BMC Bioinformatics

Meeting report

Proceedings of the Eighth Annual UT-ORNL-KBRIN Bioinformatics Summit 2009

Eric C Rouchka*[1] and Julia Krushkal[2]

Address: [1]Department of Computer Engineering and Computer Science, University of Louisville, 123 JB Speed Building, Louisville, KY 40292, USA and [2]Department of Preventive Medicine and Center of Genomics and Bioinformatics, University of Tennessee Health Science Center, 66 N. Pauline Street, Suite 633, Memphis, TN 38163, USA

Email: Eric C Rouchka* - eric.rouchka@louisville.edu

* Corresponding author

This article is available from: http://www.biomedcentral.com/1471-2105/10/S7/I1

Over the past decade, the University of Tennessee (UT), Oak Ridge National Laboratory (ORNL), and the Kentucky Biomedical Research Infrastructure Network (KBRIN) have collaborated to share their extensive bioinformatics research and educational expertise to further strengthen bioinformatics in the Tennessee and Kentucky region. One of the results of these collaborations is joint sponsorship in an annual regional bioinformatics summit that brings together researchers, educators, and students with interests in bioinformatics from research and educational institutions in Kentucky, Tennessee, and other states. These summits provide unique opportunities for enhancing collaborative links in the region and for further integration of multidisciplinary research efforts across institutions. As a result, a number of new collaborative research and educational projects have been fostered across institutions. The Eighth Annual UT-ORNL-KBRIN Bioinformatics Summit was held at Fall Creek Falls State Park in Pikeville, Tennessee from March 20–22, 2009. A total of 202 participants registered for the summit, with 94 from various Tennessee institutions and 68 from various Kentucky institutions. A number of additional participants came from universities and research institutions from other states and countries, e.g. the National Institutes of Health, Virginia Commonwealth University, University of Cincinnati, Emory University, and the University of British Columbia. Seventy-seven registrants

were faculty, with an additional 62 student, 43 staff, and 20 postdoctoral level participants.

The conference program included three days of presentations. The first day was devoted to workshops, including two Geospiza/Digital World Biology workshops, along with a bioinformatics education and a microarray analysis workshop. The last day and a half were dedicated to scientific sessions in bioinformatics divided into three plenary sessions: Medical and Translational Informatics, Systems Biology, and Next-Generation Sequencing and Epigenetics.

Geospiza/Digital World Biology Workshops

Dr. Sandra Porter, president of Digital World Biology, kicked off the Summit with two workshops from Geospiza and Digital World Biology. The first workshop, "Polymorphism/SNP Discovery" focused on discovering single nucleotide polymorphisms (SNPs) using raw Sanger sequencing trace files [1,2] and the associated Phred [3,4] quality values in regions with low quality scores. These SNPs can be visually represented by using techniques for viewing sequence chromatograms such as Geospiza's FinchTV. The example of the SNP with dbSNP [5] entry rs671 was used as an illustrative case. This SNP results in a single base difference in the nucleotide sequence from an A in the wild type to a G in the mutant,

causing a change from a glutamine to a lysine in amino acid position 487 of the alcohol dehydrogenase (ALDH2) gene. Conformational changes in the ALDH2 protein structure cause an individual ingesting alcohol to lose the capability of efficiently metabolizing acetaldehyde [6]. HapMap [7] information for this SNP indicates that 31% of the Han Chinese population is heterozygous for this SNP. In fact, this allele is not typically found in any population outside of Asia [8]. Dr. Porter discussed using NCBI's Cn3D [9] to view structural locations to form hypotheses as to the possible molecular interactions of amino acids at specific locations and how these interactions can be affected by polymorphisms. This discussion was illustrated with the wild type and mutant for ALDH2 structures with Protein Data Bank (PDB) [10] entries 1O05 and 1ZUM, respectively, where the change from a negatively to a positively charged amino acid causes a change in how the protein subunits of ALDH2 interact.

