Curriculum Vitae

Lei Sun

(Last updated May 1, 2017)

Department of Statistical Sciences	100 St. George Street
Faculty of Arts and Science	Toronto, ON M5S 3G3, Canada
Division of Biostatistics	P: (416) 978-4327, F: (416) 978-5133
Dalla Lana School of Public Health	sun@utstat.toronto.edu
University of Toronto	http://www.utstat.toronto.edu/sun

Degrees

08/2001	Ph.D.	Statistics, University of Chicago, USA *Two statistical problems in human genetics: I. Detection of pedigree errors; II. Identification of polymorphisms that explain a linkage result* Supervisor: Mary Sara McPeek
07/1995	B.S.	Mathematics, Fudan University, Shanghai, China

Academic Positions

07/2014-present	Full Professor	Department of Statistical Sciences Faculty of Arts and Science Division of Biostatistics Dalla Lana School of Public Health University of Toronto (75% FTE in Statistics and 25% in Biostatistics)
07/2007-06/2014	Associate Professor	Division of Biostatistics Dalla Lana School of Public Health University of Toronto
07/2005-06/2014	Adjunct Professor	Department of Statistical Sciences Faculty of Arts and Science University of Toronto
09/2001-06/2007	Assistant Professor	Division of Biostatistics Dalla Lana School of Public Health University of Toronto

Interruptions/Delays

04/2004-11/2004	Maternity Leave (30 weeks)
01/2007-08/2007	Maternity Leave (30 weeks)

Awards

2017	CRM-SSC prize in Statistics, award by the Centre de recherches mathématiques (CRM) and the Statistical Society of Canada (SSC) *in recognition of a statistical scientist's professional accomplishments in research during the first fifteen years after having received a doctorate*
2015	Dean's Excellence Award, Faculty of Arts and Science 5% Merit Pool for outstanding performance across the Faculty, University of Toronto
2014	Dean's Excellence Award, Faculty of Arts and Science 5% Merit Pool for outstanding performance across the Faculty, University of Toronto
2006	Dean's Award, Faculty of Medicine's 5% Merit Pool for academic achievement in research, teaching and service, University of Toronto
2004	Dean's Award, Faculty of Medicine's 5% Merit Pool for academic achievement in research, teaching and service, University of Toronto
1996-2001	University of Chicago Fellowship
1996-1997	Paul Meier Fellowship, Department of Statistics, University of Chicago

Research Grants

2016-2019	Cystic Fibrosis Canada (CFC) *Genetic epidemiology of cystic fibrosis* PI: Lisa Strug 5 co-Is: Tanja Gonska, Felix Ratjen, Peter Durie, Johanna Rommens, **Lei Sun** $120,571/year; $361,713 in total
2014-2019	Canadian Institutes of Health Research (CIHR) *Statistical methods and computational tools for next-generation genetic studies of complex traits* PI: **Lei Sun** 2 co-Is: Andrew Paterson, Lisa Strug $133,528/year; $667,641 in total
2013-2018	Natural Sciences and Research Council of Canada (NSERC) *Joint analyses of multiple genetic variants and multiple trait phenotypes* PI: **Lei Sun** (Individual) $23,000/year; $115,000 in total
2012-2017	Canadian Institutes of Health Research (CIHR) *Modification of cystic fibrosis phenotypes in the pancreas and gut* PI: Lisa Strug 3 co-Is: Peter Durie, Johanna Rommens, **Lei Sun** $214,079/year; $1,070,395 in total

2012-2015 Juvenile Diabetes Research Foundation (JDRF)
Genetics of the decline in glomerular filtration rate in type 1 diabetes
PI: Andrew Paterson
6 co-Is: Ian de Boer, Andrew Boright, Shelley Bull, Barbara Klein, Ronald Klein, **Lei Sun**
US $485,827/year; US $1,457,483 in total

2011-2012 McLaughlin Centre Accelerator Grant in Genomic Medicine
Developing novel hypothesis driven sequence analysis methodology towards improved therapeutics for cystic fibrosis
PI: **Lei Sun**
3 co-Is: Peter Durie, Johanna Rommens, Lisa Strug
$75,000 in total
Featured as one of the "10 Big Stories in Personalized Medicine"

2011-2012 McLaughlin Centre Accelerator Grant in Genomic Medicine
Identification of genes for severe early diabetic nephropathy
PI: Andrew Paterson
3 co-Is: Andrew Boright, Shelley Bull, **Lei Sun**
$75,000 in total

2008-2013 Natural Sciences and Research Council of Canada (NSERC)
Statistical methods for complex genetic data
PI: **Lei Sun** (Individual)
$14,000/year; $70,000 in total

2007-2012 Canadian Institutes of Health Research (CIHR)
Design and analysis of genome-wide studies of complex diseases and traits
co-PI: **Lei Sun**; 2 other co-PIs: Shelley Bull, Radu V. Craiu
$108,606/year; $543,030 in total

2006-2009 National Institute of Health (NIH)
Genome-wide association of common alleles with long-term diabetic complications
PI: Andrew Paterson
5 co-Is: Andrew Boright, Shelley Bull, Patricia Cleary, John Lachin, **Lei Sun**
US $1,391,927/year; US $4,175,781 in total

2003-2006 Canadian Institutes of Health Research (CIHR)
Statistical methods to improve the reliability of results from genome-wide studies of complex disease and quantitative traits
co-PI: **Lei Sun**; 1 other co-PI: Shelley Bull
$87,000/year $261,000 in total

2002-2007 Natural Sciences and Research Council of Canada (NSERC)
Mapping genetic variants for complex disease via statistical methods for positional cloning
PI: **Lei Sun** (Individual)
$14,500/year; $72,500 in total

2001-2002 University of Toronto Connaught Start-up Grant
PI: **Lei Sun** (Individual)
$10,000 in total

Training/Workshop/Equipment Grants

2017-2022 Genome Canada - Genomics Technology Platforms:
Operations Support and Technology Development Funds
The Centre for Applied Genomics (TCAG)
co-Leaders: Stephen Scherer, Lisa Strug
co-Is: **Lei Sun** among over 10 co-Is
$9,174,603 in total

2017-2020 Canadian Statistical Sciences Institute (CANSSI)
Health Science Collaborating Centre: Collaborating Centre for Statistical Omics
co-PIs: Shelley Bull, Laurent Biollais, Rafal Kustra, Lisa Strug, **Lei Sun**
$10,000 in total

2014-2015 The Canadian Statistical Sciences Institute (CANSSI) - Workshop Grant
Statistical issues in biomarker and drug co-development
Coordinator: Judy-Anne Chapman
Organizing committee members: **Lei Sun** among 14 members
$17,145 in total

2010-2016 Canadian Institutes of Health Research (CIHR) Training Grant
STAGE (Strategic Training for Advanced Genetic Epidemiology): An integrated program in statistical & epidemiological training for genetics with a population health impact
co-PIs and co-Directors: France Gagnon, Shelley Bull
3 other co-PIs: Steven Narod, Andrew Paterson, **Lei Sun**
$294,969/year and 5,000 for equipment &1,774,814 in total

2005-2006 Natural Sciences and Research Council of Canada (NSERC) Equipment Grant
Enhanced computing resources for statistical research
PI: Radford Neal
co-Is: **Lei Sun** among over 10 co-Is
$88,000 in total

2002-2003 Natural Sciences and Research Council of Canada (NSERC) Equipment Grant
High performance parallel computer server for statistics and biostatistics
PI: James Stafford
co-Is: **Lei Sun** among over 10 co-Is
$44,728 in ottal

Research Contributions (*Google Citations*)

Trainees underlined and **Sun L*** for senior or co-senior corresponding authorship

Refereed Book Chapters

5. **Sun L*** (2017). Detecting pedigree relationship errors. In *Statistical Human Genetics, 2nd Edition*, Elston R (Editor). Human Press, Inc. Springer.

4. **Sun L***, Dimitromanolakis A, Chen WM (2017). Identifying cryptic relationships. In *Statistical Human Genetics, 2nd Edition*, Elston R (Editor). Human Press, Inc. Springer.

3. Craiu RV, **Sun L*** (2014). Bayesian methods in Fisher's statistical genetics world. In *Statistics in Action: A Canadian Perspective*, Lawless JF (Editor).

2. **Sun L*** (2012). Detecting pedigree relationship errors. In *Statistical Human Genetics: Methods and Protocols*, Elston R, Satagopan J and Sun S (Editors), Human Press, Inc. Springer, pp.25-46.

1. **Sun L***, Dimitromanolakis A (2012). Identifying cryptic relationships. In *Statistical Human Genetics: Methods and Protocols*, Elston R, Satagopan J and Sun S (Editors), Human Press, Inc. Springer, pp.47-58.

Refereed Journal Publications

60. Soave D, **Sun L*** (in press). A generalized Levene's scale test for variance heterogeneity in the presence of sample correlation and group uncertainty. *Biometrics*.

59. Yoo YJ, **Sun L**, Poirier J, Paterson AD, Bull SB (2017). Multiple-linear-combination (MLC) regression tests for common variants adapted to linkage disequilibrium structure. *Genetic Epidemiology* 41(2):108-121.

58. Strug LJ, Gonska T, He G, Keenan K, Ip W, Boelle PY, Lin F, Panjwani N, Gong J, Li W, Soave D, Xiao B, Tullis E, Rabin H, Parkins MD, Price A, Zuberbuhler PC, Corvol, H, Ratjen F, **Sun L**, Bear CE, Rommens JM (2016). Cystic fibrosis gene modifier SLC26A9 modulates airway response to CFTR-directed therapies. *Human Molecular Genetics* 25 (20): 4590-4600.

57. Xu L, Craiu RV, **Sun L***, Paterson AD (2016). Parameter expanded algorithms for Bayesian latent variable modeling of genetic pleiotropy data. *Journal of Computational and Graphical Statistics* 25(2):405-425.

56. Derkach A, Lawless JF, **Sun L*** (2015). Score tests for association under response-dependent sampling designs for expensive covariates. *Biometrika* 102(4):988-994.

55. Corvol H, Blackman S, Boelle PV, Gallins P, Pace R, Stonebraker J, Accurso F, Clement A, Collaco J, Dang H, Dang A, Franca A, Gong J, Guillot L, Keenan K, Li W, Lin F, Patrone M, Raraigh K, **Sun L**, Zhou YH, O'Neal W, Sontag M, Levy H, Durie P, Rommens J, Drumm M, Wright F, Strug L, Cutting G, Knowles M (2015). Genome-wide association meta-analysis identifies five modifier loci of lung disease severity in cystic fibrosis. *Nature Communications* 6. doi:10.1038/ncomms9382

54. Poirier JG, Faye LL, Dimitromanolakis A, Paterson AD, **Sun L**, Bull SB (2015). Resampling to address the winner's curse in genetic association analysis of time to event. *Genetic Epidemiology* 39(7):518-528.

53. Soave D, Corvol H, Panjwani N, Gong J, Wei Li, Boelle PV, Durie P, Paterson AD, Rommens JM, Strug LJ, **Sun L*** (2015). A joint location-scale test improves power to detect associated SNPs, gene-sets and pathways. *The American Journal of Human Genetics* 97(1):125-138.

    Selected for the inaugural "Trainee Paper Spotlight" that highlights outstanding publications and research done by trainees, by the American Society of Human Genetics' Training and Development Committee.

    Recommended by Faculty of 1000 In F1000Prime.

52. Miller MR, Soave D, Li W, Gong J, Pace RG, BoŚlle PV, Cutting GR, Drumm ML, Knowles, MR, **Sun L**, Rommens JM, Accurso F, Durie PR, Corvol H, Levy H, Sontag MK, Strug LJ (2015). Variants in solute carrier SLC26A9 modify prenatal exocrine pancreatic damage in cystic fibrosis. *Journal of Pediatrics* 166(5):1152-1157.

51. Hosseini SM, Boright AP, **Sun L**, Canty AJ, Bull SB, Klein BE, Klein R, the DCCT/EDIC Research Group, Paterson AD (2015). The association of previously reported polymorphisms for microvascular complications in a meta-analysis of diabetic retinopathy. *Human Genetics* 134(2):247-257.

