William Valdar

Associate Professor, Department of Genetics, University of North Carolina at Chapel Hill
Work Address: 120 Mason Farm Road, Suite 5113, Chapel Hill, NC 27599-7264
Tel: +1 919 843 2833 · Email: william.valdar@unc.edu · Web: http://valdarlab.unc.edu

EDUCATION

2002-2003	Oxford University, UK	MS Applied Statistics – Distinction	Awarded: 2004-05-08
1997-2001	University College London, UK	PhD Protein Structure Bioinformatics	Awarded: 2001-11-30
1996-1997	Manchester University, UK	MS Bioinformatics	Awarded: 1998-06-30
1993-1996	Manchester University, UK	BS (Hons) Biochemistry	Awarded: 1996-06-27

PROFESSIONAL EXPERIENCE

2015-present	**Associate Professor, Department of Genetics, UNC-Chapel Hill**
2015-present	Adjunct Associate Professor, Department of Biostatistics, UNC-Chapel Hill
2015-present	Adjunct Associate Professor, Department of Computer Science, UNC-Chapel Hill
2013-present	Associate Director, Graduate Program in Bioinformatics and Computational Biology, UNC-Chapel Hill
2012-present	Adjunct Assistant Professor, Department of Computer Science, UNC-Chapel Hill
2010-present	Member, Center for Integrated Systems Genetics, UNC-Chapel Hill
2010-present	Member, Nutrition Obesity Research Center, UNC-Chapel Hill
2009-present	Member, Program in Bioinformatics and Computational Biology, UNC-Chapel Hill
2009-present	Member, Lineberger Comprehensive Cancer Center UNC-Chapel Hill
2010-2015	Adjunct Assistant Professor, Department of Biostatistics, UNC-Chapel Hill
2009-2015	Assistant Professor, Department of Genetics, UNC-Chapel Hill
2008-2009	MRC Career Development Fellow in Biostatistics, Wellcome Trust Centre for Human Genetics, University of Oxford, UK. Sponsor: Peter Donnelly, PhD FRS.
2001-2008	Post-doctoral research scientist, bioinformatics and statistical genetics at the Wellcome Trust Centre for Human Genetics, University of Oxford, UK. Supervisors: Richard Mott, PhD, Jonathan Flint, MD.
1997-2001	Doctoral student in protein structure bioinformatics. Biochemistry Dept, University College London, UK. Thesis: Residue conservation at protein-protein interfaces. Supervisor: Janet M Thornton, PhD FRS CBE.

HONORS AND AWARDS

12/2011- 12/2012	IBM Junior Faculty Development Award, University of North Carolina
7/2008 – 7/2009	Career Development Fellowship in Biostatistics, The Medical Research Council, UK
1/2005 – 5/2005	Access to Research Infrastructures Fellowship, European Commission, Sweden

BIBLIOGRAPHY: BOOK CHAPTERS (1)

1. **Valdar WSJ**, Jones DT (**2003**) Amino acid residue conservation. In *Bioinformatics: Genes, Proteins and Computers.* BIOS Scientific Publishers Ltd, Oxford, 48-63.

BIBLIOGRAPHY: PEER REVIEWED ARTICLES (54)

* indicates equal contribution authorship; **bold** = W Valdar; underline = under direct supervision of W Valdar

ORIGINAL RESEARCH ARTICLES (50)

1. Xie Y, Liu Y, **Valdar W** (**2016**) Joint estimation of multiple dependent Gaussian graphical models with applications to mouse genetics. *Biometrika* 103(3):493-511.

   This theoretical work along with associated analysis was performed by my student Yuying Xie under the joint supervision of me and his co-advisor Dr Liu. In particular, I contributed to the concept, the development of the model, the design of all simulations, the analysis of real data, and I extensively edited the manuscript.

2. Gralinski LE, Ferris MT, Aylor DL, Whitmore AC, Green R, Frieman M, Deming D, Menachery VD, Miller DR, Buus RJ, Bell TA, Churchill GA, Threadgill DW, Katze MG, McMillan L, **Valdar W**, Heise MT, Baric RS (**2015**) Genome wide identification of SARS-CoV susceptibility loci using the Collaborative Cross. *PLoS Genetics.* 11(10):e1005504.

3. Li J, Lange LA, Sabourin J, Duan Q, **Valdar W**, Willis MS, Li Y, Wilson JG, Lange EM (**2015**) Genome-wide and Exome-wide Association Study of Serum Lipoprotein (a) in the Jackson Heart Study. *Journal of Human Genetics* 60(12):755-61.

4. Durda P, Sabourin J, Lange EM, Nalls MA, Mychaleckyj JC, Jenny NS, Li J, Walston J, Harris TB, Psaty BM, **Valdar W**, Liu Y, Cushman M, Reiner AP, Tracy* RP, Lange* LA (**2015**) Plasma levels of sIL-2Rα: associations with clinical cardiovascular events and genome-wide association scan. Accepted in *Arteriosclerosis, Thrombosis, and Vascular Biology* 35(10):2246-53.

5. Sabourin J, **Valdar W**, Nobel AB (**2015**) A permutation approach for selecting the penalty parameter in penalized model selection. *Biometrics* 71(4):1185-94.

6. Crowley JJ, Zhabotynsky V, Sun W, Huang S, Pakatci IK, Kim Y, Wang JR, Morgan AP, Calaway JD, Aylor DL, Yun Z, Bell TA, Buus RJ, Calaway ME, Didion JP, Gooch TJ, Hansen SD, Robinson NN, Shaw GD, Spence JS, Quackenbush CR, Barrick CJ, Nonneman RaJ, Kim K, Xenakis J, Xie Y, **Valdar W**, Lenarcic AB, Wang W, Welsh CE, Fu C, Zhang Z, Holt J, Guo Z, Threadgill DW, Tarantino LM, Miller DR, Zou F, McMillan L, Sullivan PF, Pardo Manuel de Villena F (**2015**) Analyses of allele-specific gene expression in highly divergent mouse crosses identifies pervasive allelic imbalance. *Nature Genetics* 47(4):353-60.