In the second Geospiza workshop, "Next Generation DNA Sequencing", Dr. Porter provided an overview of three next generation sequencing platforms (454 [11], Illumina [12], and SOLiD [13]) which are technologies enabling for an increased availability of DNA sequence information at a reduced cost per run compared to the traditional Sanger sequencing technique. The properties of each approach in terms of methodologies, data preparation, raw data, analysis, and sequence types interrogated (i.e. genomes, transcriptomes, miRNAs, copy number variants, SNP analysis, epigenetic methylation) were explained. Dr. Porter also discussed the pipeline that Geospiza have in place for dealing with data from sample to results. Use of next generation sequencing data and RNA-Seq [14] for transcriptome analysis as opposed to microarrays [14] is becoming a more real possibility. One of the main advantages of such an approach is that it becomes possible to study all possible isoforms and SNPs, including those not previously described. Studies performed for transcriptome analysis [15,16] were analyzed using Geospiza's GeneSifter™ software.

Bioinformatics Education Workshop
Dr. Steven Jennings of the University of Arkansas-Little Rock led a discussion on the state of bioinformatics education. Building upon his experience in creating a Ph.D. program in Bioinformatics as well as serving on national society level bioinformatics education committees, Dr. Jennings offered several insights into how a student interested in bioinformatics should be trained. The issue of training students often leads to a struggle of breadth versus depth of training. As Dr. Jennings pointed out, many of the techniques students learn will be out-of-date within five years of graduation. Therefore, importance in bioinformatics training should be placed on producing students who are independent thinkers who are able to adapt

to changing technologies. A methodology for constructing a program in bioinformatics was proposed by constructing a cube, where each dimension represents topics in the fields of biology, computer science, and mathematical modeling/computation. The intersection of these areas shows the difficulty in producing a "one size fits all" program.

Statistical Analysis of Microarrays Workshop
Issues involved in analyzing microarray data from a statistical perspective were the topic of the workshop provided by Dr. Arnold Stromberg from the University of Kentucky. Dr. Stromberg has examined a number of research issues with microarrays, including pooling samples [17]. The main focus of this workshop was to encourage researchers to reduce the number of tests and gene lists used in order to increase the p-values and reduce the false discovery rate (FDR). A quadratic regression analysis technique was discussed that allows researchers to classify the behavior of genes into one of nine basic patterns over time [18]. Such an approach can be favorable to cluster analysis by showing the actual behavior of the gene(s) of interest. Dr. Stromberg suggested that the best approach to solving issues with microarrays is to consider the experimental design from the outset, keeping in mind three key questions: 1) What do you want to know? 2) What is the simplest design that will do the job? 3) Can the design be modified to reduce variability?

Medical and Translational Informatics
This year's Medical and Translational Informatics session included a plenary presentation by Bruce Aronow of the University of Cincinnati and Cincinnati Children's on "Integrative Biology and Disease." Dr. Aronow presented his perspective of building upon systems biology techniques for understanding systems dynamics across concepts to allow for a higher level of abstractions. Inclusion of databases of prior knowledge such as molecular, clinical and phenomic sources is key to the understanding of what is going on biologically. For instance, a pathway can be analyzed by first understanding how miRNAs can knock down transcription factor expression, thereby altering gene expression which in turn may affect a particular pathway. A discussion of ontological models for drug and disease correlation was included, which will hopefully lead to better personalized medicine by developing a greater connectivity of knowledge between drug interactions and their effects on genes, gene products, as well as transcriptional and translational control elements. The Systems Biology of Disease and Drug Ontology (SBD) as well as GATACA and ToppGene [19] were discussed as tools that allow for a better understanding of disease.