50. Soave D, Miller M, Keenan K, Li W, Gong J, Ip W, Accurso F, **Sun L**, Rommens JM, Sontag M, Durie PR, Strug LJ (2014). Evidence for a causal relationship between early exocrine pancreatic disease and cystic fibrosis-related diabetes: a Mendelian randomization study. *Diabetes* 63(6):2114-2119.

49. Gonçalves VF, Zai CC, Tiwari AK, Derkach A, Meltzer HY, Lieberman JA, Mueller DJ, **Sun L***, Kennedy JL (2014). A hypothesis driven association study of 28 nuclear-encoded mitochondrial genes with antipsychotic-induced weight gain in schizophrenia. *Neuropsychopharmacology* 39:1347-1354.

48. Derkach A, Lawless J, **Sun L*** (2014). Pooled association tests for rare genetic variants: a review and some new results. *Statistical Science* 29(2):302-321.

47. Blue EM, **Sun L***, Tintle NL, Wijsman EM (2014). Value of Mendelian laws of segregation in families: data quality control, imputation and beyond. *Genetic Epidemiology* 38(S1):21-28.

46. Xu L, Craiu RV, Derkach A, Paterson AD, **Sun L*** (2014). Using a Bayesian latent variable approach to detect pleiotropy in the GAW18 Data. *BMC Proceedings* 8(S1):S77, 1-5.

45. **Sun L***, Dimitromanolakis A (2014). PREST-plus identifies pedigree errors and cryptic relatedness in the GAW18 sample using genome-wide SNP data. *BMC Proceedings* 8(S1):S23, 1-6.

44. Derkach A, Lawless J, Merico D, Paterson AD, **Sun L*** (2014). Evaluation of gene-based association tests for analyzing rare variants using Genetic Analysis Workshop 18 data. *BMC Proceedings* 8(S1):S9, 1-6.

43. Bickeboller H, Bailey JN, Beyene J, Cantor RM, Cordell HJ, Culverhouse RC, Engelman CD, Fardo DW, Ghosh S, KŽnig IR, Bermejo JL, Melton PE, Santorico SA, Satten GA, **Sun L**, Tintle NL, Ziegler A, MacCluer JW, Almasy L (2014). Genetic Analysis Workshop 18: Methods and strategies for analyzing human sequence and phenotype data in members of extended pedigrees. *BMC Proceedings* 8(S1):1-4.

42. Li W, Soave D, Miller MR, Keenan K, Lin F, Gong J, Chiang T, Stephenson AL, Durie P, Rommens J, **Sun L***, Strug LJ (2014). Unraveling the complex genetic model for Cystic Fibrosis: pleiotropic effects of modifier genes on early CF-related morbidities. *Human Genetics* 133(2):151-161.

41. Yoo YJ, **Sun L**, Bull SB (2013). Gene-based multiple regression association testing for combined examination of common and low frequency variants in quantitative trait analysis. *Frontiers in Genetics* 4(233):1-17.

40. Blackman S, Commander C, Watson C, Arcara K, Strug L, Stonebraker J, Wright F, Rommens J, **Sun L**, Pace R, Norris S, Durie P, Drumm M, Knowles M, Cutting G (2013). Genetic modifiers of cystic fibrosis-related diabetes. *Diabetes* 62(10):3627-3635.

39. Faye LL, Machiela MJ, Kraft P, Bull SB, **Sun L*** (2013). Re-ranking sequencing variants in the post-GWAS era. *PLoS Genetics* 98(8):e1003609:1-16.

38. Acar E, **Sun L*** (2013). A generalized Kruskal-Wallis test incorporating group uncertainty with application to genetic association studies. *Biometrics* 69(2):427-435.

37. Derkach A, Lawless J, **Sun L*** (2013). Robust and powerful tests for rare variants using Fisher's method to combine evidence of association from two or more complementary tests. *Genetic Epidemiology* 37(1):110-121.

    In the Genetic Epidemiology Publisher's September 2014 report, this work was among the top 10 most downloaded and cited articles published in 2013.

36. Gonçalves VF, Tiwari AK, de Luca V, Kong SL, Zai C, Tampakeras M, Mackenzie B, **Sun L**, Kennedy JL (2012). DRD4 VNTR polymorphism and age at onset of severe mental illnesses. *Neuroscience Letters* 519(1):9-13.

35. **Sun L***, Rommens JM, Corvol H, Li W, Li X, Chiang T, Lin F, Dorfman R, Busson PF, Parekh RV, Zelenika D, Blackman S, Corey M, Doshi V, Henderson L, Naughton K, O'Neal WK, Pace RG, Stonebraker JR, Wood SD, Wright FA, Zielenski J, Clement A, Drumm ML, BoŚlle PY, Cutting GR, Knowles MR, Durie PR, Strug LJ (2012). Multiple apical plasma membrane constituents are associated with susceptibility to meconium ileus in individuals with cystic fibrosis. *Nature Genetics* 44(5):562-569.

    Featured as one of the "10 Big Stories in Personalized Medicine" in the McLaughlin Centre Biennial Report, and in the Hospital for Sick Children News Release.

34. Mirea L, Infante-Rivard C, **Sun L**, Bull SB (2012). Strategies for genetic association analyses combining unrelated case-control individuals and family trios. *American Journal of Epidemiology* 176(1):70-79.

33. Wright F, Strug LJ, Doshi VK, Commander CW, Blackman SM, **Sun L**, Berthiaume Y, Cutler D, Cojocaru A, Collaco JM, Corey M, Dorfman R, Goddard K , Green D, Kent Jr JW, Lange EM, Lee S, Li W, Luo J, Mayhew GM, Naughton KM, Pace RG, ParŐ P, Rommens J, Sandford A, Stonebraker JR, Sun W, Taylor C, Vanscoy LL, Zou F, Blangero J, Zielenski J, OÕNeal WK, Drumm ML, Durie PR, Knowles MR, Cutting GR (2011). Genome-wide association and linkage identify modifier loci of lung disease severity in cystic fibrosis at 11p13 and 20q13.2. *Nature Genetics* 43(6):539-548.

32. Faye LL, **Sun L***, Dimitromanolakis A, Bull SB (2011). A flexible genome-wide bootstrap method that accounts for ranking- and threshold-selection bias in GWAS interpretation and replication study design. *Statistics in Medicine* 30(15):1898-1912.

31. **Sun L*** (2011). On the efficiency of genome-wide scans: a multiple hypothesis testing perspective. U.P.B. Sci. Bull., Series A., 73(1):19-26.

30. **Sun L***, Dimitromanolakis A, Faye L, Paterson AD, Waggott D, The DCCT/EDIC Research Group, Bull SB (2011). BR-squared: a practical solution to the winner's curse in genome-wide scans. *Human Genetics* 129(5):545-552.

29. Dorfman R, Taylor C, Lin F, **Sun L**, Sandford A, Pare P, Berthiaume Y, Corey M, Durie P, Zielenski J, on behalf of the members of the Canadian Consortium for CF Genetic Studies (2011). Modulatory effect of the SLC9A3 gene on susceptibility to infections and pulmonary function in children with cystic fibrosis. *Pediatric Pulmonology* 46(4):385-392.

28. Li W, **Sun L**, Corey M, Zou F, Lee S, Cojocaru AL, Taylor C, Blackman SM, Stephenson A, Sandford AJ, Dorfman R, Drumm ML, Cutting GR, Knowles MR, Durie P, Wright FA, Strug LJ (2011). Understanding the population structure of North American patients with Cystic Fibrosis. *Clinical Genetics* 79(2):136-146.

27. Xu L, Craiu RV, **Sun L*** (2011). Bayesian methods to overcome the winner's curse in genetic studies. *Annals of Applied Statistics* 5(1):201-231.

26. Mirea L, **Sun L***, Stafford JE, Bull SB (2010). Using evidence for population stratification bias in combined individual- and family-level genetic association analyses of quantitative traits. *Genetic Epidemiology* 34:502-511.

25. Paterson AD, Waggott D, Boright AP, Hosseini SM, Shen E, Sylvestre MP, Wong I, Bharaj B, Cleary PA, Lachin JM, MAGIC, Below JE, Nicolae D, Cox NJ, Canty AJ, **Sun L**, Bull SB, and the DCCT/EDIC Research Group (2010). A genome-wide association study identifies a novel major locus for glycemic control in type 1 diabetes, as measured by both HbA1c and glucose. *Diabetes* 59:539-549.

24. Yoo YJ, Bull SB, Paterson AD, Waggott D, DCCT/EDIC Research Group, **Sun L*** (2010). Were genome-wide linkage studies a waste of time? Exploiting candidate regions within genome-wide association studies. *Genetic Epidemiology* 34:107-118.

23. Paterson AD, Lopes-Virella MF, Waggott D, Boright AP, Hosseini SM, Carter RE, Shen E, Mirea L, Bharaj B, **Sun L**, Bull SB, and the DCCT/EDIC Research Group (2009). Genome-wide association identifies the ABO blood group as a major locus associated with serum levels of soluble E-Selectin. *Arteriosclerosis, Thrombosis, and Vascular Biology* 29:1958-1967.

22. Dorfman R, Li W, **Sun L**, Lin F, Wang Y, Sandford S, Pare PD, McKay K, Kayserova H, Macek M, Bal J, Sands D, Tiddens H, Castro S, Sontag M, Accurso FJ, Blackman S, Cutting GR, Tsui LC, Corey M, Durie P, Zielenski J, Strug LJ (2009). Modifier gene study of meconium ileus in cystic fibrosis: statistical considerations and gene mapping results. *Human Genetics* 126:763-778.

21. Yoo YJ, Pinnaduwage D, Waggott D, Bull SB, **Sun L*** (2009). Genome-wide association analyses of North American Rheumatoid Arthritis Consortium and Framingham Heart Study data utilizing genome-wide linkage results. *BMC Proceedings* 3:S103.

20. Asimit J, Yoo YJ, Waggott D, **Sun L**, Bull SB (2009). Region-based analysis in genome-wide association study of Framingham Heart Study blood lipid phenotypes. *BMC Proceedings* 3:S127.

19. Craiu RV, **Sun L*** (2008). Choosing the lesser evil: trade-off between false discovery rate and non-discovery rate. *Statistica Sinica* 18:861-879.

18. Lee SSF, **Sun L***, Kustra R, Bull SB (2008). EM-random forest and new measures of variable importance for multi-Locus quantitative trait linkage analysis. *Bioinformatics* 24:1603-1610.

17. Dorfman R, Sandford A, Taylor C, Huang B, Frangolias D, Wang Y, Sang R, Pereira L, **Sun L**, Berthiaume Y, Tsui LC, Pare PD, Durie P, Corey M, Zielenski J (2008). Complex two-gene modulation of lung disease severity in children with cystic fibrosis. *Journal of Clinical Investigation* 118:1040-1049.

16. Al-Kateb H, Boright AP, Xie X, Mirea L, Sutradhar R, Mowjoodi A, Bharaj B, Liu M, Bucksa JM, Arends VL, Steffes MW, Cleary PA, Sun W, Lachin JM, Thorner PS, Ho M, McKnight AJ, Maxwell PA, Savage DA, Kidd KK, Kidd JR, Speed WC, Orchard TJ, Miller RG, **Sun L**, Bull SB, Paterson AD and the DCCT/EDIC research group (2008). Multiple superoxide dismutase 1/splicing factor serine alanine 15 variants are associated with the development and progression of diabetic nephropathy. *Diabetes* 57:218-228.

15. Al-Kateb H, Mirea L, Xie X, **Sun L**, Liu M, Chen H, Bull SB, Boright AP, Paterson AD, The DCCT/EDIC Research Group (2007). Multiple variants in Vascular Endothelial Growth Factor (VEGF) are risk factors for time to severe retinopathy in type 1 diabetes: The DCCT/EDIC genetics studies. *Diabetes* 56:2161-2168.

14. Huang B, Rangreg J, Paterson AD, **Sun L*** (2007). The multiplicity problem in linkage analysis of gene expression data - the power of differentiating *cis*- and *trans*-acting regulators. *BMC Proceedings* 1:S142.

