xu, X. Zhang, A. Amei, J. Gelernter, **H. Zhao**, A. C. Justice, Z. Wang (2019) Retrospective association analysis of longitudinal binary traits identifies important loci and pathways in cocaine use. *Genetics*, 213(4):1225-1236.

**2020 (4)**

[443] D. Tang, S. Park, **H. Zhao** (2019) NITUMID: Nonnegative Matrix Factorization-based Immune-TUmor MIcroenvironment Deconvolution. *Bioinformatics*, 36: 1344-1350.

[444] S. Song, W. Jiang, L. Hou, **H. Zhao** (2020) Leveraging effect size distributions to improve polygenic risk scores derived from summary statistics of genome-wide association studies. *PLOS Computational Biology*, 16: e1007565.

[445] P. D. Harvey, N. Sun, T. B. Bigdel, A. H. Fanous, M. Aslan, A. K. Malhotra, Q. Lu, Y. Hu, B. L, Q. Chen, S. Mane, P. Miller, N. Rajeevan, F. Sayward, K. H. Cheung, Y. Li, T. A. Greenwood, R. E. Gur, D. L. Braff; Consortium on the Genetics of Schizophrenia (COGS), M. Brophy, S. Pyarajan, T. J. O'Leary, T. Gleason, R. Przygodszki, S. Muralidhar, J. M. Gaziano, J. Concato, **H. Zhao**, L. J. Siever (2020) Genome-wide association study of cognitive performance in US veterans with schizophrenia or bipolar disorder. *Neuropsychiatric Genetics*, 183: 181-194.

[446] S. Dalvie, A. X. Maihofer, J. R. I. Coleman, B. Bradley, G. Breen, L. A. Brick, C. Y. Chen, K. W. Choi, L. E. Duncan, G. Guffanti, M. Haas, S. Harnal, I. Liberzon, N. R. Nugent, A. C. Provost, K. J. Ressler, K. Torres, A. B. Amstadter, A. S. Bryn Austin, D. G. Baker, E. A. Bolger, R. A. Bryant, J. R. Calabrese, D. L. Delahanty, L. A. Farrer, N. C. Feeny, J. D. Flory, D. Forbes, S. Galea, A. Gautam, J. Gelernter, R. Hammamieh, M. Jett, A. G. Junglen, M. L. Kaufman, R. C. Kessler, A. Khan, H. R. Kranzler, L. A. M. Lebois, C. Marmar, M. R. Mavissakalian, A. McFarlane, M. O. Donnell, H. K. Orcutt, R. H. Pietrzak, V. B. Risbrough, A. L. Roberts, A. O. Rothbaum, P. Roy-Byrne, K. Ruggiero, A. V. Seligowski, C. M. Sheerin, D. Silove, J. W. Smoller, M. B. Stein, M. H. Teicher, R. J. Ursano, M. Van Hooff, S. Winternitz, J. D. Wolff, R. Yehuda, **H. Zhao**, L. A. Zoellner, D. J. Stein, K. C. Koenen, C. M. Nievergelt (2020) Genomic influences on self-reported childhood maltreatment. *Translational Psychiatry* 10(1): 38.

**In Press (12)**

[447] S. Mahajan, J. Gu, Y. Lu, R. Khera, E. S. Spatz, M. Zhang, N. L. Sun, X. Zheng, **H. Zhao**, H. Lu, Z. J. Ma, H. M. Krumholz (2020) Hemodynamic phenotypes of hypertension based on cardiac output and systemic vascular resistance. The *American Journal of Medicine*, in press.

[448] Y. Song, **H. Zhao**, T. Wang (2020) An adaptive independence test for microbiome community data. *Biometrics*, in press

[449] M. A. Munn-Chernoff, ………**H. Zhao**, H., B. M. Neale, T. D. Wade, A. C. Heath, G. W. Montgomery, N. G. Martin, P. F. Sullivan, J. Kaprio, G. Breen, J. Gelernter, H. J. Edenberg, C. M. Bulik, A. Agrawal (2020) Shared Genetic Risk between Eating Disorder- and Substance-Use-Related Phenotypes: Evidence from Genome-Wide Association Studies. *Addiction Biology*, in press.

[450] S. Park, H. Xu, **H. Zhao** (2020) Integrating multidimensional data for clustering analysis with applications to cancer patient data. *Journal of American Statistical Association*, in press.