7. Sabourin J, Nobel AB, **Valdar W** (**2015**) Fine-mapping additive and dominant SNP effects using group-LASSO and Fractional Resample Model Averaging. *Genetic Epidemiology* 39(2):77-88.

   I designed and supervised this study, performed by my student (and latterly my postdoc) Dr Sabourin; I supported the work and with my student co-wrote the manuscript.

8. Gatti DM, Svenson KL, Shabalin A, Wu L, **Valdar W**, Simecek P, Goodwin N, Cheng R, Pomp D, Palmer A, Chesler EJ, Broman KW, Churchill GA (**2014**) Quantitative trait locus mapping methods for Diversity Outbred mice. *G3: Genes, Genomes, Genetics* 4(9):1623-1633.

9. Zhang Z, Wang W, **Valdar W** (**2014**) Bayesian modeling of haplotype effects in multiparent populations. *Genetics* 198:139-156.

   I designed and supervised this study performed by my student, co-supported the work, and co-wrote and extensively edited/rewrote the manuscript.

9. Matson BC, Corty RW, Karpinich NO, Murtha AP, **Valdar W**, Grotegut CA, Caron KM. (**2014**) Midregional pro-adrenomedullin plasma concentrations are blunted in severe preeclampsia. *Placenta* 35(9):780-3.

10. Crowley JJ*, Kim Y*, Lenarcic AB*, Quackenbush CR, Barrick CJ, Adkins DE, Shaw GS, Miller DR, Pardo-Manuel de Villena F, Sullivan PF, **Valdar W (2014)** Genetics of adverse reactions to haloperidol in a mouse diallel: A drug-placebo experiment and Bayesian causal analysis. *Genetics* 196(1):321-47.

    This collaborative study comprised a drug-placebo mouse experiment (wet lab), development of a novel and extensive statistical approach for its analysis (methods development), and application of that statistical approach to the mouse data (data analysis). I led the methods development and data analysis portion of this project, which due to the complex and unprecedented nature of the experiment included defining an entire theoretical framework for how to report the results, as well as: supervising my postdoc Dr Lenarcic in extensive methods development, resulting in 11 of the final 27 double-column journal pages being devoted to statistical theory and methods; supervising the data analysis performed by Dr Kim (Res. Assist. Prof. in the Sullivan lab); interpreting the data, along with Drs Crowley (Res. Assist. Prof. in the Sullivan lab), Kim and Lenarcic. I co-wrote and extensively re-wrote this complex manuscript through multiple drafts.

    This article has since received attention in several medical online news outlets ("The Genetics of Drug Tolerance", in Healthcanal, MedicalXpress, Newswise, Futurity). By the journal's own metrics of "attention" it was, as of Oct 29, 2014, ranked #24 of 1,487 articles in Genetics.

11. Leamy LJ, Elo K, Nielsen MK, Thorn SR, **Valdar W**, Pomp D (**2014**) Quantitative trait loci for energy balance traits in an advanced intercross line derived from mice divergently selected for heat loss. *PeerJ* 2:e392

12. Phillippi J*, Xie Y*, Miller DR, Bell TA, Zhang Z, Lenarcic AB, Aylor DL, Krovi SH, Threadgill DW, Pardo-Manuel de Villena F, Wang W, **Valdar W***, Frelinger JA* (**2014**) Using the Collaborative Cross to probe the immune system. *Genes and Immunity* 15(1):38-46.

    This gene mapping study was a collaboration between an experimental group, led by immunologist Dr Frelinger, and a computational group, led by me. Specifically, I led the computational/statistical genetic analysis portion of the study, supervising my students Mr Xie and Mr Zhang, and postdoc Dr Lenarcic, to perform statistical genetic analysis of the data. I co-wrote the article with Dr Frelinger and his student Mr Phillippi, and I am co-corresponding author.

13. Calaway J, Lenarcic AB, Didion JP, Wang JR, Searle JB, McMillan L, **Valdar W**, Pardo-Manuel de Villena F **(2013)** Genetic architecture of skewed X inactivation in the laboratory mouse. *PLoS Genetics* 9(10):e1003853.

This project was led by Dr Pardo-Manuel de Villena, but a key component of it – quantifying X-inactivation skew – formed the basis of an equal collaboration between his group and mine. Specifically, a crucial part of the work involved quantifying statistical uncertainty about X-inactivation proportions from pyrosequencing data, a problem for which no existing statistical technique was readily applicable. Dr Pardo-Manuel de Villena and I formed a collaboration to develop and apply such a technique; in that technique's application, my group (postdoc, Dr Lenarcic) guided decisions about subsequent experimental design, and the technique itself is currently under review at a statistical journal (see Lenarcic et al, under review). In this collaboration I supervised and supported my postdoc, contributed to the interpretation of the analysis, and wrote and edited sections of the paper.

14. Ferris MT, Aylor DL, Bottomly D, Whitmore AC, Aicher LD, Bell TA, Bradel-Tretheway B, Bryan JT, Buus RJ, Gralinski LE, Haagmans BL, McMillan L, Miller DR, Rosenzweig E, **Valdar W**, Wang J, Churchill GA, Threadgill DW, McWeeney SK, Katze MG, Pardo-Manuel de Villena F, Baric RS, Heise MT (**2013**) Modeling host genetic regulation of influenza pathogenesis in the Collaborative Cross. *PLoS Pathogens* 9(2):e1003196.

15. Solberg Woods LC, Holl KL, Oreper D, Xie Y, Tsaih S-W, **Valdar W** (**2012**) Fine-mapping diabetes-related traits, including insulin resistance, in heterogeneous stock rats. *Physiological Genomics* 44(21):1013-26.

    This article arose from an ongoing wet-lab/dry-lab collaboration between my group and the experimental group of Dr Solberg Woods. I led the statistical genetic analysis team, supervising my students Mr Xie and Mr Oreper to analyze the data, including development and testing of new statistical techniques in order to satisfy reviewer comments. I co-wrote the paper. The article was highlighted as an editor's pick for that month.

16. **Valdar W***, Sabourin J*, Nobel A, Holmes C (**2012**) Reprioritizing genetic associations in hit regions using LASSO-based resample model averaging. *Genetic Epidemiology* 36(5):451-462.