13. Greenwood C, Rangrej J, **Sun L*** (2007). Optimal selection of markers for validation from genome-wide association studies. *Genetic Epidemiology* 31:396-407.

12. Wu LY, **Sun L***, Bull SB (2006). Locus-specific heritability estimation via the bootstrap in linkage scans for quantitative trait loci. *Human Heredity* 62:84-96.

11. **Sun L***, Craiu RV, Paterson AD and Bull SB (2006). Stratified false discovery control for large-scale hypothesis testing with application to genome-wide association studies. *Genetic Epidemiology* 30:519-530.

10. Wu LY, Lee SSF, Shi HS, **Sun L**, Bull SB (2005). Resampling methods to reduce the selection bias in genetic effect estimation in genome-wide scans. *Genetic Analysis Workshop 14:* Microsatellite and single-nucleotide polymorphism. *BMC Genetics* 6:S24.

9. Biernacka J, **Sun L***, Bull SB (2005). Tests for the presence of two linked disease susceptibility genes. *Genetic Epidemiology* 29:389-401.

8. **Sun L***, Bull SB (2005). Reduction of selection bias in genome-wide genetic studies by resampling. *Genetic Epidemiology* 28:352-367.

7. Biernacka J, **Sun L***, Bull SB (2005). Simultaneous localization of two linked disease susceptibility genes. *Genetic Epidemiology* 28:33-47.

6. Paterson A, **Sun L**, Liu XQ (2003). Transmission ratio distortion in families from the Framingham Heart Study. *Genetic Analysis Workshop 13:* Analysis of longitudinal family data for complex diseases and related risk factors. *BMC Genetics* 4:S48.

5. Strug LJ, **Sun L**, Corey M (2003). The genetics of cross-sectional and longitudinal BMI. *Genetic Analysis Workshop 13:* Analysis of longitudinal family data for complex diseases and related risk factors. *BMC Genetics* 4:S14.

4. **Sun L***, Wilder K, McPeek MS (2002). Enhanced pedigree error detection. *Human Heredity* 54:99-110.

3. **Sun L***, Cox NJ, McPeek MS (2002). A statistical method for identification of polymorphisms that explain a linkage result. *American Journal of Human Genetics* 70:399-411.

2. **Sun L***, Abney M, McPeek MS (2001). Detection of misspecified relationships in inbred and outbred pedigrees. *Genetic Analysis Workshop 12*: Analysis of complex genetic traits: Applications to asthma and simulated data. *Genetic Epidemiology* 21:S36-S41.

1. McPeek MS, **Sun L*** (2000). Statistical tests for detection of misspecified relationships by use of genome-screen data. *American Journal of Human Genetics* 66:1076-1094.

Publicly-Released Research Software

An important component of this research program is the implementation of developed and tested methodology as user-friendly and open-resource software and programs as part of the Research Dissemination and Knowledge Transfer and Exchange. All programs are freely available at http://www.utstat.toronto.edu/sun/

10. Soave D, **Sun L*** (2017). *gJLS*, a generalized Joint Location-Scale association testing framework that tests the null hypothesis of equal mean and equal variance between genotypes, allowing for sample correlation and group membership uncertainty.

9. Soave D, Strug LJ, **Sun L*** (2015). *JLS*, a Joint Location-Scale association testing framework that detects both main and interaction effects.

8. Acar E, **Sun L*** (2011). *GKW*, a Generalized Kruskal-Wallis test that incorporates group uncertainty.

7. Dimitromanolakis A, **Sun L*** (2010). *PREST-plus*, for detecting pedigree errors and cryptic relationships.

6. Dimitromanolakis A, Faye L, Bull SB, **Sun L*** (2009). *BR-squared*, Bias-Reduced estimates via Bootstrap Resampling for overcoming the winner's curse due to selective inference

5. Yoo YJ, **Sun L*** (2008). *sFDR*, a stratified False Discovery Rate control for multiple hypothesis testing.

4. Lee SSF, **Sun L**, Kustra R, Bull SB (2007). *EMRF*, a EM-Random Forest approach for multi-locus quantitative trait linkage analysis.

3. Wu LY, **Sun L***, Bull SB (2006). *BR-squared - Linkage* for calculation of Bias-Reduced Bootstrap Resampling-based estimates of locus-specific gene-effect size in linkage analyses

2. Wen W, **Sun L*** (2004). *STEPC*, STatistical Explanation of Positional Cloning for identification of polymorphisms that explain a linkage result.

1. **Sun L***, Wilder K, McPeek MS (2000). *PREST*, Pedigree RElationship Statistical Test and ALTERTEST for detecting pedigree errors and estimation of relationships using genome-wide marker data.

Patents

1. Patent application for "Modifiers of CFTR-directed therapy", WIPO (PCT) Patent Appln No. CA2016/051044, filed September 2, 2016. Inventors are

    - Johanna Rommens, Hospital for Sick Children, Toronto, Ontario, Canada
    - Lisa Strug, Hospital for Sick Children, Toronto, Ontario, Canada
    - Lei Sun, University of Toronto, Toronto, Ontario, Canada

Refereed Conference Abstracts

For dissemination purpose, each project/topic is usually presented at multiple venues. Most of the conference abstracts have resulted in journal paper publications. The first author is the presenter.

177. Zhang L, **Sun L*** (2017). A robust allele-based regression framework for testing association and Hardy-Weinberg equilibrium. Abstract #16 presented at *the 2017 Statistics Graduate Student Research Day: 40 years of Statistical Sciences at the University of Toronto*. The Fields Institute, Toronto, Canada.

176. Chen B, Craiu RV, **Sun L***. Bayesian model averaging for the X-chromosome dilemma in genetic association studies. Abstract #2 presented at *the 2017 Statistics Graduate Student Research Day: 40 years of Statistical Sciences at the University of Toronto*. The Fields Institute, Toronto, Canada.

175. Yoo YJ, **Sun L**, Poirier J, Paterson AD, Bull SB (2016). Multi-variant linear regression tests with reduced degrees of freedom for association analysis of common variants. Abstract #172 presented at *the International Genetic Epidemiology Society (IGES) annual meeting*. Toronto, Canada.

174. Gossens ET, **Sun L***. Combining evidence after selection: a powerful framework for testing the global null hypothesis (2016). Abstract #81 presented at *the International Genetic Epidemiology Society (IGES) annual meeting.* Toronto, Canada.

173. Deng W, **Lei Sun*** (2016). X-Inclusion: analyzing X-chromosome in whole genome association studies of variance heterogeneity. Abstract #61 presented at *the International Genetic Epidemiology Society (IGES) annual meeting.* Toronto, Canada.

172. Chen B, Craiu RV, **Lei Sun*** (2016). Bayesian model averaging approach for the X-Inactivation dilemma in genetic association studies. Abstract #50 presented at *the International Genetic Epidemiology Society (IGES) annual meeting.* Toronto, Canada.

171. He G, Gonska T, Keenan K, Ip W, BŽelle PY, Lin F, Panjwani N, Gong J, Li W, Soave D, Xiao B, Tullis E, Rabin H, Parkins M, Price A, Zuberbuhler P, Corvol H, Ratjen F, **Sun L**, Bear CE, Strug LJ, Rommens JM (2016). Airway response to CFTR-directed therapeutics for cystic fibrosis is modulated by SLC26A9 . Abstract #2644W presented at *the 66th Annual Meeting of The American Society of Human Genetics (ASHG).* Vancouver, Canada.

170. Hosseini SM, Snell-Bergeon JK, Boright AP, Canty AJ, **Sun L**, Bull SB, Marcovina SM, Brunzell JD, Paterson AD, the DCCT/EDIC Research Group (2016). Identifying genetic variants for serum lipoprotein(a) independent of kringle repeat polymorphism. Abstract #1229T presented at *the 66th Annual Meeting of The American Society of Human Genetics (ASHG).* Vancouver, Canada.

169. Gossens ET, **Sun L*** (2016). Is selective inference more powerful in meta-, gene-based and pathway analyses? Abstract #550W presented at *the 66th Annual Meeting of The American Society of Human Genetics (ASHG).* Vancouver, Canada.

168. Xiao B, Panjwani N, Gong J, He G, Keenan K, Lin F, Soave D, Drumm M, Cutting G, Knowles M, Corvol H, **Sun L**, Rommens JM, Strug LJ (2016). Meta-GWAS in cystic fibrosis indicates common variation in regulatory regions of modifier genes contributes to meconium ileus. Abstract #491T presented at *the 66th Annual Meeting of The American Society of Human Genetics (ASHG).* Vancouver, Canada.

167. Deng W, **Lei Sun*** (2016). X-inclusion: Integrating X chromosome in whole genome association studies of variance heterogeneity. Abstract #466W presented at *the 66th Annual Meeting of The American Society of Human Genetics (ASHG).* Vancouver, Canada.

166. Bull SB, Yoo YJ, **Sun L** (2016). Multiple-linear-combination (MLC) regression tests for gene-based association analysis of common variants adapted to linkage disequilibrium structure. Abstract presented at the *XXVIIIth International Biometric Conference (IBC 2016).* Victoria, Canada (Selected for platform presentation).

165. Soave D, **Sun L*** (2015). A generalized joint location-scale association test for uncertain genotypes and related individuals. Abstract #1361F presented at *the 65th Annual Meeting of The American Society of Human Genetics (ASHG)*. Baltimore USA.

164. Soave D, **Sun L*** (2015). A generalized joint location-scale association test for uncertain genotypes and related individuals. Abstract #141 presented at *the International Genetic Epidemiology Society (IGES) annual meeting*. Baltimore, USA.

163. Soave D, **Sun L*** (2015). Generalized Levene's test of homoscedasticity for correlated data with group uncertainty. Abstract presented at the *Joint Statistical Meetings (JSM)*. Seattle USA.

162. Lawless JF, Derkach A, **Sun L*** (2015). Outcome-dependent sampling in two-phase genetic association studies. Invited abstract presented at the *annual meeting of the Statistical Society of Canada (SSC)*. Halifax, Canada.

161. Chen B, Craiu RV, **Sun L*** (2015). Statistical analysis of sex chromosome in genome-wide association studies. Abstract presented at the *annual meeting of the Statistical Society of Canada (SSC)*. Halifax, Canada.

160. Hiraki L et al. (2015). Genetic signal for renal function decline in type 1 diabetes (T1D) identified by analysis of repeated longitudinal measures. Abstract #2015-A-4399-Diabetes presented at the *American Diabetes Association's 75th Scientific Sessions*. Boston, USA. (Selected for platform presentation).

159. Yoo Y, Kim S, **Sun L**, Bull S (2014). Linkage disequilibrium clustering can improve power of weighted-sum-type multi-marker tests for genetic association analysis. Abstract #1791S presented at *the 64th Annual Meeting of The American Society of Human Genetics (ASHG)*. San Diego, USA.

158. Soave D, Miller M, Keenan K, Li W, Gong J, Ip W, Durie P, **Sun L**, Rommens J, Strug L (2014). What are genome-wide association studies detecting? Our experience predicting cystic fibrosis-related diabetes onset. Abstract #1778T presented at *the 64th Annual Meeting of The American Society of Human Genetics (ASHG)*. San Diego, USA.

157. Hosseini S, Howard K, **Sun L**, Boright AP, Tregouet DA, Sandholm N, Hietala K, Toppila I, Lajer MS, Marre M, Rossing P, Groop PH, Canty AJ, Hadjadj S, Klein BE, Bull SB, Klein R, Paterson AD, the DCCT/EDIC Research Group (2014). Polymorphism upstream of cryopyrin gene (NLRP3) is associated with severe retinopathy in type 1 diabetes. Abstract #946S presented at *the 64th Annual Meeting of The American Society of Human Genetics (ASHG)*. San Diego, USA.

156. Gonçalves VF (2014). A hypothesis-driven analysis of genome-wide association summary results from the psychiatric genomics consortium identifies novel nuclear-encoded

mitochondria susceptibility loci for schizophrenia. Abstract presented at the *World Congress of Psychiatric Genetics*. Copenhagen, Denmark.
(Selected as one of the three ECIP Oral Presentation Award Finalists).