[451] M. Li, R. Liu, J. Lin, V. Bunn, **H. Zhao** (2020) Bayesian semi-parametric design (BSD) for adaptive dose-finding with multiple strata. *Journal of Biopharmaceutical Statistics*, in press.

[452] S. Mahajan, J. Gu, C. Caraballo, Y. Lu, E. S. Spatz, **H. Zhao**, M. Z. Zhang, N. L. Sun, X. Zheng, H. Lu, H. Yuan, Z. J. Ma, H. M. Krumholz (2020) Relationship of age with the hemodynamic parameters in individuals with elevated blood pressure. *Journal of the American Geriatrics Society,* in press.

[453] K. A. Schalper, M. Carleton, M. Zhou, T. Chen, Y. Feng, S.-P. Huan, A. Walsh, V. Baxi, D. Pandya, T. Baradet, D. Locke, Q. Wu, T. P. Reilly, P. Philips, V. Nagineni, J. Gu, H. Zhao, J. L. Pérez-Gracia, M. F. Sanmamed, I. Melero (2020) Elevated serum interleukin 8 is associated with enhanced intratumor neutrophils and reduced clinical benefit with immune checkpoint inhibitors. *Nature Medicine*, in press.

[454] H. Hunter-Zinck, Y. Shi, M. Li, B. R. Gorman, S.-G. Ji, N. Sun, T. Webster, A. Liem, P. Hsieh, P. Devineni, P. Karnam, L. Radhakrishnan, J. Schmidt, T. L. Assimes, J. Huang, C. Pan, D. Humphries, M. Brophy, J. Moser, S. Muralidhar, G. D. Huang, R. Przygodzki, J. Concato, J. M. Gaziano, J. Gelernter, C. J. O’Donnell, E. R. Hauser, **H. Zhao**, T. J. O’Leary, P. S. Tsao, S. Pyarajan, on behalf of the VA Million Veteran Program (2020) Measuring genetic variation in the multi-ethnic Million Veteran Program (MVP). *American Journal of Human Genetics*, in press.

[455] G. Peng, Y. Tang, T. Cowan, G. Enns, **H. Zhao**, C. Scharfe (2020) Reducing false-positive results in newborn screening using machine learning. *International Journal of Neonatal Screening*, in press.

[456] K-Y Lee, T. Liu, B. Li, **H. Zhao** (2020) Learning causal networks via additive faithfulness. *Journal of Machine Learning Research*, in press.

[457] Y. Kong, H. Li, Y. Ren, G. Genchev, X. Wang, **H. Zhao**, Z. Xie, H. Lu (2020) Automated yeast cells counting using a parallel U-Net based two-stage framework. *OSA Continuum*, in press.

[458] G. Peng, Y. Tang, N. Gandotra, G. M. Enns, T. M. Cowan, **H. Zhao**, C. Scharfe (2020) Ethnic variability in newborn metabolic screening markers associated with false-positive outcomes. *Journal of Inherited Metabolic Disease*, in press.

**Refereed Proceeding Articles (24)**

[1] **H. Zhao**, L. J. Sheffield, A. J. Pakstis, M. P. Knauert, and K. K. Kidd. (1999) A more powerful method to evaluate p-values in GENEHUNTER. *Genetic Epidemiology*, 17: S415-S420.

[2] L. J. Sheffield, M. P. Knauert, A. J. Pakstis, **H. Zhao**, and K. K. Kidd. (1999) Analyses of the COGA data set in one ethnic group with examinations of alternative definitions of alcoholism. *Genetic Epidemiology*, 17, S319-S324.

[3] S. Zhang, K. Zhang, J. Li, F. Sun, and **H. Zhao**. (2001) Test of association for quantitative traits in general pedigrees: the quantitative pedigree disequilibrium test. *Genetic Epidemiology*, 21: S370-S375.

[4] J. Li, D. Wang, J. Dong, R. Jiang, K. Zhang, S. Zhang, **H. Zhao**, and F. Sun.(2001) The power of transmission disequilibrium tests for quantitative traits. *Genetic Epidemiology*, 21: S632-S637.

[5] S. Zhang, K. Zhang, J. Li, and **H. Zhao**. (2002) On a family-based haplotype pattern mining method for linkage disequilirbium mapping. *Pacific Symposium on Biocomputing*, 100-111.