    The initial idea for this statistical method was co-developed by myself and collaborator Dr Holmes while I was an independent researcher at Oxford (the recipient of a "K"-like award from UK). When I moved to the US this collaboration was retained, but the project itself was led by me. Specifically, I supervised and supported my student Mr Sabourin throughout development of the method and simulations; I co-wrote the paper with my student and I am corresponding author.

17. Lenarcic AB, Svenson KL, Churchill GA, **Valdar W** (**2012**) A general Bayesian approach to analyzing diallel crosses of inbred strains. *Genetics* 190(2):413-435.

    I formulated the concept for this paper, which used unpublished data from Drs Svenson and Churchill; I developed initial models, supervised my postdoc in developing all methods, simulations and data analyses, and wrote most of the paper and extensively edited all of it.

18. Zhang W, Korstanje R, Thaisz J, Staedtler F, Harttman N, Xu L, Feng M, Yanas L, Yang H, **Valdar W**, Churchill GA, DiPetrillo K (**2012**) Genomewide association mapping of quantitative traits in outbred mice. *G3: Genes, Genomes, Genetics* 2(2):167-174.

    My role in this paper was specific: I supervised Dr Zhang (at the Jackson Laboratory) for a component part of the statistical genetic analysis, namely the detection of multiple QTL using resampling techniques that I devised. I edited the manuscript.

19. Svenson KL, Gatti DM, **Valdar W**, Welsh CE, Cheng R, Chesler EJ, Palmer AA, McMillan L, Churchill GA (**2012**) High resolution genetic mapping with the mouse Diversity Outbred population. *Genetics* 190(2):437-447.

    My role in this paper was specific: I developed and executed a set of computer simulations that form a crucial component of the of the research, and I wrote the corresponding text describing this component.

20. Collaborative Cross Consortium: (Iraqi FA, Mahajne M, Salaymah Y, Sandovski H, Tayem H, Vered K, Balmer L, Hall M, Manship G, Morahan G, Pettit K, Scholten J, Tweedie K, Wallace A, Weerasekera L, Cleak J, Durrant C, Goodstadt L, Mott R, Yalcin B, Hill C, Aylor DL, Baric RS, Bell TA, Bendt KM, Brennan J, Brooks JD, Buus RJ, Crowley JJ, Calaway JD, Calaway ME, Cholka A, Darr DB, Didion JP, Dorman A, Everett ET, Ferris MT, Mathes WF, Fu CP, Gooch TJ, Goodson SG, Gralinski LE, Hansen SD, Heise MT, Hoel J, Hua K, Kapita MC, Lee S, Lenarcic AB, Liu EY, Liu H, McMillan L, Magnuson TR, Manly KF, Miller DR, O'Brien DA, Odet F, Pakatci IK, Pan W, de Villena FP, Perou CM, Pomp D, Quackenbush CR, Robinson NN, Sharpless NE, Shaw GD, Spence JS, Sullivan PF, Sun W, Tarantino LM, **Valdar W**, Wang J, Wang W, Welsh CE, Whitmore A, Wiltshire T, Wright FA, Xie Y, Yun Z, Zhabotynsky V, Zhang Z, Zou F, Powell C, Steigerwalt J, Threadgill DW, Chesler EJ, Gary A.) (**2012**) The genome architecture of the Collaborative Cross mouse genetic reference population. *Genetics* 190(2):389-401.

21. Ahlqvist E, Ekman D, Lindvall T, Popovic M, Förster M, Hultqvist M, Klaczkowska D, Teneva I, Johannesson M, Flint J, **Valdar W**, Nandakumar KS, Holmdahl R (**2011**) High resolution mapping of a complex disease, a model for rheumatoid arthritis, using heterogeneous stock mice. *Human Molecular Genetics* 20(15):3031-3041.

22. Rönnegård L, **Valdar W** (**2011**) Detecting major genetic loci controlling phenotypic variability in experimental crosses. *Genetics* 188(2):435-447.

Dr Ronnegard is a collaborator in Sweden. This article proposes a new statistical method for finding genes affecting trait variability, coins the term "vQTL" (variability-controlling QTL) and demonstrates the method in simulations accompanied by theory. We worked together on the theory, simulations, and co-wrote the paper.

23. Johnsen AK*, **Valdar W***, Golden L, Ortiz-Lopez A, Hitzemann R, Flint J, Mathis D, Benoist C (**2011**) Genetics of arthritis severity: a genome-wide and species-wide dissection in HS mice. *Arthritis & Rheumatism* 63(9):2630-2640.

    This study began as a wet lab-dry lab collaboration between the laboratory of Dr Benoist at Harvard, whose postdoc Dr Johnsen performed experiments, and my former mentor Dr Flint at Oxford, with me performing the extensive statistical analysis. I led most of (and performed all of) the computational and statistical analyses while at UNC, and co-wrote the paper with Dr Benoist and Dr Johnsen.

24. Aylor DL, **Valdar W**, Foulds-Mathes W, Buus RJ, Verdugo RA, Baric RS, Ferris MT, Frelinger JA, Heise M, Frieman MB, Gralinski LE, Bell TA, Didion JP, Hua K, Nehrenberg DL, Powell CL, Steigerwalt J, Xie Y, Kelada SNP, Collins FS, Yang IV, Schwartz DA, Branstetter LA, Chesler EJ, Miller DR, Spence J, Liu EY, McMillan L, Sarkar A, Wang J, Wang W, Zhang Q, Broman KW, Korstanje R, Durrant C, Mott R, Iraqi FA, Pomp D, Threadgill D, Pardo-Manuel de Villena F, Churchill GA (**2011**) Genetic Analysis of Complex Traits in the Emerging Collaborative Cross. *Genome Research* 21(8):1213-1222.

25. Solberg Woods L, Holl K, Tschannen M, **Valdar W** (**2010**) Fine-mapping a locus for glucose tolerance using heterogeneous stock rats. *Physiological Genomics* 41(1):102-8.

    This article arose from an ongoing wet-lab/dry-lab collaboration between my group and the experimental group of Dr Solberg Woods. I led and wholly performed the computational/statistical genetic analysis, and I co-wrote the paper.