155. Strug LJ, Soave D, Miller MR, Gong J, Li W, Keenan K, Durie PR, **Sun L**, Rommens J (2014). Risk of CF-related diabetes can be predicted within the first years of life. Abstract presented at the *North American Cystic Fibrosis Annual Conference*. Atlanta, USA.

154. Miller MR, Soave D, Li W, Gong J, Levy H, Corvol H, Cutting GR, Drumm ML, Knowles MR, Durie P, **Sun L**, Rommens JM, Accurso F, Sontag MK, Strug LJ (2014). Variant in the solute carrier SLC26A9 modifies newborn weight and early weight gain in Cystic Fibrosis. Abstract presented at the *North American Cystic Fibrosis Annual Conference*. Atlanta, USA.

153. Hiraki L, Boer ID, Canty AJ, **Sun L**, Boright AP, Klein R, Klein BE, Bull SB, DCCT/EDIC Research Group, Paterson AD (2014). Genetic variants associated with repeated longitudinal measures of renal function in type 1 diabetes (T1D). Abstract #2014-A-1772-Diabetes presented at the *American Diabetes Association's 74th Scientific Sessions*. San Francisco, USA.

152. Gonçalves VF, Zai C, Tiwari A, Derkach A, Meltzer H, Lieberman J, Paterson AD, **L. Sun**, Mueller D, Kennedy JL (2013). A hypothesis driven association analysis of nuclear-encoded mitochondrial genes with antipsychotic-induced weight gain in schizophrenia subjects. Abstract #699W presented at *the 63rd Annual Meeting of The American Society of Human Genetics (ASHG)*. Boston, USA.

151. Hosseini SM, Howard K, **Sun L**, Boright AP, Tregouet DA, Sandholm N, Lajer MS, Hietala K, Forsblom C, Marre M, Rossing P, Groop PH, Canty AJ, Hadjadj S, Klein BE, Bull SB, Klein R, Paterson AD, the DCCT/EDIC Research Group (2013). Dipeptidyl peptidase 10 (DPP10) is associated with severe retinopathy in type 1 diabetes. Abstract #889T presented at *the 63rd Annual Meeting of ASHG*. Boston, USA.

150. Soave D, Paterson AD, Strug LJ, **Sun L** (2013). Improved detection of variants with main or interaction effects using a robust location-scale testing framework. Abstract #1848T presented at *the 63rd Annual Meeting of ASHG*. Boston, USA.

149. Gong J, Lin F, Chiang T, Keenan K, Miller M, Soave D, Li W, **Sun L**, Rommens JM, Strug LJ (2013). Abstract #1889T presented at *the 63rd Annual Meeting of ASHG*. Boston, USA.

148. **Sun L**, Derkach A, Lawless JF (2013). Response dependent sampling designs and analysis in studies with rare variants. Abstract #1929T presented at *the 63rd Annual Meeting of ASHG*. Boston, USA.

147. Eny KM, **Sun L**, Canty AJ, Bull SB, Boright AP, Hosseini SM, Cleary PA, Lachin JM, Paterson AD, the DCCT/EDIC Research Group (2013) Are loci associated with low density lipoprotein cholesterol (LDL-C) in non-diabetics similarly associated with LDL-C in type 1 diabetes (T1D)? Abstract #2108F presented at *the 63rd Annual Meeting of ASHG*. Boston, USA.

146. Miller MR, Soave D, Li W, Chiang T, Gong J, Levy H, **Sun L**, Rommens JM, Accurso F, Durie P, Sontag MK, Strug LJ (2013). The Solute Carrier SLC26A9 Accounts for Variability in Biomarkers of Cystic Fibrosis-Related Prenatal Exocrine Pancreatic Damage. Abstract #2941F presented at *the 63rd Annual Meeting of ASHG*. Boston, USA.

145. Gonçalves VF, Zai C, Tiwari A, Derkach A, Meltzer H, Lieberman J, Paterson AD, **L. Sun**, Mueller D, Kennedy JL (2013). A hypothesis driven association study of 28 nuclear-encoded mitochondrial genes with antipsychotic-induced weight gain in schizophrenia Abstract #74 presented at *the XXI_{st} World Congress of Psychiatric Genetics (WCPG)*. Boston, USA.

144. Blackman SM, Commander CW, Watson C, Arcara K, Strug LJ, Stonebraker JR, Wright FA, Rommens JM, **Sun L**, Pace RG, Norris S, Durie PR, Drumm ML, Knowles M, Cutting GR (2013). Genome-wide association studies for Type 2 Diabetes and CFRD reveal common risk loci. Abstract # 165 presented at the 27th Annual North American Cystic Fibrosis Conference (NACFC). Salt Lake City, USA.

143. Miller MR, Soave D, Li W, Chiang T, Gong J, Levy H, **Sun L**, Rommens JM, Accurso F, Durie PR, Sontag MK, Strug LJ (2013). Genetic modification of prenatal exocrine pancreatic disease. Abstract # 166 presented at the NACFC 27th Annual North American Cystic Fibrosis Conference. Salt Lake City, USA.

142. Gonska T, Dupuis A, Keenan K, Sontag M, Castellani C, Cipolli M, Naehrlich L, Dorfman R, Taylor C, **Sun L**, Ooi KY, Rommens JM, Strug LJ, Durie PR (2013). Prevalence of meconium ileus as a phenotypic marker for the severity of CFTR mutations. Abstract # 164 presented at the North American Society for Pediatric Gastroenterology, Hepatology and Nutrition (NASPGHAN) Annual Meeting. Chicago, USA.

141. Derkach A, **Sun L**, Lawless JF (2013). Response dependent sampling designs and analysis in studies with rare variants. Abstract presented at the *Joint Statistical Meetings (JSM)*. Montreal, Canada.

140. Acar E, **Sun L** (2013). A generalized kruskal-wallis test incorporating group uncertainty with application to genetic association studies. Abstract presented at the *Joint Statistical Meetings (JSM)*. Montreal, Canada.

139. Craiu RV, Xu L, **Sun L** (2013). Bayesian latent variable modelling of longitudinal family data for genetic pleiotropy studies. Abstract presented at the *Joint Statistical Meetings (JSM)*. Montreal, Canada.

138. **Sun L**, Derkach A, Lawless JF (2013). Statistical methods for association studies of rare variants. *The 8th Conference on Extreme Value Analysis*. Shanghai, China.

137. Hosseini MS, Boright AP, **Sun L**, Howard K, Canty AJ, Bull SB, Klein BE, Klein R, Paterson AD and the DCCT/EDIC research group (2013). A meta genome-wide association study of time-to diabetic retinopathy. 2^{nd} *annual Canadian Human and Statistical Genetics Meeting*. EstŐrel, Quebec

136. Li W, Su D, Chiang T, Li X, Miller MR, Keenan K, Corvol H, Wright FA, Blackman S, Drumm ML, Cutting GR, Knowles MR, Durie PR, Rommens JM, **Sun L**, Strug LJ (2012). Do severity of early lung disease and meconium ileus in cystic fibrosis have common genetic contributors? Exocrine and endocrine pancreatic damage in cystic fibrosis are associated with SLC26A9. Abstract #1322W presented at *the 62nd Annual Meeting of The American Society of Human Genetics (ASHG)*. San Francisco, USA.

135. Derkach A, Lawless J, **Sun L** (2012). Robust and powerful tests for rare variants using Fisher's method to combine evidence of association from two or more complementary tests. Abstract #1412W presented at *the 62nd Annual Meeting of ASHG*. San Francisco, USA.

134. Hosseini SM, Boright AP, **Sun L**, Howard K, Canty AJ, Bull SB, Klein BE, Klein R, Paterson AD, the DCCT/EDIC research group (2012). Genome-wide association study of time-to diabetic retinopathy. Abstract #2109W presented at *the 62nd Annual Meeting of ASHG*. San Francisco, USA.

133. Gonçalves VF, Zai C, Paterson AD, **Sun L**, Kennedy JL, Knight J (2012). GWAS of alcohol abuse and dependence in the CATIE schizophrenia sample. Abstract #2376W presented at *the 62nd Annual Meeting of ASHG*. San Francisco, USA.

132. Soave D, Chiang T, Miller MR, Su D, Keenan K, Li W, Ip W, Wright FA, Blackman S, Corvol H, Knowles MR, Cutting GR, Drumm M, **Sun L**, Rommens JM, Durie PR, Strug LJ (2012). Exocrine and endocrine pancreatic damage in cystic fibrosis are associated with SLC26A9. Abstract #2953W presented at *the 62nd Annual Meeting of ASHG*. San Francisco, USA.

131. Derkach A, Lawless J, **Sun L** (2012). Combining p-values from linear and quadratic tests for rare variants provides robust statistics across genetic models. Abstract # 7 presented at *the International Genetic Epidemiology Society (IGES) annual meeting*. Stevenson, USA.
(Selected for platform presentation and nominated as one of the three finalists for the Williams Award; selected for the IGES Travel Award).

130. Faye LL, Bull SB, Kraft P, **Sun L** (2012). Re-ranking next generation sequencing variants for accurate causal variant identification. Abstract #8 presented at *the*

IGES annual meeting.Stevenson, USA.
(Selected for platform presentation and won the Williams Award).

129. Taleban J Faye LL, Dimitromanolakis A, Paterson AD, **Sun L**, Bull, SB (2012). A general bias-reduction procedure to address the winner's curse in genetic association analysis: evaluation for time-to-event phenotypes. Abstract #86 presented at *the IGES annual meeting*. Stevenson, USA.

128. Chen Z, Paterson AD, Canty AJ, **Sun L**, Bull SB (2012). Joint modelling of repeated quantitative trait measures and time to event in longitudinal genetic association studies. Abstract #104 presented at *the IGES annual meeting*. Stevenson, USA.

127. Canty A, Wang Tao, Bull SB, **Sun L**, Boright A, DCCT/EDIC Research Group, Paterson AD (2012). GWAS of repeated lipid measures in type 1 diabetes identifies a novel locus for low-density lipoprotein cholesterol. Abstract #105 presented at *the IGES annual meeting*. Stevenson, USA.

126. Xu L, Craiu RV, Derkach A, Paterson AD, **Sun L** (2012). Using a Bayesian latent variable approach to detect pleiotropy in the GAW18 Data. Abstract presented at the *Genetic Analysis Workshop 18*. Stevenson, USA.
(Selected for GAW Student Travel Award and the University of Toronto McLaughlin Centre Training Award).

125. **Sun L**, Dimitromanolakis A (2012). PREST-plus identifies pedigree errors and cryptic relatedness in the GAW18 sample using genome-wide SNP data. Abstract presented at the *Genetic Analysis Workshop 18*. Stevenson, USA.

124. Derkach A, Lawless J, Merico D, Paterson AD. **Sun L** (2012). Evaluation of association tests and gene annotations for analyzing rare variants using GAW18 data. Abstract presented at the *Genetic Analysis Workshop 18*. Stevenson, USA.
(Selected for GAW Student Travel Award and the University of Toronto McLaughlin Centre Training Award).

123. Gonska T, Keenan K, Sontag M, Castellani C, Cipolli M, Naehrlich L, Dupuis A, Dorfman R, Taylor C, **Sun L**, Ooi CY, Rommens J, Strug LJ, Durie, PR. (2012). The meconium ileus prevalence score as a new phenotypic marker for the severity of CFTR mutations. Abstract #445 presented at the 26th Annual North American Cystic Fibrosis Conference (NACFC). Orlando, USA.

122. Durie PR, Soave D, Gonska T, Ip W, Keenan K, Miller M, **Sun L**, Rommens J, Strug LJ (2012). Early ecocrine pancreatic damage determined by serum immunoreactive trypsinogen is significant predictor of CF-related diabetes at a later age. Abstract #530 presented at the 26th Annual NACFC. Orlando, USA.

121. Yoo YJ, Bull SB, **Sun L** (2012). Comparisons of SNP partitioning strategies for regression-based multi-marker scanning of genetic association. Abstract presented at the *Joint Statistical Meetings (JSM)*. San Diego, USA.