[6] K. H. Cheung, K. White, J. Hager, M. Gerstein, V. Reinke, K. Nelson, P. Masiar, R. Srivastava, Y. Li, J. Li, J. M. Li, **H. Zhao**, D. B. Allison, M. Snyder, P. Miller, K. Williams. (2002) YMD: A microarray Database for Large-scale Gene Expression Analysis. *Proceedings of AMIA Symposium* 140-144.

[7] S. Wang, S. Huang, N. Liu, L. Chen, C. Oh, **H. Zhao** (2005) Whole-genome linkage analysis in mapping alcoholism genes using single nucleotide polymorphisms and microsatellites. *BMC Genetics,* 6(Suppl 1): S28.

[8] C. Oh, S. Wang, N. Liu, L. Chen, **H. Zhao** (2005) A Bayesian genome screening of maximum number of drinks as an alcoholism phenotype with the new Haseman-Elston method. *BMC Genetics,* 6(Suppl 1): S116.

[9] N. Liu, L. Chen, S. Wang, C. Oh, **H. Zhao** (2005) Comparison of single nucleotide polymorphisms and microsatellites in inference of population structure. *BMC Genetics,* 6(Suppl 1): S26*.*

[10] S. Huang, S. Wang, N. Liu, L. Chen, C. Oh, **H. Zhao** (2005) Whole-genome association analysis to identify markers associated with recombination rates using single nucleotide polymorphisms and microsatellites. *BMC Genetics,* 6(Suppl 1): S51.

[11] L. Chen, N. Liu, S. Wang, C. Oh, N. J. Carriero, **H. Zhao** (2005) Whole genome association studies on alcoholism comparing different phenotypes using single nucleotide polymorphisms and microsatellites. *BMC Genetics,* 6(Suppl 1): S130.

[12] L. A. Cupples, J. Bailey, K. C. Cartier, C. T. Falk, K-Y Liu, Y. Ye, R. Yu, **H. Zhao** (2005) Group 14: Data Mining. *Genetic Epidemiology*, 29 Suppl 1: S103-9.

[13] S. Huang, D. Ballard, **H. Zhao** (2007) The role of heritability in mapping expression quantitative trait loci. *BMC Proceedings,* 1: S86*.*

[14] C. Aporntewan, D. H. Ballard, J. Y. Lee, J. S. Lee, Z. Wu, **H. Zhao** (2009) Gene hunting of the GAW16 rheumatoid arthritis data using rough set theory. *BMC Proceedings,* Supplement 7:S126.

[15] Z. Wu, C. Aporntewan, D. H. Ballard, J. Y. Lee, J. S. Lee, **H. Zhao** (2009)Two-stage joint selection method to identify potential candidate markers from genome-wide association studies. *BMC Proceedings,* Supplement 7:S29.

[16] D. H. Ballard, C. Aporntewan, J. Y. Lee, J. S. Lee, Z. Wu, **H. Zhao** (2009) A pathway analysis applied to GAW 16 genome wide rheumatoid arthritis data. *BMC Proceedings,* Supplement 7:S91.

[17] Z. He, **H. Zhao**, W. Yu (2011) [Score regularization for peptide identification.](http://www.ncbi.nlm.nih.gov/pubmed/21342549) *BMC Bioinformatics*, 12 Supplement 1: S2.

[18] G. Li, J. Ferguson, W. Zheng, J. S. Lee, X. Zhang, L. Li, J. Kang, X. Yan, **H. Zhao** (2011) Large-scale risk prediction applied to Genetic Analysis Workshop 17 mini-exome sequence data. *BMC Proceedings,* 5 Supplement 9: S46.

[19]X. Yan, L. Li, J. S. Lee, W. Zheng, J. Ferguson, **H. Zhao** (2011) Detecting functional rare variants by collapsing and incorporating functional annotation in Genetic Analysis Workshop 17 mini-exome data. *BMC Proceedings,* 5 Supplement 9: S27.

[20] L. Li, W. Zheng, J. S. Lee, X. Zhang, J. Ferguson, X. Yan, **H. Zhao** (2011) Collapsing-based and kernel-based single gene analyses applied to Genetic Analysis Workshop 17 mini-exome data. *BMC Proceedings,* 5 Supplement 9: S117.

[21] J. Kang, W. Zheng, L. Li, J. S. Lee, X. Yan, **H. Zhao** (2011) Utilization of Bayesian Network to dissect the complexity of genetic disease: Application to the GAW17 Simulated Data. *BMC Proceedings,* 5 Supplement 9: S37.