A second plenary presentation entitled "Slim-Prim: A bioinformatics database bridging basic and clinical science"

was made by Ian Brooks of the University of Tennessee Health Science Center. Slim-Prim is a HIPPA compliant management system for managing information for either scientific laboratories or patient-care research. Slim-Prim was initially developed for use by members of the University of Tennessee Health Science Center's Clinical and Translational Science Institute (CTSI). At its base is an Oracle data/knowledge management system. Built upon this core is a web-based API for building forms for individual projects or patient studies. Each project can then be linked to additional information such as patient history and biorepository information both locally and in a federated fashion. Dr. Brooks discussed two such sources of electronic health records currently housed in Slim-Prim, the Kids' Inpatient Database (KID) [20] which contains 7 million records; and the Mid-South eHealth Alliance (MSeHA) [21] which produces "RHIO" for electronic health records at a rate of 1.5 million records per year. A web-based report generator, Knowledge Informatics for Science and Medical Education and Training (KISMET), allows for access to local and national resources, including caBIG [22], for more complete analysis of the Slim-Prim data. The main benefits of the Slim-Prim system are that it is user-friendly, secure, versatile, and portable.

Systems Biology

The Systems Biology plenary session featured four speakers from Virginia Commonwealth University (VCU). Dr. Michael Miles presented his research on genetic characterization of robust ethanol-responsive gene networks in mouse prefrontal cortex [23-28]. Analysis by his group of QTL mapping of genome-wide expression changes to ethanol in mice response pinpointed multiple genome loci showing strong signals, indicating the role of these loci in gene expression changes. A number of loci were suggested to influence regulation of response to ethanol for gene networks. Epistatic interactions were observed for a number of loci, suggesting the role for DNA modification in regulation of gene expression in response to ethanol.

The presentation on systems vaccinology for *Cryptosporidium*, an important apicomplexan parasite, was made by Dr. Gregory Buck, head of the Center for the Study of Biological Complexity at VCU. He summarized his research, which yielded the genome sequences of *C. hominis* and *C. parvum* [29,30]. He further described the successful identification by his group of promising vaccine targets by employing a joint strategy of comparative analysis of gene expression and proteomics of different stages of the life cycle of *Cryptosporidium*, using genome analysis identifying predicted membrane- or surface-associated proteins, secreted proteins, and other relevant candidates, and by employing a combination of experimental and *in silico* analysis [31-33].

Dr. Zhongming Zhao presented his research on gene networks and pathways in schizophrenia. In his presentation, he discussed his bioinformatic approach to identify candidate genes for schizophrenia by combining results from gene mapping studies including genome-wide association analysis, linkage analysis, gene expression information, and literature search, and by employing screening criteria of connectivity in the human protein-protein interaction networks [34]. He also outlined his research on the role of microRNA interaction networks in schizophrenia and the successful development of an online database for schizophrenia genes.

Dr. Ping Xu presented the final talk at this session, in which he described his integrative study of streptococcal virulence by employing comparative genomics and systems biology. He described the devastating effect of streptococcal infections and summarized a systematic experimental genome-wide deletion analysis of each open reading frame in the *Streptococcus sanguinis* genome, which will lead to better understanding of the phenotypic role of each of these genes [35,36].

Next-Gen Sequencing and Epigenetics

Robert Hanson of NIH/NIDDK was the first presenter in this session with a talk on "Genetic and Epigenetic Studies of Type 2 Diabetes in American Indians." His presentation included a discussion of the complexity of Type 2 diabetes, specifically in understanding the role of potential epigenetic factors. These studies involved looking at the birth weight of babies in addition to familial history in the American Indian population. Genome-wide linkage analysis and association mapping studies indicate potential candidates [37-57]. Some variants show significantly weaker effects in American Indians than in Europeans, indicating the importance of epigenetics in terms of parent-of-origin effects and interaction with the diabetic intrauterine environment.

Jarret Glasscock of Cofactor Genomics followed with a presentation titled "New aspects of bioinformatics introduced by next-generation sequencing technologies." Dr. Glasscock has been involved in early testing and characterization of many of the Next-Gen sequencing platforms, including the llumina, 454, and SOLiD technologies. His presentation covered the possibilities these technologies now provide, including large scale and single nucleotide polymorphism discovery, gene expression quantification, and epigenetic studies through bisulphate sequencing. An overview and contrast of these technologies were given in terms of which types of studies are most suitable for each. Dr. Glasscock's presentation led to an engaging discussion. These exciting technologies are rapidly evolving and lead to many interesting research questions both with the

data generated and in methodologies for handling and annotating the data itself.