120. **Sun L**, Derkach A, Lawless J (2012). Combining p-values from linear and quadratic tests for rare variants provides robust statistics across genetic models. Abstract presented at the *Joint Statistical Meetings (JSM)*. San Diego, USA.

119. Derkach A, Lawless JF, **Sun L** (2012). Combining linear and quadratic tests for rare variants provides a robust test across genetic models. Abstract presented at the *annual meeting of the Statistical Society of Canada (SSC)*. Guelph, Canada.

118. Taleban J, Faye LL, **Sun L**, Dimitromanolakis A, Bull SB (2012). Bootstrap bias-reduction in genetic association analysis of time-to-event outcomes. Abstract presented at the *annual meeting of the Statistical Society of Canada (SSC)*. Guelph, Canada.

117. Gonçalves VF, Tiwari AT, de Luca V, Kong SL, Zai C, Tampakeras M, Mackenzie B, **Sun L**, Kennedy JL (2012). A functional polymorphism in DRD4 gene is associated with age at onset of severe mental illnesses. Abstract presented at the XX_{th} *World Congress of Psychiatry Genetics*. Hamburg, Germany.

116. Derkach A, Lawless J, **Sun L** (2012). Combining p-values from linear and quadratic tests for rare variants provides robust statistics across genetic models. *The* 1^{st} *annual Canadian Human and Statistical Genetics Meeting*. Niagara-on-the-Lake, Canada. (Selected for platform presentation).

115. Chen Z, Paterson AD, Canty AJ, **Sun L**, Bull SB (2012). Specification and interpretation of joint phenotype models in genetic association of complex traits. *The* 1^{st} *annual Canadian Human and Statistical Genetics Meeting*. Niagara-on-the-Lake, Canada.

114. Faye LL, Bull SB, **Sun L** (2012). Re-ranking sequencing variants in the post-GWAS era. 1^{st} *Canadian Human and Statistical Genetics Meeting*. Niagara-on-the-Lake, Canada.

113. Gonçalves VF, Zai CC, Paterson A, **Sun L**, Kennedy JL, Knight J (2012). Genome-wide association study of alcohol abuse and dependence in two schizophrenia samples. 1^{st} *Canadian Human and Statistical Genetics Meeting*. Niagara-on-the-Lake, Canada.

112. Durie PR, Soave D, Gonska T, Ip W, Keenan K, Miller M, **Sun L**, Rommens J, Strug LJ (2012) Early exocrine pancreatic damage determined by serum immunoreactive trypsinogen is a significant predictor of CF-related diabetes at a later age. Abstract presented at the *North American Cystic Fibrosis Conference*, Orlando, USA. *Pediatric Pulmonology* 47(S35):408

111. Derkach A, Lawless J, **Sun L** (2012), Combining p-values from linear and quadratic tests for rare variants provides robust statistics across genetic models. *Statistics Graduate Student Research Day - Models for Dependent Data.* (Selected for platform presentation)

110. Li X, Corvol, H, Li W, Chiang T, Lin F, Boelle P-Y, Drumm M, Cutting G, Knowles M, Durie P, Rommens J, **Sun L**, Strug L (2011). Replication evidence that constituents of the apical plasma membrane contribute to Meconium ileus in Cystic Fibrosis. Abstract #1134T presented at *the 12th International Congress of Human Genetics (ICHG)/61st Annual Meeting of The American Society of Human Genetics (ASHG)*. Montreal, Canada.

109. Eny KM, Bull SB, Canty AJ, **Sun L**, Boright AP, Hosseini SM, Cleary PA, Lachin J, Paterson AD, DCCT/EDIC Research Group (2011). Genome-wide association analysis of lactose consumption measured longitudinally identifies a novel variant 500kb downstream of the LCT gene region. Abstract #882T presented at *the 12th ICHG/61st annual meeting of ASHG*. Montreal, Canada.

108. Faye L, Bull SB, **Sun L** (2011). Re-ranking sequencing variants in the post-GWAS era. Abstract #743W presented at *the 12th ICHG/61st annual meeting of ASHG*. Montreal, Canada.

107. Taleban J, Faye LL, Dimitromanolakis A, Paterson AD, **Sun L**, Bull SB (2011). Extensions of Bootstrap bias-reduction to address the winner's curse in genome-wide association analysis of time-to-event phenotypes. Abstract #729F presented at *the 12th ICHG/61st annual meeting of ASHG*. Montreal, Canada.

106. Derkach A, Lawless J, **Sun L** (2011). A generalized pooled association statistic for analyzing rare variants. Abstract #723F presented at *the 12th ICHG/61st annual meeting of ASHG*. Montreal, Canada.

105. Chen Z, Paterson AD, Canty AJ, **Sun L**, Bull SB (2011). Joint modelling of repeated measures and time-to-event data in genetic association analysis of type 1 diabetes Abstract #722F presented at *the 12th ICHG/61st annual meeting of ASHG*. Montreal, Canada.

104. Li W, Li X, Lin F, Chiang T, Drumm M, Knowles M, Cutting G, Durie P, Rommens J, **Sun L**, Strug L (2011). Investigating the ÔmissingheritabilityÕ of Meconium Ileus in Cystic Fibrosis: contributions from a hypothesis-driven GWAS (GWAS-HD). Abstract #686F presented at *the 12th ICHG/61st annual meeting of ASHG*. Montreal, Canada.

103. Xu L, Craiu RV, Paterson AD, **Sun L** (2011). Detecting pleiotropy via Bayesian latent variable modeling. Abstract #613F presented at *the 12th ICHG/61st annual meeting of ASHG*. Montreal, Canada.

102. Hosseini SM, Howard K, Boright AP, **Sun L**, Canty AD, Bull SB, Klein BE, Klein R, Paterson AD (2011) Using time-to-event analysis for genome-wide association study in a long-term cohort of diabetic retinopathy. Abstract #423W presented at *the 12th ICHG/61st annual meeting of ASHG*. Montreal, Canada.

101. Gagnon F, Dimitromanolakis A, Antoni G, Martinez A, Greliche N, de Buil A, Soria JM, Morange PE, Wells PS, Tregouet DA, **Sun L** (2011) Prioritized-GWAS based on linkage information identifies novel putative loci influencing coagulation. Abstract #368W presented at *the 12th ICHG/61st annual meeting of ASHG*. Montreal, Canada.

100. Gagnon F, Dimitromanolakis A, Antoni G, Martinez A, Greliche N, de Buil A, Soria JM, Morange PE, Wells PS, Tregouet DA, **Sun L** (2011). Prioritized-GWAS based on linkage information identifies novel putative loci influencing coagulation. Abstract #18 presented at *the International Genetic Epidemiology Society (IGES) annual meeting*. Heidelberg, Germany.

99. Eny KM, Bull SB, Canty AJ, **Sun L**, Boright AP, Hosseini SM, Cleary PA, Lachin J, DCCT/EDIC Research Group, Paterson AD (2011). Genome-wide association study (GWAS) of lactose consumption measured longitudinally identifies a novel variant 500kb downstream of the LCT gene region. Abstract #101 presented at *the IGES annual meeting*. Heidelberg, Germany.

98. Yoo YJ, **Sun L**, Bull SB (2011). Regression based multi-marker tests for gene-based analysis of genetic association. Abstract #302135 presented at
it the annual Joint Statistical Meetings (JSM). Miami Beach, USA.

97. Gagnon F, Dimitromanolakis A, Antoni G, Martinez A, Greliche N, Buil A, Soria JM, Morange P, Wells PS, Tregouet D, **Sun L** (2011). Prioritized-GWAS based on linkage information identifies novel putative loci influencing Factor XII plasma levels. Abstract presented at *the XXIII International Society on Thrombosis & Haemostasis (ISTH) Congress with 57th Annual SSC Meetings.* Kyoto, Japan.

96. Gonska T, Keenan K, Dorfman R, Taylor C, **Sun L**, Ooi CY, Zielenski J, Corey M, Rommens J, Strug L, Durie P (2011). The relation between meconium ileus prevalence score and functional severity of CFTR mutations. Abstract presented at *the 34th European Cystic Fibrosis Conference*. Hamburg, Germany.
Selected for platform presentation

95. Derkach A, Lawless J, **Sun L** (2011). A unified statistical framework for association methods for rare variants. Abstract presented at *the* 6^{th} *Canadian Genetic Epidemiology and Statistical Genetics Workshop*. King City, Canada.
(Selected for platform presentation)

94. Li W, Li X, Lin F, Chiang T, Durie P, Rommens J, **Sun L**, Strug, LJ (2011). Determining the proportion of missing heritability recovered by hypothesis-driven GWAS (GWAS-HD) of Meconium Ileus: an application of penalized regression. Abstract presented at *the* 6^{th} *Canadian Genetic Epidemiology and Statistical Genetics Workshop*. King City, Canada.

93. Xu L, Craiu RV, **Sun L** (2011). Bayesian modeling using latent variables for genetic pleiotropy studies. Abstract presented at *the* 6^{th} *Canadian Genetic Epidemiology and Statistical Genetics Workshop*. King City, Canada.

92. Faye L, Bull SB, **Sun L** (2011). Re-ranking of sequencing variants improves accuracy in targeted sequencing studies. Abstract presented at *the* 6^{th} *Canadian Genetic Epidemiology and Statistical Genetics Workshop*. King City, Canada.
(Selected for platform presentation)

91. Derkach A, Lawless J, **Sun L** (2011). Evaluation of composite statistics for association analysis of rare variants. *The Inaugural Biostatistics Research Day.*
(Recipient of the 1st Prize of the Student Poster Presentation Award).

90. Faye L, Bull SB, **Sun L** (2011). Accuracy in targeted sequencing studies. *The Inaugural Biostatistics Research Day.*

89. Jiang H, **Sun L**, Dimitromanolakis A, Gagnon F (2011). Genome-wide linkage and association analyses of APCR trait. *The Inaugural Biostatistics Research Day.*

88. Li W, Li X, Lin F, Chiang T, Durie P, Rommens J, **Sun L**, Strug LJ (2011). Recovering the missing heritability in Meconium Ileus using hypothesis-driven GWAS (GWAS-HD): an application of penalized logistic regression. *The Inaugural Biostatistics Research Day.*

87. Faye L, Bull SB, **Sun L** (2011). Re-ranking of sequencing variants improves accuracy in targeted sequencing studies. *The CRM Workshop - Computational Statistical Methods for Genomics and Systems Biology.*

86. Acar E, **Sun L** (2011). A rank-based association test that incorporates uncertainty in imputed SNPs. *The CRM Workshop - Computational Statistical Methods for Genomics and Systems Biology.*

85. Acar E, **Sun L** (2010). A generalized Kruskal-Wallis test for association analysis of imputed SNPs. Abstract #2794 presented at *the annual meeting of the American Society of Human Genetics (ASHG)*. Washington DC, USA.

84. Hosseini SM, Waggott D, Shen E, Boright AP, Sylvestre MP, Sun W, Cleary PA, Canty AJ, **Sun L**, Bull SB, Paterson AD, the DCCT/EDIC Research Group (2010). Genome-wide association study of refractive error in 1304 individuals. Abstract #1114 presented at *the annual meeting of the American Society of Human Genetics*. Washington DC, USA.

83. Xu L, Craiu RV, **Sun L** (2010). Bayesian methods to overcome the winner's curse in genetic studies. Abstract #1108 presented at *the annual meeting of the American Society of Human Genetics*. Washington DC, USA.

82. Strug LJ, **Sun L**, Li W, Dorfman R, Taylor C, Wright F, Henderson L, Drumm M, Knowles M, Cutting G, Rommens J, Durie P, on behalf of the North American Cystic Fibrosis Gene Modifier Consortium (2010). Coming full circle: an hypothesis-driven GWAS (GWAS-HD) with application to Meconium Ileus in Cystic Fibrosis (2010). Abstract #219 presented at *the annual meeting of the American Society of Human Genetics*. Washington DC, USA.
(Selected for platform presentation).