26. **Valdar W**, Holmes C, Mott R, Flint J (**2009**) Mapping in structured populations by resample model averaging. *Genetics* 182(4):1263-77.

27. Kover PX, **Valdar W**, Trakalo J, Scarcelli N, Ehrenreich IM, Purugganan MD, Durrant C, Mott R. (**2009**) A Multiparent Advanced Generation Inter-Cross to fine-map quantitative traits in Arabidopsis thaliana. *PLoS Genet.* 5(7).

28. López-Aumatell R, Vicens-Costa E, Guitart-Masip M, Martínez-Membrives E, **Valdar W**, Johannesson M, Cañete T, Blázquez G, Driscoll P, Flint J, Tobeña A, Fernández-Teruel A (**2009**) Unlearned anxiety predicts learned fear: a comparison among heterogeneous rats and the Roman rat strains. *Behav Brain Res*. 202(1):92-101.

29. Huang GJ, Shifman S, **Valdar W**, Johannesson M, Yalcin B, Taylor MS, Taylor JM, Mott R, Flint J (**2009**) High resolution mapping of expression QTLs in heterogeneous stock mice in multiple tissues. *Genome Res* 19(6):1133-40.

30. Ramagopalan SV, **Valdar W**, Dyment DA, DeLuca GC, Yee IM, Giovannoni G, Ebers GC, Sadovnick AD; Canadian Collaborative Study Group (**2009**) Association of infectious mononucleosis with multiple sclerosis. a population-based study. *Neuroepidemiology*. 32(4):257-62.

31. Ramagopalan SV, **Valdar W**, Criscuoli M, DeLuca GC, Dyment DA, Orton SM, Yee IM, Ebers GC, Sadovnick AD; Canadian Collaborative Study Group (**2009**) Age of puberty and the risk of multiple sclerosis: a population based study. *Eur J Neurol* 16(3):342-7.

32. Bice P, **Valdar W**, Zhang L, Liu L, Lai D, Grahame N, Flint J, Li TK, Lumeng L, Foroud T (**2009**) Genomewide SNP Screen to Detect Quantitative Trait Loci for Alcohol Preference in the High Alcohol Preferring and Low Alcohol Preferring Mice. *Alcohol Clin Exp Res* 33(3):531-7.

33. Johannesson M, Lopez-Aumatell R, Stridh P, Diez M, Tuncel J, Blázquez G, Martinez-Membrives E, Cañete T, Vicens-Costa E, Graham D, Copley RR, Hernandez-Pliego P, Beyeen AD, Ockinger J, Fernández-Santamaría C, Gulko PS, Brenner M, Tobeña A, Guitart-Masip M, Giménez-Llort L, Dominiczak A, Holmdahl R, Gauguier D, Olsson T, Mott R, **Valdar W**, Redei EE, Fernández-Teruel A, Flint J. (**2009**) A resource for the simultaneous high-resolution mapping of multiple quantitative trait loci in rats: the NIH heterogeneous stock *Genome Res* 19(1):150-8.

34. Ramagopalan SV, Herrera BM, **Valdar W**, Dyment DA, Orton SM, Yee IM, Criscuoli M, Atkins K, Ebers GC, Sadovnick AD (**2008**) No Effect of Birth Weight on the Risk of Multiple Sclerosis. A Population-Based Study. *Neuroepidemiology* 31(3):181-184.

35. Ramagopalan SV, **Valdar W**, Dyment DA, DeLuca GC, Orton SM, Yee IM, Criscuoli M, Ebers GC, Sadovnick AD; Canadian Collaborative Study Group (**2008**) No effect of preterm birth on the risk of multiple sclerosis: a population based study. *BMC Neurol* 8:30.

36. Lopez-Aumatell R, Guitart-Masip M, Vicens-Costa E, Gimenez-Llort L, **Valdar W**, Johannesson M, Flint J, Tobena A, Fernandez-Teruel (**2008**) Fearfulness in a large N/Nih genetically heterogeneous rat stock: differential profiles of timidity and defensive flight in males and females. *Behavioural Brain Research* 188(1):41-55.

37. Ramagopalan SV, Dyment DA, **Valdar W**, Herrera BM, Criscuoli M, Yee IM, Sadovnick AD, Ebers GC; Canadian Collaborative Study Group (**2007**) Autoimmune disease in families with multiple sclerosis: a population-based study. *Lancet Neurol* 6(7):604-10.

38. Taylor M, **Valdar W**, Kumar A, Flint J, Mott R (**2007**) Management, presentation and interpretation of genome scans using GSCANDB. *Bioinformatics* 15;23(12):1545-9.

39. Keays DA, Tian G, Poirier K, Huang GJ, Siebold C, Cleak J, Oliver PL, Fray M, Harvey RJ, Molnar Z, Pinon MC, Dear N, **Valdar W**, Brown SD, Davies KE, Rawlins JN, Cowan NJ, Nolan P, Chelly J, Flint J (**2007**) Mutations in alpha-tubulin cause abnormal neuronal migration in mice and lissencephaly in humans. *Cell* 128(1):45-57.

40. Keays DA, Clark TG, Campbell TG, Broxholme J, **Valdar W** (**2007**) Estimating the number of coding mutations in genotypic and phenotypic driven N-ethyl-N-nitrosourea (ENU) screens: revisited. *Mammalian Genome* 18(2):123-4.

41. **Valdar W**, Solberg LC, Gaugier D, Cookson WO, Rawlins JNP, Mott R, Flint J (**2006**) Genetic and environmental effects on complex traits in mice. *Genetics* 174(2):959-84.

42. Orton S, Herrera BM, Yee IM, **Valdar W**, Sadovnick AD, Ebers GC (**2006**) Increasing sex ratio of multiple sclerosis in Canada. *Lancet Neurology* 5(11):932-6.

43. **Valdar W**, Solberg LC, Gaugier D, Burnett S, Klenerman P, Cookson WO, Taylor M, Rawlins JNP, Mott R, Flint J (**2006**) Genome-wide genetic association of complex traits in outbred mice. *Nature Genetics* 38(8):879-87.