Educational Opportunities

Dr. Cynthia Peterson, the director of the UT/ORNL Graduate School of Genome Science and Technology presented an update on the educational opportunities at UT/ORNL. She discussed the progress made with SCALE-IT (scalable computing and leading edge innovative technologies) program over the past year. In addition, she discussed the National Institute for Mathematical & Biological Synthesis (NIMBioS), a one-of-a-kind institute housed at the University of Tennessee. NIMBioS is the result of a $16 million National Science Foundation award to the University of Tennessee, Knoxville that will draw more than 600 national and international researchers each year to participate in working groups, workshops, and conferences. Support for working groups, postdoctoral and sabbatical fellowships, as well as graduate assistantships are all available through NIMBioS http://www.nimbios.org/. PEER, The Program for Excellence & Equity in Research, was also discussed as an avenue to increase the diversity of student populations in the STEM areas through graduate fellowships, scientific training, and career skills workshops.

Poster session

Thirty-five posters were presented on Saturday afternoon during a two-hour poster session. Abstracts (many of which appear in this supplement) were divided into the general groupings of Bioimaging, Bioinformatics Infrastructure, Bioinformatics of Health and Disease, Comparative Genomics, Databases, Functional Genomics, Gene Regulation, Genomics, Machine Learning and Algorithms, Microarrays, Ontologies and Text Mining, Proteomics, Structure and Function Prediction, and Systems Biology.

Five of these abstracts were included in the summit program as short platform presentations. They included "A Framework for Layered Integration of Heterogeneous Data: A Case Study" (Vida Abedi); "Motif Tool Manager: a web-based platform for motif discovery" (Vinhthuy Phan); "The Ontological Discovery Environment: Integrating gene-centered data across diverse experiments " (Jeremy Jay); "Transcriptional Profiling of CD4 T-cells Reveals Abnormal Gene Expression in Young Prediabetic NOD Mice" (Dorothy Kakoola); and "Extensive Parent-of-Origin Genetic Effects on Fetal Growth" (Ron Adkins). A sixth presentation was given by Ramin Homayouni of ComputableGenomix on their GeneIndexer toolkit.

Future plans

The 2010 Bioinformatics Summit will rotate back to Lake Barkley State Park in western Kentucky for the spring of

2010. Areas of interest will likely be on the use of next-generation sequencing technologies in research laboratories, clinical informatics, and integrative systems biology.

Competing interests

The authors declare that they have no competing interests.

References

1. Sanger F, Nicklen S, Coulson AR: **DNA sequencing with chain-terminating inhibitors.** *Proc Natl Acad Sci USA* 1977, **74:**5463-5467.
2. Smith LM, Sanders JZ, Kaiser RJ, Hughes P, Dodd C, Connell CR, *et al.*: **Fluorescence detection in automated DNA sequence analysis.** *Nature* 1986, **321:**674-679.
3. Ewing B, Green P: **Base-calling of automated sequencer traces using phred. II. Error probabilities.** *Genome Res* 1998, **8:**186-194.
4. Ewing B, Hillier L, Wendl MC, Green P: **Base-calling of automated sequencer traces using phred. I. Accuracy assessment.** *Genome Res* 1998, **8:**175-185.
5. Smigielski EM, Sirotkin K, Ward M, Sherry ST: **dbSNP: a database of single nucleotide polymorphisms.** *Nucleic Acids Res* 2000, **28:**352-355.
6. Crabb DW, Edenberg HJ, Bosron WF, Li TK: **Genotypes for aldehyde dehydrogenase deficiency and alcohol sensitivity. The inactive ALDH2(2) allele is dominant.** *J Clin Invest* 1989, **83:**314-316.
7. **The International HapMap Project.** *Nature* 2003, **426:**789-796.
8. Goedde HW, Agarwal DP, Fritze G, Meier-Tackmann D, Singh S, Beckmann G, *et al.*: **Distribution of ADH2 and ALDH2 genotypes in different populations.** *Hum Genet* 1992, **88:**344-346.
9. Hogue CW: **Cn3D: a new generation of three-dimensional molecular structure viewer.** *Trends Biochem Sci* 1997, **22:**314-316.
10. Bernstein FC, Koetzle TF, Williams GJ, Meyer EF Jr, Brice MD, Rodgers JR, *et al.*: **The Protein Data Bank: a computer-based archival file for macromolecular structures.** *J Mol Biol* 1977, **112:**535-542.
11. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, *et al.*: **Genome sequencing in microfabricated high-density picolitre reactors.** *Nature* 2005, **437:**376-380.