81. Strug LJ, **Sun L**, Li W, Dorfman R, Taylor C, Wright F, Henderson L, Drumm M, Knowles M, Cutting G, Rommens J, Durie P, on behalf of the North American Cystic Fibrosis Gene Modifier Consortium (2010). Constituents of the apical plasma membrane may contribute to Meconium Ileus. *The annual North American Cystic Fibrosis Conference* (selected for platform presentation).

80. Faye L, **Sun L**, Dimitromanolakis A, Bull SB (2010). Genome-wide bootstrap bias reduction for point and interval estimation that accounts for ranking- and threshold-selection bias in discovery GWAS, with implications for replication study sample size. Abstract #214 presented at the IGES annual meeting. *Genetic Epidemiology* 34:977.

79. Ye C, Canty AJ, Waggott D, Sylvestre MP, Shen E, Hosseini M, Boright AP, **Sun L**, Bull SB, Paterson AD, the DCCT/EDIC Research Group (2010). A repeated measures genome wide association study of blood pressure in type 1 diabetes. Abstract #203 presented at the IGES annual meeting. *Genetic Epidemiology* 34:973.

78. Paterson AD, Hosseini SM, Waggott D, Boright AP, Shen E, Sylvestre MP, Cleary PA, Lachin JM, Below JE, Nicolae D, Cox NJ, Sandholm N, Forsblom C, Groop PH. Canty AJ, **Sun L**, Bull SB, the DCCT/EDIC Research Group (2010). Genetic variation at adenylate cyclase 5 (ADCY5) is associated with glycemic control in type 1 diabetes. Abstract #192 presented at the IGES annual meeting. *Genetic Epidemiology* 34:970.

77. Acar E, **Sun L** (2010). A rank-based association test that incorporates uncertainty in imputed SNPs. Abstract #122 presented at the IGES annual meeting. *Genetic Epidemiology* 34:950.

76. Craiu RV, **Sun L**, Xu, L (2010). Bayesian Methods to Overcome the Winner's Curse in Genetic Studies. *The* 19^{th} *International Conference on Computational Statistics.*

75. Hosseini SM, Waggott D, Boright AP, Shen E, Sylvestre MP, Cleary PA, Lachin JM, Below JE, Nicolae D, Cox NJ, Canty AJ, **Sun L**, Bull SB, Paterson AD, the DCCT/EDIC Research Group (2010). Genetic variation at adenylate cyclase 5 (ADCY5) is associated with glycemic control in type 1 diabetes. *The American Diabetes Association* Scientific Sessions.

74. Mirea L, **Sun L**, Stafford JE, Infante-Rivard C, Bull SB (2010). Incorporating evidence for population stratification bias in genetic association analysis combining individual and family data. *The 5th Canadian Genetic Epidemiology and Statistical Genetics Workshop* (Winner of the "Best Trainee Platform Presentation Award").

73. Faye L, **Sun L**, Dimitromanolakis A, Bull SB (2010). A flexible genome-wide bootstrap method that accounts for ranking- and threshold- selection bias in GWAS interpretation and replication study design. *The 5th Canadian Genetic Epidemiology and Statistical Genetics Workshop.*

72. Hosseini SM, Shen E, Waggott D, Wood AR, Boright AP, Canty AJ, **Sun L**, Bull SB, Weedon MN, Frayling TM, Marcovina SM, Brunzell JD, Paterson AD, and the DCCT/EDIC Research Group (2010). Genome-wide association study of serum lipoprotein(a) identifies multiple associated SNPs in addition to the kringle repeat polymorphism. *The 5th Canadian Genetic Epidemiology and Statistical Genetics Workshop.*

71. Paterson AD, Waggott D, Shen E, Boright A, Hosseini SM, Cleary PA, Lachin JM, **Sun L**, Bull SB, the DCCT/EDIC Research Group (2010). Genome-wide association study of risk for diabetic retinopathy in type 1 diabetes. *The American Diabetes Association* Scientific Sessions.

70. Dimitromanolakis A, Paterson AD, **Sun L** (2009). Accurate IBD inference identifies cryptic relatedness in 9 HapMap populations. Abstract #1768 presented at *the annual meeting of the American Society of Human Genetics.*

69. Li W, **Sun L**, Taylor C, Dorfman R, Cojocaru A, Zielenski J, Durie P, Corey M, Strug LJ (2009). Cystic Fibrosis: an 'inclusive' disease P demonstrated by analysis of population stratification. Abstract #1749 presented at *the annual meeting of the American Society of Human Genetics.*

68. Bull SB, Yoo YJ, **Sun L** (2009). Regression multi-marker tests for gene-based genetic association analysis. Abstract #1698 presented at *the annual meeting of the American Society of Human Genetics.*

67. Faye L, Bull SB, **Sun L** (2009). Effect estimates in a 2-stage design: from genome-wide association to sequencing. Abstract #1678 presented at *the annual meeting of the American Society of Human Genetics.*

66. Hosseini SM, Shen E, Waggott D, Boright AP, **Sun L**, Bull SB, Marcovina SM, Brunzell JD, Paterson AD, the DCCT/EDIC Research Group (2009). Genome-wide association study of serum lipoprotein(a) concentration identifies multiple associated SNPs in addition to the kringle repeat. Abstract #989 presented at *the annual meeting of the American Society of Human Genetics*

65. **Sun L**, Bull SB, Dimitromanolakis A, Faye L, Waggott D, Paterson AD, the DCCT/EDIC Research Group (2009). BR-squared: a practical solution to the winner's curse in genome-wide scans. Abstract #186 presented at *the annual meeting of the American Society of Human Genetics* (**selected for platform presentation**).

64. Paterson AD, Waggott D, Boright A, Hosseini M, Shen E, Sylvestre MP, Wong I, Bharaj B, Cleary P, Lachin J, Canty A, **Sun L**, Bull SB, the DCCT/EDIC Research Group (2009). A genome-wide association study identifies variation near SORCS1 as a major locus for glycemic control in type 1 diabetes, as measured by both HbA1c and glucose. Abstract #141 presented at the IGES annual meeting. *Genetic Epidemiology* 33:793-793.

63. Faye L, **Sun L**, Dimitromanolakis A, Bull SB (2009). A comprehensive look at the likelihood and bootstrap approaches to overcome the winnerUs curse in GWAS. Abstract #126 presented at the IGES annual meeting. *Genetic Epidemiology* 33:788-789.

62. Dimitromanolakis A, Paterson AD, **Sun L** (2009). Accurate IBD inference identifies cryptic relatedness in 9 HapMap populations. Abstract #106 presented at the IGES annual meeting. *Genetic Epidemiology* 33:782-783.

61. Mirea L, **Sun L**, Stafford JE, Infante-Rivard C, Bull SB (2009). Incorporating evidence for population stratification bias in combined analyses of case-control and case-trio data. Abstract #38 presented at the IGES annual meeting. *Genetic Epidemiology* 33:763-763.

60. Yoo YJ, **Sun L**, Bull SB (2009). Gene-based genetic association analysis by multiple regression tests. The annual Joint Statistical Meetings (selected for platform presentation).

59. Li W, Zou F, Lee S, Corey M, **Sun L**, Wright FA, Strug LJ (2009). Analysis of population tratification in North American CF patients. *The Annual North American Cystic Fibrosis Conference* (selected for platform presentation).

58. Cojocaru A, Li W, **Sun L**, Corey M, Dorfman R, Taylor C, Zielenski J, Durie P, Keenan K, L.J. Strug L (2009). Association analysis for Meconium Ileus incorporating phenotype misclassification. *The 4th Canadian Genetic Epidemiology and Statistical Genetics Workshop.*

57. Dimitromanolakis A, Paterson AD, **Sun L** (2009). PREST-plus: An improved software tool for the detection of cryptic relations and pedigree errors using high-throughput genotype data. *The 4th Canadian Genetic Epidemiology and Statistical Genetics Workshop.*

56. Faye L, Bull SB, **Sun L** (2009). Two-stage study designs combining GWA and sequencing: accuracy of genetic effect estimates. *The 4th Canadian Genetic Epidemiology and Statistical Genetics Workshop*
(Winner of the "Best Trainee Platform Presentation Award").

55. Li W, **Sun L**, Taylor C, Dorfman R, Cojocaru A, Zielenski J, Durie P, Corey M, L.J. Strug L (2009). Cystic Fibrosis: an 'inclusive' disease P demonstrated by population stratification analysis. *The 4th Canadian Genetic Epidemiology and Statistical Genetics Workshop.*

54. Yoo YJ, **Sun L**, Bull SB (2009). Regression-based multi-marker tests for gene-based analysis of genetic association data. *The 4th Canadian Genetic Epidemiology and Statistical Genetics Workshop.*

53. Paterson AD, Waggott D, Boright AP, Hosseini SM, Shen E, Sylvestre MP, Wong I, Bharaj B, Cleary PA, Lachin JM, Canty AJ, **Sun L**, Bull SB, and the DCCT/EDIC Research Group (2009). A genome-wide association study identifies variation near SORCS1 as a major locus for glycemic control in type 1 diabetes, as measured by both HbA1c and glucose. *The American Diabetes Association* Scientific Sessions (selected for platform presentation).

52. Boright AP, Shen E, Waggott D, Hosseini SM, Cleary PA, Lachin JM, Canty AJ, **Sun L**, Bull SB, Paterson AD and the DCCT/EDIC Research Group (2009). Genome Wide Association Analysis identifies genetic risk factors for Coronary Calcium in subjects with Type 1 Diabetes in the DCCT/EDIC Genetics Study. *The American Diabetes Association meeting.*

51. Faye L, Bull SB, **Sun L** (2009). Effect estimates in 2-stage design: from genome-wide association to sequencing. *Statistical Society of Canada* annual meeting.

50. **Sun L**, Chiang T, Li, SS (2008). PREST-plus: detection of pedigree error and cryptic relatedness among individuals that allows for high-throughput genotype data and adjusts for large-scale multiple hypothesis testing. Abstract #2295 presented at *the annual meeting of the American Society of Human Genetics.*

49. Faye L, Bull SB, **Sun L** (2008). Bootstrap- vs. Likelihood-based methods to reduce selection bias: variance, confidence interval estimation and computational feasibility in large-scale genetic mapping studies. Abstract #2396 presented at *the annual meeting of the American Society of Human Genetics.*

48. Mirea L, **Sun L**, Stafford JE, Bull SB (2008). Combined Individual- and Family-for Population Stratification Bias. Abstract #2439 presented at *the annual meeting of the American Society of Human Genetics.*

47. Li W, **Sun L**, Corey M, Dorfman R, Zielenski J, Durie P, Strug LJ (2008). Complexities in Linkage Analysis for Modifier Genes. Abstract #2505 presented at *the annual meeting of the American Society of Human Genetics.*

46. Strug L, Li W, **Sun L**, Huang B, Corey M, Dorfman R, Blackman S, Cutting G, Zielenski J, Durie P (2008). Modifier gene loci for meconium ileus map to chromosomes 12 and 4. *2008 North American Cystic Fibrosis Conference* (selected for platform presentation).

45. Paterson AD, Boright AP, Waggott D, Zuo Y, Mirea L, Zhu L, Huang OL, Yoo YJ, Hosseini M, **Sun L**, Bull SB, and DCCT/EDIC Research Group (2008). Genome-wide association study of 867,874 SNPs with time to long-term diabetic complications in The Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications Research Group (DCCT/EDIC). Late-breaking abstract for *American Diabetes Association*.

44. **Sun L**, Yoo YJ, Bull SB, Paterson AD, Waggott D (2008) Preserving candidate regions in the era of genome-wide association via stratified false-discovery rate control improves power. *Genetic Epidemiology* 32:675-675 (**selected for platform presentation at the International Genetic Epidemiology Society annual meeting 2008**).

43. Paterson AD, Boright AP, Waggott D, Zuo Y, Shen E, Mirea L, Zhu L, Huang O, Yoo YJ, Hosseini M, Cleary PA, Lachin JM, **Sun L**, Bull SB (2008). Genome-wide association study of time to long-term diabetic complications. *Genetic Epidemiology* 32:710-711.