44. Solberg LC, **Valdar W**, Gauguier D, Nunez G, Taylor A, Hernandez P, Davidson S, Burns P, Cookson W, Deacon R, Rawlins JNP, Mott R, Flint J (**2006**) A protocol for high throughput phenotyping, suitable for quantitative trait analysis in mice. *Mammalian Genome* 17(2):129-46.

45. **Valdar W,** Flint J, Mott R (**2006**) Simulating the collaborative cross: power of QTL detection and mapping resolution in large sets of recombinant inbred strains of mice. *Genetics* 172(3):1783-97.

46. Flint J, **Valdar W**, Shifman S, Mott R (**2005**) Experimental strategies for mapping and cloning quantitative trait genes in rodents. *Nature Reviews Genetics* 6(4):271-286.

47. **Valdar WSJ**, Flint J, Mott R (**2003**) QTL fine-mapping with recombinant-inbred heterogeneous stocks and in-vitro heterogeneous stocks. *Mammalian Genome* 14(12):830-838.

48. Nobeli I, Laskowski RA, **Valdar WS**, Thornton JM (**2001**) On the molecular discrimination between adenine and guanine in proteins. *Nucleic Acids Research*. 29(21): 4294-4309.

49. **Valdar WS**, Thornton JM (**2001**) Conservation helps to identify biologically relevant crystal contacts. *Journal of Molecular Biology*. 313(2): 399-416.

50. **Valdar WS**, Thornton JM (**2001**) Protein-protein interfaces: Analysis of amino acid conservation in homodimers. *Proteins: Structure, Function, and Genetics* 42(1):108-124.

OTHER PEER-REVIEWED ARTICLES (4)

1. Rönnegård L, **Valdar W** (**2012**) Recent developments in statistical methods for detecting genetic loci affecting phenotypic variability. *BMC Genetics* 13(1):63. (7 pages)

   This article is a "correspondence", ie, a combination of a review with original content, and quickly earned the distinction of being "highly accessed" in BMC Genetics. I co-wrote this article.

2. Willis-Owen SAG, **Valdar W** (**2009**) Deciphering gene-environment interactions through mouse models of allergic asthma. *J Allergy Clin Immunol.* 123(1):14-23.

3. Churchill GA, Airey DC, Allayee H, Angel JM, Attie AD, Beatty J, Beavis WD, Belknap JK, Bennett B, Berrettini W, Bleich A, Bogue M, Broman KW, Buck KJ, Buckler E, Burmeister M, Chesler EJ, Cheverud JM, Clapcote S, Cook MN, Cox RD, Crabbe JC, Crusio WE, Darvasi A, Deschepper CF, Doerge RW, Farber CR, Forejt J, Gaile D, Garlow SJ, Geiger H, Gershenfeld H, Gordon T, Gu J, Gu W, de Haan G, Hayes NL, Heller C, Himmelbauer H, Hitzemann R, Hunter K, Hsu HC, Iraqi FA, Ivandic B, Jacob HJ, Jansen RC, Jepsen KJ, Johnson DK, Johnson TE, Kempermann G, Kendziorski C, Kotb M, Kooy RF, Llamas B, Lammert F, Lassalle JM, Lowenstein PR, Lu L, Lusis A, Manly KF, Marcucio R, Matthews D, Medrano JF, Miller DR, Mittleman G, Mock BA, Mogil JS, Montagutelli X, Morahan G, Morris DG, Mott R, Nadeau JH, Nagase H, Nowakowski RS, O'Hara BF, Osadchuk AV, Page GP, Paigen B, Paigen K, Palmer AA, Pan HJ, Peltonen-Palotie L, Peirce J, Pomp D, Pravenec M, Prows DR, Qi Z, Reeves RH, Roder J, Rosen GD, Schadt EE, Schalkwyk LC, Seltzer Z, Shimomura K, Shou S, Sillanpaa MJ, Siracusa LD, Snoeck HW, Spearow JL, Svenson K, Tarantino LM, Threadgill D, Toth LA, **Valdar W**, de Villena FP, Warden C, Whatley S, Williams RW, Wiltshire T, Yi N, Zhang D, Zhang M, Zou F; Complex Trait Consortium (**2004**) The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nature Genetics* 36(11):1133-1137.

4. **Valdar WSJ (2002)** Scoring residue conservation *Proteins: Structure, Function, and Genetics.* 43(2): 227-241.

TEACHING RECORD: COURSES AND LECTURES

COURSES:

BCB 720 "Introduction to Statistical Modeling"
Type: 3-credit, full semester course of 75min lectures to graduate students. For graduate students in the BCB curriculum this course is compulsory. (Prior to 2015, this was 2-credit, 2/3 semester.)
Role: **Course director**, course creator, instructor; co-instructed with Ethan Lange.
Years:

- Fall 2016: 20/29 lectures, 13/15 homeworks, 25 contact hours, 20 enrolled students.
- Fall 2015: 18/30 lectures, 10/14 homeworks, 22.75 contact hours, 21 enrolled students.
- Fall 2014: 15/22 lectures, 7/9 homeworks, 18.75 contact hours, 13 enrolled students.
- Fall 2013: 14/21 lectures, 8/10 homeworks, 17.5 contact hours, 10 enrolled students.
- Fall 2012: 14/20 lectures, 7/9 homeworks, 17.5 contact hours, 14 enrolled students.
- Fall 2011: 11/19 lectures, 7/9 homeworks, 13.75 contact hours, 9 enrolled students.

GNET 743 "Introduction to R and Statistics"
Type: 1-credit, 1/3 semester course of 75min lectures to graduate students.
Role: **Course director**, course creator, instructor; co-instructed with Leslie Lange.
Years:

- Spring 2014: 9/12 lectures, 3/4 homeworks, 11.25 contact hours, 13 enrolled students.
- Spring 2013: 6/12 lectures, 3/4 homeworks, 7.5 contact hours, 22 enrolled students.

GNET 641 "Bioinformatics" – Course Module **"Introduction to R: calculate, plot, analyze"**
Type: 3-credit, full semester, lectures and computer practicals for graduate students; split into 3 modules.
Role: **Module director**, module creator, instructor.