12. Turcatti G, Romieu A, Fedurco M, Tairi AP: **A new class of cleavable fluorescent nucleotides: synthesis and optimization as reversible terminators for DNA sequencing by synthesis.** *Nucleic Acids Res* 2008, **36:**e25.

13. Pandey V, Nutter RC, Prediger E: **Applied Biosystems SOLiD™ System: Ligation-Based Sequencing.** In *Next-Generation Genome Sequencing: Towards Personalized Medicine* Edited by: *Janitz M. Hoboken, NJ: Wiley*; 2008:29-41.

14. Wang Z, Gerstein M, Snyder M: **RNA-Seq: a revolutionary tool for transcriptomics.** *Nat Rev Genet* 2009, **10:**57-63.

15. Cloonan N, Forrest AR, Kolle G, Gardiner BB, Faulkner GJ, Brown MK, *et al.*: **Stem cell transcriptome profiling via massive-scale mRNA sequencing.** *Nat Methods* 2008, **5:**613-619.

16. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B: **Mapping and quantifying mammalian transcriptomes by RNA-Seq.** *Nat Methods* 2008, **5:**621-628.

17. Peng X, Wood CL, Blalock EM, Chen KC, Landfield PW, Stromberg AJ: **Statistical implications of pooling RNA samples for microarray experiments.** *BMC Bioinformatics* 2003, **4:**26.

18. Li H, Wood CL, Liu Y, Getchell TV, Getchell ML, Stromberg AJ: **Identification of gene expression patterns using planned linear contrasts.** *BMC Bioinformatics* 2006, **7:**245.

19. Chen J, Xu H, Aronow BJ, Jegga AG: **Improved human disease candidate gene prioritization using mouse phenotype.** *BMC Bioinformatics* 2007, **8:**392.

20. Steiner C, Elixhauser A, Schnaier J: **The healthcare cost and utilization project: an overview.** *Eff Clin Pract* 2002, **5:**143-151.

21. Frisse ME: **State and community-based efforts to foster interoperability.** *Health Aff (Millwood)* 2005, **24:**1190-1196.

22. Kakazu KK, Cheung LW, Lynne W: **The Cancer Biomedical Informatics Grid (caBIG): pioneering an expansive network of information and tools for collaborative cancer research.** *Hawaii Med J* 2004, **63:**273-275.

23. Goldowitz D, Matthews DB, Hamre KM, Mittleman G, Chesler EJ, Becker HC, *et al.*: **Progress in using mouse inbred strains, consomics, and mutants to identify genes related to stress, anxiety, and alcohol phenotypes.** *Alcohol Clin Exp Res* 2006, **30:**1066-1078.

24. Guo AY, Webb BT, Miles MF, Zimmerman MP, Kendler KS, Zhao Z: **ERGR: An ethanol-related gene resource.** *Nucleic Acids Res* 2009, **37:**D840-D845.

25. Kerns RT, Ravindranathan A, Hassan S, Cage MP, York T, Sikela JM, *et al.*: **Ethanol-responsive brain region expression networks: implications for behavioral responses to acute ethanol in DBA/2J versus C57BL/6J mice.** *J Neurosci* 2005, **25:**2255-2266.

26. Kerns RT, Miles MF: **Microarray analysis of ethanol-induced changes in gene expression.** *Methods Mol Biol* 2008, **447:**395-410.