[22] C. Yang, C. Li, M. Chen, X. Chen, L. Hou, **H. Zhao** (2014) A penalized linear mixed model for genomic prediction using pedigree structures. *The Proceedings of Genetic Analysis Workshop 18*, *BMC Proceedings,* 8(Suppl 1): S67*.*

[23] C. Li, C. Yang, M. Chen, X. Chen, L. Hou, **H. Zhao** (2014) Adjustment of familial relatedness in association test for rare variants. *The Proceedings of Genetic Analysis Workshop 18*, *BMC Proceedings,* 8(Suppl 1): S39*.*

[24] M. Chen, C. Yang, C. Li, L. Hou, X. Chen, **H. Zhao** (2014) Admixture mapping analysis in the context of GWAS with GAW18 data. *The Proceedings of Genetic Analysis Workshop 18*, *BMC Proceedings,* 8(Suppl 1): S3*.*

**Letters/Comments (3)**

[1] **H. Zhao**. (1999) The interpretation of the parameters in the transmission/disequilibrium test. *American Journal of Human Genetics*, 64: 326-328.

[2] **H. Zhao**, K. R. Merikangas, and K. K. Kidd. (2000) Reply to Kong and Nicolae. *American Journal of Human Genetics*, 67: 1355-1356.

[3] N. Sun, **H. Zhao** (2014) Putting things in order. *Proc Natl Acad Sci USA* 111: 16236-16237.

**Journal Review Articles (21)**

[1] H. Zhang, **H. Zhao**, and K. R. Merikangas. (1997) Strategies to identify genes for complex diseases. *Annals of Medicine*, 29: 493-498.

[2] **H. Zhao**.(2000) Family-based association studies. *Statistical Methods in Medical Research*, 9: 563-587.

[3] **H. Zhao**, R. Pfiffer, M. H. Gail. (2003) Haplotype analysis in population genetics and association studies. *Phamacogenomics*, 4: 171-178.

[4] N. Sun, **H. Zhao**. (2004) Genomic approaches in dissecting complex biological pathways. *Pharmacogenomics*, 5: 163-179.

[5] Y. Xia, H. Yu, R. Jansen, M. Seringhaus, S. Baxter, D. Greenbaum, **H. Zhao**, M. Gerstein. (2004) Analyzing cellular biochemistry in terms of molecular networks. *Annual Review of Biochemistry*, 73: 1051-1087.

[6] Stone KL, Bjornson RD, Blasko GG, Bruce C, Cofrancesco R, Carriero NJ, Colangelo CM, Crawford JK, Crawford JM, Dasilva NC, Deluca JD, Elliott JI, Elliott MM, Flory PJ, Folta-Stogniew EJ, Gulcicek E, Kong Y, Lam TT, Lee JY, Lin A, Lopresti MB, Mane SM, McMurray WJ, Tikhonova IR, Westman S, Williams NA, Wu TL, **Zhao H**, Williams KR. (2007) Keck foundation biotechnology resource laboratory, Yale university. *Yale J Biol Med*. 80: 195-211.

[7] N. Liu, K. Zhang, **H. Zhao** (2008) Haplotype Association Analysis. *Advances in Genetics,* 60: 335-405.

[8] Y. Liu, I. Kim, **H. Zhao** (2008) Protein interaction predictions from diverse sources. *Drug Discovery Today*, 13: 409-416.

[9] N. Sun, **H. Zhao** (2009) Reconstructing transcriptional regulatory networks through genomics data. *Statistical Methods in Medical Research*, 18: 595-617.

[10] W. Zheng, T. A. Gianoulis, K. J. Karczewski, **H. Zhao**, M. Snyder (2011) Regulatory variation within and between species. *Annual Reviews of Genomics and Human Genetics*, 12: 327-346.

[11] J. Cao, X. Qi, **H. Zhao** (2012) Modelling gene regulation networks using ordinary differential equations. *Methods in Molecular Biology*, 802: 185-197.

[12] R. Luo, **H. Zhao** (2012) Protein quantitation using iTRAQ: Review on the sources of variations and analysis of nonrandom missingness. *Statistics and Its Interface*, 5: 99-107.

[13] L. Li, X. Zhang, **H. Zhao** (2012) eQTL. *Methods in Molecular Biology*, 871: 265-279.

[14] H. Ma, **H. Zhao** (2013) Drug target inference through pathway analysis of genomics data. *Advanced Drug Delivery Reviews*, 65: 966-972.