- Spring 2012: 6/6 2hr lectures, 12 contact hours, 2 homeworks, 8 enrolled students.
- Spring 2011: 5/5 1.5hr lectures, 7.5 contact hours, 2 homeworks, 7 enrolled students.

EXAMINATIONS:

Examiner, BCB Graduate Curriculum Qualifying Exam
Semesters (Role):

- Summer 2015 (exam creation and grading)
- Summer 2013, 2014 (exam creation and grading; **Chair of Exam Committee**)
- Summer 2012 (exam creation and grading)
- Summer 2011 (exam creation and grading)
- Summer 2010 (exam grading only)

LECTURES:

Guest lecture "Statistical genetics in model organisms: QTL mapping in experimental crosses" (75min)
In BCB 725 "Statistical Genetics" (Course director: Yun Li), a 3-credit, full semester, course to graduate students.

- Apr 2015
- Apr 2012

Sep 2010: **Short course lecture** "Mapping QTLs in Outbred Populations" (90min)
In "Short Course on Systems Genetics", Jackson Laboratory, Bar Harbor, ME; a residential training course for graduate students, postdocs and professionals.

Aug 2008: **Short course lectures/labs (2)** "Complex Trait Genetics in Model Organisms" and "Introduction to R" (6 hr lecture/lab)
In "The 22nd International Statistical Genetics Methods Workshop", Leuven, Belgium; a residential training course for graduate students, postdoc and professionals.

Mar 2008: **Short course lecture/lab** "Introduction to R"

In "The International Workshop on Methodology of Twin and Family Studies: Introductory Course", Boulder, CO; a residential training course for graduate students, postdoc and professionals.

May 2007: **Guest lecture** "Genetic mapping of Human Behavioural Traits: Anxiety/Depression and Rodents" (60min)
Lecture to undergraduates in Human Sciences program, Psychology Dept, University of Oxford.

Mar 2007: **Short course lectures/labs (2)** "Animal models I & II" (3hr lecture/lab, co-lectured with Dr Jonathan Flint).
In "The International Workshop on Methodology of Twin and Family Studies: Advanced Course", Boulder, CO; a residential training course for graduate students, postdoc and professionals.

TEACHING RECORD: POSTDOCTORAL TRAINING AND SUPERVISION

Dates	Postdoc	Primary Supervisor(s)	Department	Current Position
Aug 2010-May 2014	Alan Lenarcic	Valdar	Genetics, UNC	Postdoctoral Research Fellowship, Quantitative Analytics Unit, Securities and Exchange Commission, New York, NY
Jul 2013-Apr 2014	Jeremy Sabourin	Valdar	Genetics, UNC	Postdoctoral Research Fellowship, NHGRI Computational and Statistical Genetics Branch, Johns Hopkins University, Baltimore, MD

TEACHING RECORD: GRADUATE TRAINING AND SUPERVISION

Note: BBSP = Biological and Biomedical Sciences Program, an umbrella program encompassing multiple graduate curricula at UNC Chapel Hill. BCB = Bioinformatics and Computational Biology curriculum, a specialization within BBSP.

PHD THESIS STUDENTS

Dates	Student	Degree program	Primary Supervisor(s)	Department [curriculum, if applicable]; Thesis title; Current or last position
Apr 2016-present	Yanwei Cai	PhD	Valdar	Genetics, UNC [BCB]
Apr 2015-present	Wesley Crouse	PhD	Samir Kelada, Valdar	Genetics, UNC [BCB]
Jul 2013-present	Robert Corty	MD/PhD	Valdar	Genetics, UNC [BCB]
May 2013-present	Paul Maurizio	PhD	Mark Heise, Valdar	Genetics, UNC [BCB]
May 2013-present	Gregory Keele	PhD	Valdar	Genetics, UNC [BCB]
Apr 2013-present	Daniel Oreper	PhD	Valdar	Genetics, UNC [BCB]
Aug 2012-Aug 2015	Yuying Xie	PhD	Yufeng Liu, Valdar	Statistics and Operations Research, UNC; "Estimation of graphical models with biomedical applications"; Assistant Professor, Michigan State University, MI
Feb 2010-May 2014	Zhaojun Zhang	PhD	Wei Wang, Valdar	Computer Science, UNC; "Efficient computational genetics methods for multiparent crosses"; Software Engineer, Coursera Inc., Mountain View, CA
Jan 2010-Jun 2013	Jeremy Sabourin	PhD	Valdar, Andrew Nobel	Statistics and Operations Research, UNC; "LASSO based Resample Model Averaging for Genetic Association Studies"; Postdoc at Computational and Statistical Genomics Branch, NHGRI, Baltimore, MD

MS THESIS STUDENTS

Dates	Student	Degree program	Primary Supervisor(s)	Department [Curriculum, if applicable]

Jan 2013-Apr 2013	Brittany Baur	MS	Leah Solberg Woods, Valdar	Graduate School, Marquette University
Jul 2008- Aug 2008	Jon Krohn	MS	Jonathan Flint, Valdar	WTCHG, Oxford University [Neuroscience Doctoral Training Program]

SHORT PROJECT / ROTATION STUDENTS

Dates	Student	Degree program	Primary Supervisor(s)	Department [Curriculum, if applicable]
May 2016-Aug 2016	Kathie Sun	PhD	Valdar	Genetics, UNC [BBSP]
Mar 2016-Apr 2016	Yanwei Cai	PhD	Valdar	Genetics, UNC [BBSP]
Nov 2015-Feb 2016	Youli Xia	PhD	Valdar	Genetics, UNC [BBSP]
Nov 2015-Feb 2016	Thanh Hoang	PhD	Valdar	Genetics, UNC [BBSP]
Nov 2014-Feb 2015	Wesley Crouse	PhD	Samir Kelada, Valdar	Genetics, UNC [BBSP]
Jun 2013-Aug 2013	Robert Corty	MD/PhD	Valdar	Genetics [MD/PhD program and BCB]
Jan 2013-Mar 2013	Bailey Peck	PhD	Valdar	Genetics, UNC [BBSP]
Jan 2013-Mar 2013	Paul Maurizio	PhD	Mark Heise, Valdar	Genetics, UNC [BBSP]
Sep 2012-Dec 2012	Gregory Keele	PhD	Valdar	Genetics, UNC [BBSP]
Sep 2011-Dec 2011	Daniel Oreper	PhD	Valdar	Genetics, UNC [BBSP]
Jul 2007-Aug 2007	Afidalina Tumian	PhD	Valdar, Jordana Tzenova Bell	WTCHG, Oxford University [Life Sciences Interface Doctoral Training Program]