27. Khisti RT, Wolstenholme J, Shelton KL, Miles MF: **Characterization of the ethanol-deprivation effect in substrains of C57BL/6 mice.** *Alcohol* 2006, **40:**119-126.

28. Mulligan MK, Ponomarev I, Boehm SL, Owen JA, Levin PS, Berman AE, *et al.*: **Alcohol trait and transcriptional genomic analysis of C57BL/6 substrains.** *Genes Brain Behav* 2008, **7:**677-689.

29. Abrahamsen MS, Templeton TJ, Enomoto S, Abrahante JE, Zhu G, Lancto CA, *et al.*: **Complete genome sequence of the apicomplexan, Cryptosporidium parvum.** *Science* 2004, **304:**441-445.

30. Xu P, Widmer G, Wang Y, Ozaki LS, Alves JM, Serrano MG, *et al.*: **The genome of Cryptosporidium hominis.** *Nature* 2004, **431:**1107-1112.

31. Mercado R, Buck GA, Manque PA, Ozaki LS: **Cryptosporidium hominis infection of the human respiratory tract.** *Emerg Infect Dis* 2007, **13:**462-464.

32. Puiu D, Enomoto S, Buck GA, Abrahamsen MS, Kissinger JC: **CryptoDB: the Cryptosporidium genome resource.** *Nucleic Acids Res* 2004, **32:**D329-D331.

33. Widmer G, Akiyoshi D, Buckholt MA, Feng X, Rich SM, Deary KM, *et al.*: **Animal propagation and genomic survey of a genotype 1 isolate of Cryptosporidium parvum.** *Mol Biochem Parasitol* 2000, **108:**187-197.

34. Sun J, Kuo PH, Riley BP, Kendler KS, Zhao Z: **Candidate genes for schizophrenia: a survey of association studies and gene ranking.** *Am J Med Genet B Neuropsychiatr Genet* 2008, **147B:**1173-1181.

35. Ge X, Kitten T, Chen Z, Lee SP, Munro CL, Xu P: **Identification of Streptococcus sanguinis genes required for biofilm formation and examination of their role in endocarditis virulence.** *Infect Immun* 2008, **76:**2551-2559.

36. Xu P, Alves JM, Kitten T, Brown A, Chen Z, Ozaki LS, *et al.*: **Genome of the opportunistic pathogen Streptococcus sanguinis.** *J Bacteriol* 2007, **189:**3166-3175.

37. Rong R, Hanson RL, Ortiz D, Wiedrich C, Kobes S, Knowler WC, *et al.*: **Association analysis of variation in/near FTO, CDKAL1, SLC30A8, HHEX, EXT2, IGF2BP2, LOC387761, and CDKN2B with type 2 diabetes and related quantitative traits in Pima Indians.** *Diabetes* 2009, **58:**478-488.

38. Pavkov ME, Knowler WC, Hanson RL, Nelson RG: **Diabetic nephropathy in American Indians, with a special emphasis on the Pima Indians.** *Curr Diab Rep* 2008, **8:**486-493.

39. Nelson RG, Pavkov ME, Hanson RL, Knowler WC: **Changing course of diabetic nephropathy in the Pima Indians.** *Diabetes Res Clin Pract* 2008, **82(Suppl 1):**S10-S14.

40. Ma L, Hanson RL, Que LN, Guo Y, Kobes S, Bogardus C, *et al.*: **PCLO variants are nominally associated with early-onset type 2 diabetes and insulin resistance in Pima Indians.** *Diabetes* 2008, **57:**3156-3160.

41. Ma L, Hanson RL, Que LN, Mack JL, Franks PW, Infante AM, *et al.*: **Association analysis of Kruppel-like factor 11 variants with type 2 diabetes in Pima Indians.** *J Clin Endocrinol Metab* 2008, **93:**3644-3649.

42. Guan W, Pluzhnikov A, Cox NJ, Boehnke M: **Meta-analysis of 23 type 2 diabetes linkage studies from the International Type 2 Diabetes Linkage Analysis Consortium.** *Hum Hered* 2008, **66:**35-49.