[15] H. Lin, **H. Zhao** (2013) A Review of post-GWAS prioritization approaches. *Frontiers in Genetics (Statistical Genetics and Methodology)*, 4: 280.

[16] H. Chun, M. Chen, B. Li, **H. Zhao** (2013) Network inference via conditional graphical models with multiple sources of genomic data. *Frontiers in Genetics (Bioinformatics and Computational Biology)*, 4: 294.

[17] X. Zhou, C. Yang, **H. Zhao**, W. Yu (2014) Low-rank modeling and its applications in image analysis. *ACM Computing Surveys*, 47: 36.

[18] Q. Wang, Q. Lv, **H. Zhao** (2015) A review of study designs and statistical methods for genomic epidemiology studies using next generation sequencing. *Frontiers in Genetics* (*Applied Genetic Epidemiology*), 6: 149.

[19] C. Yang, C. Li, Q. Wang, D. Chung, **H. Zhao** (2015) Implications of pleiotropy: Challenges and opportunities for mining Big Data in Biomedicine. *Frontiers in Genetics* (*Applied Genetic Epidemiology*), 6: 229.

[20] M. Chen, **H. Zhao** (2019) Next-generation sequencing in liquid biopsy: cancer screening and early detection. *Human Genomics*, 13: 34.

[21] N. Sun, H. Zhao (2020) Statistical Methods in Genome-Wide Association Studies. *Annual Reviews of Biomedical Data Science*, in press.

**Book Chapters (28)**

[1] **H. Zhao** and T. P. Speed. (1999) On a Markov model for chromatid interference. *Statistics in Molecular Biology* (*IMS Lecture Notes 33*), 1-20.

[2] T. P. Speed and **H. Zhao** (2001) Chromosome maps. *Handbook of Statistical Genetics*, Chapter 1: 3-39.

[3] **H. Zhao**, B. Wu, N. Sun. (2003) DNA-protein binding and gene expression patterns. In *Science and Statistics: A Festschrift for Terry Speed*. *IMS Lecture Notes-Monograph Series*, edited by D. R. Goldstein, Vol 40, pp 259-274.

[4] E. Gulcicek, C. Colangelo, W. McMurray, K. Stone, B. Wu, T. Wu, H. Spratt, A. Kurosky, **H. Zhao**, K. Williams (2005) Proteomics and the Analysis of Proteomic Data. *Current Protocols in Bioinformatics* (Davison, D., Page, R., Petsko, G., Stormo, G., and Stein, L. eds) John Wiley and Sons, Indianapolis, IN, Chapter 13.

[5] W. Yu, B. Wu, J. Liu, X. Li, K. Stone, K. Williams, **H. Zhao** (2006) MALDI-MS Data Analysis for Disease Biomarker Discovery. *Methods in Molecular Biology*, *New and Emerging Proteomics Techniques*, edited by D. Nedelkov and R. W. Nelson, Humana Press, 328: 199-216.

[6] W. Yu, B. Wu, T. Huang, X. Li, K. Williams, **H. Zhao** (2006) Statistical methods in proteomics. *Springer Handbook of Engineering Statistics,* edited by H. Pham, Springer, 623-638*.*

[7] N. Sun, **H. Zhao** (2006) A misclassification model for inferring transcriptional regulatory networks. *Bayesian Inference for Gene Expression and Proteomics*, edited by K-A Do, P. Mueller and M. Vannucci, 347-365.

[8] Y. Liu, N. Sun, J. Liu, L. Chen, M. McIntosh, L. Zheng, **H. Zhao** (2007) Identifying stage-specific genes by combining information from two different types of oligonucleotide arrays. *Methods of Microarray Data Analysis V,* edited by P. McConnell, S. Lin and P. Hurban, 59-74*.*

[9] Z. Feng, N. Liu, **H. Zhao** (2007) Haplotype inference and association analysis in unrelated samples. *Current Topics in Human Genetics, Studies in Human Diseases*, edited by HW Deng, H. Shen, Y-J Liu, H. Hu, 123-159.

[10] S. Huang, D. Ballard, Z. Wu, **H. Zhao** (2007) Expression quantitative trait loci mapping. *Current Topics in Human Genetics, Studies in Human Diseases*, edited by HW Deng, H. Shen, Y-J Liu, H. Hu,397-431.

[11] J. Wang, G. Zou, **H. Zhao** (2007) DNA pooling: Methods and applications in association studies. *Current Topics in Human Genetics, Studies in Human Diseases*, edited by HW Deng, H. Shen, Y-J Liu, H. Hu, 179-208.