FELLOWSHIP AWARDS TO STUDENTS UNDER MY SUPERVISION/SPONSORSHIP

Dates of Award	Student	Award type	Organization	Details
Aug 2016-July 2018	Robert Corty	F30	NIH/NIMH	F30-MH108265
Jan 2016-Dec 2017	Daniel Oreper	Fellowship	PhRMA	Award in Informatics (20k tuition/year)

TEACHING RECORD: SERVICE ON GRADUATE THESIS COMMITTEES

Dates	Student	Degree program	Primary Supervisor(s)	Department
Aug 2014-present	Bryan Quach	PhD	Terry Furey	Genetics, UNC
Jun 2014-present	Meredith Corley	PhD	Alain Laederach	Biology, UNC
Dec 2011-Aug 2014	Catie Welsh	PhD	Leonard McMillan	Computer Science, UNC
Apr 2013-May 2014	Ting-Huei Chen	PhD	Wei Sun	Biostatistics, UNC
Feb 2010-May 2014	Zhaojun Zhang	PhD	Wei Wang, Valdar	Computer Science, UNC
Jan 2013-Nov 2013	Gene Urrutia	PhD	Michael Wu	Biostatistics, UNC
Jan 2012-Nov 2013	John Calaway	PhD	Fernando Pardo Manuel de Villena	Genetics, UNC
Nov 2012-Jul 2013	Min Jin Ha	PhD	Wei Sun	Biostatistics, UNC
Jan 2010-May 2013	Jeremy Sabourin	PhD	Valdar, Andrew Nobel	Statistics and Operations Research, UNC
Jan 2013-Apr 2013	Brittany Baur	MS	Leah Solberg Woods, Valdar	Graduate School, Marquette University
Sep 2012-Mar 2013	Eric Liu Yi	PhD	Wei Wang, Yun Li	Computer Science, UNC

RESEARCH SUPPORT: ACTIVE

R01 (Solberg Woods) 07/01/2015 - 06/30/2020 1.2 calendar
NIDDK/Medical College of Wisconsin $32,815 (sub directs) Co-I/Key Personnel
Systems genetics of adiposity traits in outbred rats
The major goals of the subcontract are to assist in the statistical analysis of genetic and gene expression data from a study of adiposity phenotypes in HS rats. I am a sub-recipient leading the statistical analysis and remotely co-supervising a postdoc based at MCW.

1R21 HL126045-02 (Lange E. and Lange L.) 07/01/2014 - 06/30/2017 0.84 calendar
NHLBI $99,451 Co-I/Key Personnel
Priority score 24; Percentile 10
The interplay between genes and environment on cardiovascular disease phenotypes

The major goal of this project is to identify gene-environment interactions in human genetic epidemiological CVD data using a novel analytic approach based on recently developed statistical methods for identifying predictors of phenotypic variability. My role is to assist in the design and supervision of the statistical genetic analysis, and provide statistical genetics expertise.

1R01MH100241-04 (Tarantino, Valdar) 04/01/2013-03/31/2018 2.4 calendar
NIH/NIMH/NICHD $398,532 Multi-PI
Priority Score: 10; Percentile 1 (Genetics of Health and Disease Study Section [GHD])
Role of maternal diet and allelic imbalance in behavior
The goal of this proposal is to identify genes whose actions are modulated by maternal nutrition and that influence behavior in adult offspring. The experimental design makes use of the Collaborative Cross (CC), creating a highly specific sparse diallel cross of CC lines that will be tested for parent of origin and genetic effects on behavior. My role as Co-PI focuses on supervising the statistical design of the diallel cross, analysis of expression data from this cross, and the analysis of correlations between phenotypic and expression traits in matched mice subjected to behavioral testing, which in turn is used to prioritize candidate genes for follow up analysis. The project includes adapting hierarchical Bayesian methods diallel analysis already developed by my laboratory to a highly specific sparse diallel design.

1R01GM104125-05 (Valdar) 09/01/2012 - 08/30/2017 5.4 calendar
NIH/NIGMS $160,000 PI
Priority Score: 15; Percentile 4 (Genomics, Computational Biology and Technology Study Section [GCAT])
Statistical Modeling of Complex Traits in Genetic Reference Super-Populations
This project aims to develop a Bayesian statistical framework combining analysis and design of experiments on experimental resource populations that share common founders. My role as PI is to supervise development of statistical methods for Bayesian hierarchical modeling of haplotype effects on complex traits in eight-founder individuals, use existing hierarchical models developed by my laboratory to establish a decision theoretic framework for powering future experiments, and to develop Bayesian model selection methods that exploit shared founder ancestry in the analysis of multiple QTLs.

RESEARCH SUPPORT: PENDING

U01DA043106 (Tarantino) 11/01/2016-10/31/2021 2.4 calendar
NIH/NIDA $497,793 Co-I/Key Personnel
Identifying addiction liability genes using a two-stage multiparent design: A sequential survey, mapping and fine-mapping approach using Collaborative Cross and Diversity Outbred mice
Identify mouse strains that show contrasting liability for behaviors related to cocaine addiction, use these contrasting strains in a novel design to map and fine-map causative genes

---

RESEARCH SUPPORT: COMPLETED

R01 DK088975 (Solberg Woods) 7/01/2010 - 6/30/2015 1.5 calendar
NIDDK/Medical College of Wisconsin $23,379 (sub directs) Co-I/Key Personnel
Genome-wide fine-mapping of metabolic traits in heterogeneous stock rats
The major goals of the subcontract are to assist in the statistical analysis of genetic data from a study of diabetes phenotypes in HS rats. I am a sub-recipient leading the statistical analysis and remotely co-supervising a postdoc based at MCW.