43. Guo T, Hanson RL, Traurig M, Muller YL, Ma L, Mack J, *et al.*: **TCF7L2 is not a major susceptibility gene for type 2 diabetes in Pima Indians: analysis of 3,501 individuals.** *Diabetes* 2007, **56:**3082-3088.

44. Hanson RL, Bogardus C, Duggan D, Kobes S, Knowlton M, Infante AM, *et al.*: **A search for variants associated with young-onset type 2 diabetes in American Indians in a 100K genotyping array.** *Diabetes* 2007, **56:**3045-3052.

45. Pavkov ME, Hanson RL, Knowler WC, Bennett PH, Krakoff J, Nelson RG: **Changing patterns of type 2 diabetes incidence among Pima Indians.** *Diabetes Care* 2007, **30:**1758-1763.

46. Wolford JK, Yeatts KA, Red Eagle AR, Nelson RG, Knowler WC, Hanson RL: **Variants in the gene encoding aldose reductase (AKR1B1) and diabetic nephropathy in American Indians.** *Diabet Med* 2006, **23:**367-376.

47. Franks PW, Looker HC, Kobes S, Touger L, Tataranni PA, Hanson RL, *et al.*: **Gestational glucose tolerance and risk of type 2 diabetes in young Pima Indian offspring.** *Diabetes* 2006, **55:**460-465.

48. Muller YL, Infante AM, Hanson RL, Love-Gregory L, Knowler W, Bogardus C, *et al.*: **Variants in hepatocyte nuclear factor 4alpha are modestly associated with type 2 diabetes in Pima Indians.** *Diabetes* 2005, **54:**3035-3039.

49. Kovacs P, Ma L, Hanson RL, Franks P, Stumvoll M, Bogardus C, *et al.*: **Genetic variation in UCP2 (uncoupling protein-2) is associated with energy metabolism in Pima Indians.** *Diabetologia* 2005, **48:**2292-2295.

50. Red Eagle AR, Hanson RL, Jiang W, Han X, Matters GL, Imperatore G, *et al.*: **Meprin beta metalloprotease gene polymorphisms associated with diabetic nephropathy in the Pima Indians.** *Hum Genet* 2005, **118:**12-22.

51. Vozarova de Court , Hanson RL, Funahashi T, Lindsay RS, Matsuzawa Y, Tanaka S, *et al.*: **Common Polymorphisms in the Adiponectin Gene ACDC Are Not Associated With Diabetes in Pima Indians.** *Diabetes* 2005, **54:**284-289.

52. Baier LJ, Hanson RL: **Genetic studies of the etiology of type 2 diabetes in Pima Indians: hunting for pieces to a complicated puzzle.** *Diabetes* 2004, **53:**1181-1186.

53. Kovacs P, Hanson RL, Lee YH, Yang X, Kobes S, Permana PA, *et al.*: **The role of insulin receptor substrate-1 gene (IRS1) in type 2 diabetes in Pima Indians.** *Diabetes* 2003, **52:**3005-3009.

54. Farook VS, Hanson RL, Wolford JK, Bogardus C, Prochazka M: **Molecular analysis of KCNJ10 on 1q as a candidate gene for Type 2 diabetes in Pima Indians.** *Diabetes* 2002, **51:**3342-3346.

55. Lindsay RS, Kobes S, Knowler WC, Hanson RL: **Genome-wide linkage analysis assessing parent-of-origin effects in the inheritance of birth weight.** *Hum Genet* 2002, **110:**503-509.

56. Lindsay RS, Cook V, Hanson RL, Salbe AD, Tataranni A, Knowler WC: **Early excess weight gain of children in the Pima Indian population.** *Pediatrics* 2002, **109:**E33.

57. Lindsay RS, Kobes S, Knowler WC, Bennett PH, Hanson RL: **Genome-wide linkage analysis assessing parent-of-origin effects in the inheritance of type 2 diabetes and BMI in Pima Indians.** *Diabetes* 2001, **50:**2850-2857.