42. Yoo YJ, **Sun L**, Bull SB (2008). On using prior information to improve power of genome-wide studies: weighted p-value versus stratified false discovery control. *XXIVth International Biometric Conference* (selected for platform presentation).

41. Yoo YJ, **Sun L**, Bull SB (2008). On using prior information to improve power of genome-wide studies: weighted p-value versus stratified false discovery control. *Statistical Society of Canada* annual meeting.

40. Faye L, **Sun L**, Bull SB (2008). Reducing selection bias: efficiency and robustness of parametric & non-parametric approaches. *Statistical Society of Canada* annual meeting.

39. Li W, Huang B, **Sun L**, Corey M, Rorfman R, Zielenski J, Durie, P, Strug L (2008). Non-parametric linkage results must be interpreted with caution when searching for modifier genes. *The 3rd Canadian Genetic Epidemiology and Statistical Genetics Workshop*.

38. Faye L, **Sun L**, Bull SB (2008). Reducing selection bias: comparisons of parametric & non-parametric effect estimation in genetic association studies. *The 3rd Canadian Genetic Epidemiology and Statistical Genetics Workshop*
(Winner of the "Best Trainee Platform Presentation Award").

37. Yoo YJ, Bull SB, Paterson AD, **Sun L** (2008). On using prior information to improve power of genome-wide studies: weighted p-value versus stratified false discovery control. *The 3rd Canadian Genetic Epidemiology and Statistical Genetics Workshop* (selected for platform presentation).

36. Faye L, **Sun L**, Bull SB (2007). Reducing selection bias: efficiency and robustness of parametric and non-parametric effect estimation. Abstract 2152 presented at *the annual meeting of the American Society of Human Genetics* (`http://www.ashg.org/genetics/ashg07s/index.shtml`).

35. Mirea L, Bull SB, Stafford JE, **Sun L** (2007). A Logistic regression model for combined individual- and family-level association analyses of binary traits. *Genetic Epidemiology* 31:638-639.

34. Al-Kateb H, Boright AP, Xie X, Mirea L, Sutradhar R, **Sun L**, Bull SB, Paterson AD, and the DCCT/EDIC Research Group (2007). Multiple SOD1 SNPs are associated with the development and progression of diabetic nephropathy. *Genetic Epidemiology* 31:615-615.

33. Bull SB, Xie X, Faye L, **Sun L**, Paterson AD (2007). Bias reduction in genome-wide association studies with time-to-event phenotypes. *Genetic Epidemiology* 31:606-606.

32. **Sun L** (2007). On using linkage signals to improve genome-wide association studies - weighted or stratified false discovery control? *Genetic Epidemiology* 31:498-499.

31. Mirea L, Bull SB, Stafford J, **Sun L** (2007). Combined individual- and family-level association analyses of quantitative traits. *Genetic Epidemiology* 31:489-489.

30. Lee SSF, **Sun L**, Bull SB (2007). A random forest approach to identify important interacting markers in quantitative trait linkage analysis. *Genetic Epidemiology* 31:482-482.

29. Bull SB, **Sun L**, Xie X, Wu LY, Paterson AD (2007). Two-stage genome-wide association: power and sample size for replication. *Genetic Epidemiology* 31:464-464.

28. Al-Kateb H, Mirea L, Xie XL, **Sun L**, Sutradhar R, Bharaj B, Wong I, Liu XQ, Bull SB, Boright AB, Paterson AD (2007). Association of multiple variants in superoxide dismutase 1 (SOD1) with diabetic nephropathy: The DCCT/EDIC genetics study. *Diabetes* 56:A91-A91.

27. Kateb H, Mirea L, Xie XL, **Sun L**, Bull SB, Liu M (2007). Multiple variants in vascular endothelial growth factor (VEGF) gene are risk factors for severe retinopathy in type I diabetes: The DCCT/EDIC genetics study. *Diabetes* 56:A25-A25

26. Dorfman R, Sandford A, Corey M, Lin F, Wang V, Yuan XW, Huang Q, Markiewicz D, Mukerjee G, Master A, Taylor C, Frangolias D, **Sun L**, Pare P, Durie P, Tsui LC, Zielenski J (2006). Analysis of candidate genes as modifiers of Cystic Fibrosis. *Journal of Cystic Fibrosis* 5:S4.

25. Wu LY, **Sun L**, Bull SB (2006). Genetic effect estimation via resampling in linkage analysis of quantitative trait loci. *XXIIIrd International Biometric Conference.*

24. Craiu RV, Bull SB, Paterson AD, **Sun L** (2006). Separating the wheat from the chaff: statistical methods for false discovery control. Invited Contribution, *Statistical Society of Canada* annual meeting.

23. Wu LY, **Sun L**, Bull SB (2006). Genetic effect estimation via bootstrap in linkage analysis of quantitative trait loci. *Statistical Society of Canada* annual meeting.

22. Lee SSF, **Sun L**, Bull SB (2006). Weighted random forest to map genes jointly. *CAIMS-MITACS Joint Annual Conference*,
(2nd Prize Winner in Student Poster Competition).

21. Lee SSF, **Sun L**, Bull SB (2006) Weighted random forest to map genes in multi-locus linkage analysis. *the Public Health Sciences Research Day, University of Toronto*, (2nd Prize Winner in the poster competition).

20. Lee SSF, Bull SB, **Sun L** (2006). Weighted random forest to map genes in multi-locus linkage analysis. *The 2nd Canadian Genetic Epidemiology and Statistical Genetics Workshop*.

19. Al-Kateb H, Mirea L, Xie X, Mowjoodi A, Poloumienko A, **Sun L**, Bull SB, Boright AP, Paterson AD, and DCCT/EDIC Study Group (2006). Testing of common variations in 212 candidate genes for association with diabetic nephropathy and retinopathy: the DCCT/EDIC Genetics Study. *Diabetes* 55:A176-A176.

18. Lee SSF, Bull SB, **Sun L** (2005). Efficiency comparisons of estimates from Classical and EM Haseman-Elston regressions when IBD sharing is ambiguous. *Genetic Epidemiology* 29:261-261.

17. **Sun L**, Craiu VR (2005). Joint analysis of false discovery rate and non-discovery rate. *American Journal of Human Genetics* Supplement:A439.

16. Wu LY, **Sun L**, Bull SB (2005). Bias-reduced QTL effect size estimation via statistical resampling. *American Journal of Human Genetics* Supplement:A426.

15. Wu LY, Lee SSF, Shi HS, Lewinger JP, **Sun L**, Bull SB (2004). Resampling methods to reduce the selection bias in genetic effect estimation in genome-wide scans. *Genetic Analysis Workshop 14*.

14. Zielenski J, D Markiewicz D, Yuan X, Patel M, **Sun L**, Aznarez I, Tsui L(2004) Meconium Ileus in Cystic Fibrosis Neonates Is Associated With Polymorphic Markers in the Calcium-Activated Potassium Channel (Kcnn4) Gene. *Journal of Pediatric Gastroenterology and Nutrition* 39:S405-S406,

13. Wu LY, **Sun L**, Bull SB (2004). Robustness of resampling methods to reduce selection bias of genetic effect estimates. *Genetic Epidemiology* 27:304.

12. Biernacka J, **Sun L**, Bull SB (2004). A GEE approach for disease gene localization: Using IBD sharing proportions versus mean IBD. *Genetic Epidemiology* 27:262.

11. Biernacka J, **Sun L**, Bull SB (2004). Localization of linked genes for type 1 diabetes: A simultaneous search for two genes. *American Journal of Human Genetics* Supplement:A531.

10. Tevtoushenko I, Markiewicz D, Deng G, Patel M, Dorfman R, Corey M, Tan M, Li F, **Sun L**, Sandford A, Pare P, Durie P, Tsui LC, Zielenski J (2004). Preliminary analysis of human chromosome 6q27 region as a potential modifier locus for pulmonary disease in cystic fibrosis. *American Journal of Human Genetics* Supplement:A469.

9. Biernacka JM, Bull SB, **Sun L** (2004). Joint analyses of linked disease genes: Location estimation and hypothesis testing methods. Invited Contribution, *the Eastern North American Region of the International Biometric Society* annual meeting.

8. Biernacka JM, Bull SB, **Sun L**, Stafford JE (2003). A novel approach for estimating locations of multiple linked disease genes. *MITACS 4th annual conference* (1st Prize Winner in the poster competition).

7. **Sun L**, Craiu RV (2003). Sequential testing methods for pedigree error detection based on genome-screen data. *Genetic Epidemiology* 25:271-272.

6. Zielenski J, Markiewicz D, Yuan X, Patel M, **Sun L**, Liu X, Aznarez I, Tsui L-C, and the CF Modifier Collaborative Group (2003). Enhanced haplotype association for the identification of the cystic fibrosis modifier 1 gene using additional genetic markers in the region containing KCNN4. *Pediatric Pulmonology Supplement* 25:93.

5. **Sun L**, Bull S (2003). Resampling-based statistical methods to improve the estimation of locus-specific effects from genome-wide studies. *American Journal of Human Genetics* Supplement 73:609-609.

4. Biernacka J, **Sun L**, Stafford J, Bull SB (2003). Joint localization of two linked disease genes: Derivation, evaluation, and application of a new method. *American Journal of Human Genetics* Supplement 73:193-193.

3. **Sun L**, Cox NJ, McPeek MS (2000). A statistical method for identification of a functional polymorphism in a gene. *American Journal of Human Genetics* Supplement 67:A315.

2. **Sun L**, McPeek MS (1998). Detection of pedigree relationship errors from genotype data. *American Journal of Human Genetics* Supplement 63:A310.

Invited Talks/Lectures

- April 12, 2017, Department of Public Health Sciences, University of Chicago, Chicago, USA.
- March 30, 2017, Department of Statistics, University of Manitoba, Winnipeg, Canada.

- August 2, 2016, Joint Statistical Meetings (JSM), Chicago, USA.
- June 13, 2016, ICSA Applied Statistics Symposium, Atlanta, USA.
- May 30, 2016, The 2nd Conference on New Advances in Big Data, Nanjing, China.
- May 25, 2016, Institute of Statistics, Nankai University, Tianjin, China.
- December 3, 2015, Department of Biostatistics, Columbia University, New York, USA.
- April 29, 2015, Department of Economics, University Ca' Foscari of Venice, Venezia, Italy.
- April 21, 2015, School of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, Kent, UK.
- March 13, 2015, Department of Statistical Science, University of Roma, Rome, Italy.
- February 5, 2015, The French National Institute for Agricultural Research (INRA), Toulouse, France.
- August 4, 2014, Joint Statistical Meetings (JSM), Boston, USA
- May 27, 2014, Statistical Society of Canada (SSC) Annual Meeting, Toronto, Canada
- March 25, 2014, Department of Statistics, University of British Columbia, Canada
- March 21, 2014, Department of Statistics and Actuarial Science, Simon Fraser University, Canada
- March 17, 2014, International Biometric Society Eastern North American Region Regional Meeting (ENAR), Baltimore, USA
- September 25, 2013, Department of Statistical Science, Cornell University, USA
- June 4, 2013, Advanced Topics in Genome-Wide Association Studies, Ontario Institute for Cancer Research (OICR), Toronto, Canada
- April 10, 2013, Department of Statistics and Biostatistics, Rutgers University, USA
- April 12, 2012, Department of Statistics and Actuarial Science, University of Waterloo, Canada
- May 29, 2012, Advanced Topics in Genome-Wide Association Studies, Ontario Institute for Cancer Research (OICR), Toronto, Canada
- February 10, 2012, Department of Mathematics and Statistics, Acadia University, Canada