P50 HG006582 (Pardo-Manuel de Villena) 09/01/2011 – 08/31/2014 0.48 calendar
NHGRI Co-I
An interdisciplinary program for systems genomics of complex behaviors
The goal of this project is to align existing expertise at UNC-CH into a center of excellence in order to develop as a resource and demonstrate the utility of the murine Collaborative Cross (CC). In particular it focuses on investigating genetic and environmental determinants of psychiatric traits. The development of sophisticated, user-interactive databases to access the large, complex datasets collected represents a key component of the project. This is a continuation of the project previously funded by P50MH090338. My role is to advise on statistical genetics issues.

P50 MH0903380-02 (Pardo-Manuel de Villena) 09/30/2009-08/31/2011 1.2 calendar
NIMH Investigator
An interdisciplinary program for systems genomics of complex behaviors
The goal of this project is to align existing expertise at UNC-CH into a center of excellence in order to develop as a resource and demonstrate the utility of the murine Collaborative Cross (CC). In particular it focuses on investigating genetic and environmental

determinants of psychiatric traits. The development of sophisticated, user-interactive databases to access the large, complex datasets collected represents a key component of the project. My role was to advise on statistical genetics issues.

G0701612 Career Development Fellowship in Biostatistics (Valdar) 07/01/2008 – 07/30/2011* 12 calendar
Medical Research Council, UK PI
Multivariate dissection of quantitative trait gene networks
This award, which is similar to an NIH K-award, provided 3 years salary, equipment and travel costs for development of statistical methods for the multivariate analysis of disease phenotypes in heterogeneous stock (HS) mouse populations. The award was based at the Wellcome Trust Centre for Human Genetics, Oxford University. Sponsor: Peter Donnelly PhD FRS.
*Terminated early by PI on 07/31/2009 in order to move from Oxford to tenure track employment at UNC-CH.

R01 AR055271-01 (Benoist) 09/10/2007 – 09/09/2008 N/A
NIAMS/Joslin Institute subcontract to Oxford University Investigator
Genetics of Arthritis Susceptibility in Outbred HS Mice
Genetic mapping of quantitative trait loci affecting arthritis susceptibility in heterogeneous stock mice. I was a sub-recipient engaged in statistical analysis of the data (CFDA No. 93.846)

Programme Grant (Flint) 11/01/2001-10/31/2007 6 calendar
Wellcome Trust, UK Staff
Simultaneous fine-mapping of quantitative trait loci for multiple phenotypes
This project exploited then-novel statistical techniques developed for high resolution QTL mapping in outbred animals to find the genetic determinants of a number of phenotypes of medical interest, using animal models of asthma, diabetes, learning and memory and fear responses. I was employed as a postdoc, developing and applying bioinformatics tools and analytic methods at all levels of the project

Access to Research Infrastructures Fellowship (Valdar) 1/1/2005 – 31/5/2005 1 calendar
European Commission PI
Developing a method to fine-map multiple epistatic quantitative trait loci in heterogeneous stocks using a Lipschitz-optimized random effects model.
This was a competitive award providing one month salary (split between Jan and May) and expenses based at the University of Uppsala, Sweden. Sponsor Orjan Carlborg.

PROFESSIONAL SERVICE: PEER-REVIEWING GRANTS

Jul 2014 Medical Research Council, UK, *ad hoc* member – *Ultimately recused due to COI*
Jun 2014 NIH Study Section "Biobehavioral regulation, learning and ethology" (BRLE), *ad hoc* member.

PROFESSIONAL SERVICE: EDITORIAL APPOINTMENTS

Apr 2015- Genetics, Associate Editor for Statistical Genetics and Genomics

---

PROFESSIONAL SERVICE: PEER-REVIEWING PAPERS

I have served as an *ad hoc* reviewer for the following international scientific journals:
Bioinformatics
Genes, Brains and Behavior
Genetics
Genetics, Selection and Evolution
Hereditas
Journal of Dairy Science
Journal of Molecular Biology
Journal of the Royal Statistical Society
Journal of Theoretical Biology
Nature Genetics
Nature Review Genetics
Oncogene
Proteins

PROFESSIONAL SERVICE: CONSULTING

8/2009-5/2015 [Consulting prohibited by H1B visa status]

10/2006	Oregon Health Sciences University, Portland, OR, USA. Statistical genetics expert.
9/2006	Oak Ridge National Laboratory, Oak Ridge, TN, USA. Mouse statistical genetics expert.
9/2001	University College London. Contract programmer.
4/1999	Inpharmatica Ltd, 60 Charlotte Street, London. Consultant programmer.
7/1997 – 9/1997	GlaxoWellcome R&D, Stevenage. MS project student in bioinformatics.

–

PROFESSIONAL SERVICE: UNC-CH BIOINFORMATICS AND COMPUTATIONAL BIOLOGY GRADUATE PROGRAM

Apr 2014-	Associate Director
Mar 2013-	Chair, Exam Committee.
Apr 2012-	Chair, Curriculum Committee.
May 2011-	Member, Executive Committee.
May 2011-	Member, Graduate Exam Committee.
Dec 2009- Mar 2013	Member, Graduate Admissions Committee.

PROFESSIONAL SERVICE: UNC-CH CHAPEL HILL OTHER COMMITTEES

Aug 2014-	Member, Data Studies Planning Committee, The Graduate School and The Provost's Office.
Jan 2014- May 2014	Member, Faculty Search Committee for 1-2 joint positions in Dept Biostatistics and Dept Genetics [Committee Chair: Danyu Lin, Biostatistics]
Jan 2009-	Member, Advisory Board, Dept of Genetics.

PROFESSIONAL MEMBERSHIPS / AFFILIATIONS

2011-	Member, Genetics Society of America
2010-	Member, Institute of Mathematical Statistics
2010-	Member, International Society for Bayesian Analysis
2010-	Member, American Statistical Association
2007-	Collaborator, MRC Centre for Causal Analyses in Translational Epidemiology, University of Bristol, UK.