[12] H. Pang, I. Kim, **H. Zhao** (2008) Pathway-based methods for analyzing microarray data. *Analysis of Microarray Data: Network based Approaches*, edited by F. Emmert-Streib and M. Dehmer, Wiley-VCH, Weinheim, Germany, 355-384.

[13] Y. Liu, I. Kim, **H. Zhao** (2009) Protein interaction predictions from diverse sources. *New Developments in Biostatistics and Bioinformatics*, edited by J. Fan, X. Lin, and J. S. Liu, World Scientific, New Jersey, 159-178.

[14] I. Kim, Y. Liu, **H. Zhao** (2010) Sparsity priors for protein-protein interaction predictions. Bayesian Modeling in Bioinformatics, edited by D. Dey, S. Ghosh, and B. Mallick, Chapman & Hall, 255-269.

[15] N. Sun, **H. Zhao** (2009) Current trends and statistical issues in genome wide association studies. *Statistical Bioinformatics*, edited by J. Lee, Wiley-Blackwell, 283-308.

[16] X. Zhang, **H. Zhao** (2010) An evaluation of gene module concepts in the interpretation of gene expression data. *Frontiers in Computational and Systems Biology,* edited by J. Feng, W. Fu, F. Sun, Springer-Verlag, London, 331-349.

[17] K. Zhang, **H. Zhao** (2010) Family based association studies. *Handbook of Statistical Genetics*, edited by S. Lin and H. Zhao, Springer-Verlag Berlin Heidelberg, 191-240.

[18] N. Sun, **H. Zhao** (2010) Bayesian methods in genomics and proteomics studies. *Knowledge-Based Bioinformatics*, edited by M. Ramoni and G. Alterovitz, Wiley, 125-136.

[19] H. Chun, J. Kang, X. Zhang, M. Deng, H. Ma, **H. Zhao** (2011) Reverse engineering of gene regulation networks with an application to the DREAM4 in silico network challenge. *Handbook of Computational Statistics: Statistical Bioinformatics*, edited by H. H-S Lu, B. Schölkopf, and H. Zhao, Springer, 461-478.

[20] H. Ma, **H. Zhao** (2013) Application of Bayesian sparse factor analysis models in bioinformatics. *Statistical Bioinformatics,* edited by K.-A. Do, S. Qin, and M. Vannucci, Cambridge University Press, 350-365.

[21] M. Chen, **H. Zhao** (2014) Graphical modeling of biological pathways in genome-wide association studies*. Probabilistic Graphical Models for Genetics, Genomics and Postgenomics*, edited by Christine Sinoquet and Raphaël Mourad, Oxford University Press, 294-317.

[22] M. Chen, L. Hou, **H. Zhao** (2014) Statistical methods for the analysis of next generation sequence data from paired tumor-normal samples. *Statistical Analysis of Next Generation Sequencing Data*, Edited by Datta S, Nettleton D, Springer, 379-404.

[23] N. Sun, **H. Zhao** (2015) Data integration methods in genome wide association studies.*Advanced Medical Statistics* (2nd edition), edited by Y. Lu, L. Tian, J. Fan, H. Jin, World Scientific Publishing Company, 961-976.

[24] C. Li, C. Yang, G. Hather, R. Liu, **H. Zhao** (2015) Data integration and drug discovery: pathway-based approaches. *Integrating Omics Data*, edited by G. C. Tseng, X. J. Zhou and D. Ghosh, Cambridge University Press, 425-444.

[25] C. Yang, C. Li, D. Chung, M. Chen, **H. Zhao** (2016) Introduction to statistical methods in genomewide association studies. *Genome-Wide Association Studies: From Polymorphism to Personalized Medicine*. Edited by Appasani K, Cambridge University Press, 26-52.

[26] M. Chen, C. Yang, C. Li, **H. Zhao** (2016) eQTL mapping. *Genome-Wide Association Studies: From Polymorphism to Personalized Medicine*. Edited by Appasani K, Cambridge University Press, 208-228.

[27] L. Diao, Y. Zhu, N. Sestan, **H. Zhao** (2019) Modeling species specific gene expression across multiple regions in the brain. In press.

# [28] Q. Lu, H. Zhao (2019) Improving genetic association analysis through integration of functional annotations of the human genome. *Handbook of Statistical Genomics*, 4th Edition, 679-696.