- November 14, 2011, Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, USA
- July 9, 2011, Institute of Mathematical Statistics - China International Conference on Statistics and Probability, XiAn, China
- May 12, 2011, The 6th Canadian Genetic Epidemiology and Statistical Genetics Workshop, Toronto, Canada
- September 24, 2010. The 1st Bruce Kaufman Symposium on Immunoregulation and inflammatory bowel disease, Montreal, Canada
- July 20, 2010, Banff International Research Station (BIRS) 5-day Workshop - Statistical Genomics in Biomedical Research, Banff, Canada
- January 8, 2010, Cancer Epidemiology Discussion Group, Samuel Lunenfeld Research Institute Mount Sinai Hospital, Toronto, Canada
- August 3, 2009, Joint Statistical Meetings (JSM), Washington DC, USA
- July 4, 2009, Institute of Mathematical Statistics - China International Conference on Statistics and Probability, Weihai, China
- April 3, 2009, Department of Biostatistics & Medical Informatics, University of Wisconsin, Madison, USA
- March 27, 2009, CRM-ISM-GERAD Statistics Colloque (Universite de Montreal, McGill University, Concordia University, and Universite du Quebec a Montreal), Canada
- September 30, 2008, Department of Mathematics and Statistics, U. of Guelph, Canada
- April 17, 2008, Department of Statistics, University of British Columbia, Canada
- April 16, 2008, Department of Statistics and Actuarial Science, Simon Fraser University, Canada
- October 11, 2007, Genetics & Genomic Biology, HSC Research Institute, Toronto, Canada
- June 13, 2007, Statistical Society of Canada (SSC) Annual Meeting, St. John's, Newfoundland, Canada
- July 2006, SAMSI Workshop on Multiplicity and Reproducibility in Scientific Studies, Research Triangle Park, USA
- May 2006, International Workshop on Applied Probability (IWAP), University of Connecticut, USA

- April 2006, Center for Statistical Genetics, University of Michigan at Ann Arbor, USA
- April 2006, Department of Statistics, University of Chicago, USA
- March 2006, The 1st Canadian Genetic Epidemiology and Statistical Genetics Workshop, Toronto, USA
- February 2006, Department of Mathematics and Statistics, Laval University, Canada
- December 2005, Department of Statistics, University of Toronto, Canada
- December 2004, Department of Statistics, Colorado State University, USA
- September 2003, Department of Mathematics and Statistics, York University, Canada
- March 2003, Department of Statistics and Actuarial Science, University of Waterloo, Canada
- February 2002, Genetics & Genomic Biology, HSC Research Institute, Toronto, Canda
- June 2001, Applied Statistics Symposium, International Chinese Statistical Association (ICSA), Chicago, USA
- April 2001, Department of Public Health Sciences, University of Toronto, Canada
- February 2001, Department of Biostatistics, University of North Carolina at Chapel Hill, USA
- February 2001, Department of Statistics, Ohio State University, USA

Academic/Professional Activities and Services - International

Editorial Board

10/2012-present	Editorial Board	Genetic Epidemiology
09/2011-present	Review Editor	Frontiers in Statistical Genetics and Methodology
05/2015	Guest Associate Editor	PLoS Genetics
07/2007-08/2012	Associate Editor	Statistical Applications in Genetics and Molecular Biology

Scientific Advisory Board

- 09/2010-08/2015, CIHR Net Grant entitled "Immunomodulation of Regulatory Mechanisms in Mucosal Immunity: A multi-disciplinary bench-to-bedside approach to the study and treatment of IBD". McGill University Health Centre, University of Montreal.

Organizer and Member

- 2016-2019, Member, Publication Committee, International Genetic Epidemiology Society (IGES).
- 2014, Organizing Committee Member, 2-day Workshop at the Fields Institute; "Statistical issues in biomarker and drug co-development".
- 2014, Co-Organizer, 5-day Workshop at the Banff International Research Station (BIRS) for Mathematical Innovation and Discovery; ÒEmerging Statistical Challenges and Methods For Analysis of Massive Genomic Data in Complex Human Disease Studies".
- 2014, Chair, Session on New Estimation Methods, Statistical Society of Canada (SSC) Annual Meeting, Toronto, Canada.
- 2013, Member, Program Committee, IEEE International Conference on Bioinformatics and Biomedicine (BIBM).
- 2013, Member, Program Committee, Eastern North American Region/International Biometric Society meeting (ENAR).
- 2012, Group Leader, Genetic Analysis Workshop (GAW) 18.
- 2012, Member, Program Committee, IEEE International Conference on Bioinformatics and Biomedicine (BIBM).
- 2011, Co-Organizer, Invited session on ÒStatistical Genetics", STATISTICS CANADA/IMST.
- 2010, Organizer and Chair, Invited session on ÒGenomics", Institute of Mathematical Statistics (IMS)
- 2008, Member, Education Committee, International Genetic Epidemiology Society (IGES)
- 2007, Member, Education Committee, International Genetic Epidemiology Society (IGES)
- 2007, Member, Publication Committee, International Genetic Epidemiology Society (IGES)
- 2005, Organizer and Chair, Invited session on "Statistical analysis and modeling of complex traits", Joint Statistical Meeting (JSM)
- 2004, Organizer, Invited session on "Statistical genetics - modeling interaction and multi-locus analyses", Eastern North American Region/International Biometric Society meeting (ENAR)

Journal Referee

American Journal of Human Genetics (AJHG), Annals of Applied Statistics (AoAS), Annals of Human Genetics, Bioinformatics, Biometrics, BioTechniques, BMC Bioinformatics, BMC Genetics, proceedings of the Genetic Analysis Workshop, BMC Medical Genetics, BMC Medical Research Methodology, Canadian Journal of Statistics (CJS), European Journal of Human Genetics (EJHG), Genetics, Genetic Epidemiology, Human Genetics, Human Heredity, IEEE BIBM special issues, Journal of the American Statistical Association (JASA), Obesity, PLoS Genetics, PLoS One, Statistical Applications in Genetics and Molecular Biology, Statistics in Medicine, Statistica Sinica

Grant Review

Canadian Statistical Sciences Institute (CANSSI), Canadian Breast Cancer Foundation (CBCF), Canadian Institutes of Health Research (CIHR), German Research Foundation (Deutsche Forschungsgemeinschaft), Israel Science Foundation (ISF), Natural Sciences and Engineering Research Council of Canada (NSERC), Internal grant reviewer for Hospital for Sick Children (HSC) Research Institute, Toronto

Other International Services

- As external reviewer for various Full-Professor, Associate-Professor and Tenure promotions.

Professional Activities and Services - Local

Department of Statistical Sciences, FAS, University of Toronto

- 2017, Member of the Organizing Committee, Chair of the Poster Session, Statistics Graduate Student Research Day: 40 years of Statistical Sciences at the University of Toronto. The Fields Institute, Toronto, Canada.
- 2016-present, Chair, Research Committee
- 2016-present, Member, Executive Committee
- 2015-present, Member, Social Committee
- 2015-present, Member, Hiring/Search Committee
- 2015-present, Member, Graduate Studies Committee
- 2015-present, Member, Undergraduate Data Science Committee

- 2015-present, Member, Data Science Committee
- 2016, Member, PTR Committee
- 2015-2016, Co-Chair, Social Committee
- 2015-2016, Member, Promotion Committee
- 2013-2014, Member, Hiring/Search Committee
- 2010-2011, Member, Hiring/Search Committee
- 2008-2009, Member, Hiring/Search Committee

Dalla Lana School of Public Health, University of Toronto

- 2010-present, Member, Program Steering Committee, CIHR "STAGE (Strategic Training in Genetic Epidemiology): An integrated program in statistical & epidemiological training for genetics with a population health impact"
- 2010-present, Member, Admission Committee, CIHR STAGE Training Program
- 2002-present, Member, Admission Committee, Division of Biostatistics
- 2002-present, Member, Comprehensive Exam Committee, Division of Biostatistics
- 2013-2014, Member, Hiring/Search Committee, Division of Biostatistics
- 2013, Member, Executive Committee, School Council
- 2013, Chair, Comprehensive Exam Committee, Division of Biostatistics
- 2008-2012, Chair, Admission Committee, Division of Biostatistics
- 2011, Chair, Student Presentation Award Committee, The Inaugural Biostatistics Research Day
- 2008, Co-Chair, Comprehensive Exam Committee, Division of Biostatistics
- 2006-2007, Member, Curriculum Committee
- 2003-2007, Member, Space Committee
- 2001-2007, Associate Member, Graduate Faculty
- 2005-2006, Chair, Comprehensive Exam Committee, Division of Biostatistics
- 2002-2003, Member, Advisory Committee

Other Local Services

- 2017, Member, Connaught International Scholarship Committee, University of Toronto
- 2016, Chair, PhD Exam Committee, Department of Biochemistry, University of Toronto
- 2016, Member, Connaught International Scholarship Committee, University of Toronto
- 2015, Host and Resource Person, Immunology Tonight Session: Biostatistics in Medical Research. Department of Immunology, Faculty of Medicine, University of Toronto.
- 2013, Chair, PhD Exam Committee, Department of Medical Biophysics, University of Toronto
- 2011, Member, OGS Fellowship Review Committee, University of Toronto
- 2008, Member, OGS Fellowship Review Committee, University of Toronto

Teaching

Graduate Courses

- STA 4515 - Multiple Hypothesis Testing and Its Applications
  Department of Statistical Sciences, University of Toronto
  Fall 2016; Winter 2016
  Role: proposed and designed the course and as the sole instructor
- CHL 5224 - Modern Statistical Genetics
  (Required course for STAGE trainees)
  Division of Biostatistics, DLSPH, University of Toronto
  Fall 2016; Winter 2016
  Role: proposed and designed the course and as the sole instructor
- STA 2080 - Fundamentals of Statistical Genetics
  Department of Statistical Sciences, University of Toronto
  Winter 2017; Fall 2015
  Role: proposed and designed the course and as the sole instructor
- CHL 7001 - Statistical Models on Complex Human Genetic Diseases
  Division of Biostatistics, DLSPH, University of Toronto
  Summer 2016; Summer 2013; Summer 2012
  Role: Guest Lecturer

- CHL 7001 - An Introduction to the Likelihood Paradigm
  Division of Biostatistics, DLSPH, University of Toronto
  Winter 2014
  Role: co-designed the course and co-instructor; 2 other co-instructors: Dr. Lisa Strug, Hospital For Sick Children and DLSPH and Dr. Laurent Briollais, Samuel Lunenfeld Research Institute and DLSPH

- CHL 5224 - Statistical Genetics
  (Required course for STAGE trainees)
  Division of Biostatistics, DLSPH, University of Toronto
  Winter 2016; Fall 2013; Winter 2013; Fall 2006; Fall 2005; Winter 2004; Winter 2003
  Role: proposed and designed the course and as the sole instructor
  Winter 2012; Winter 2011; Winter 2010; Winter 2009; Fall 2008; Fall 2007
  Role: proposed and designed the course and as the primary instructor; co-instructor: Dr. Wei Xu, Princess Margaret Hospital and DLSPH

- CHL 5210 - Categorical Data Analysis
  (Required course for Biostatistics graduate students)
  Division of Biostatistics, DLSPH, University of Toronto
  Fall 2013; Fall 2012; Fall 2011; Fall 2010; Fall 2009
  Role: designed the course and as the primary instructor; co-instructor: Dr. Laurent Briollais, Samuel Lunenfeld Research Institute and DLSPH
  Winter 2004
  Role: designed the course and as the primary instructor; co-instructor: Dr. Joseph Beyene, Hospital for Sick Children and DLSPH

- CHL 5207/5208 - Lab in Statistical Design & Analysis
  (Required course for Biostatistics graduate students)
  Division of Biostatistics, DLSPH, University of Toronto
  2011-2012; 2010-2011
  Role: Practicum Supervisor

- CHL 7001 - Statistical Methods for Genetics and Genomics
  (Required course for STAGE trainees)
  Division of Biostatistics, DLSPH, University of Toronto
  2011-2012; 2010-2011
  Role: one of the three Faculty Discussion Leaders

- STA 4315 - Computational Methods in Statistical Genetics
  Department of Statistics, Faculty of Arts and Science, University of Toronto
  Winter 2009; Winter 2006; Winter 2005
  Role: co-proposed and co-designed the course and as co-instructor; co-instructor: Prof. Radu Craiu, Department of Statistics

- CHL 7001 - Statistical Methods in Genetic Epidemiology
  Division of Biostatistics, Department of PHS, University of Toronto

Winter 2005
Role: proposed and designed the course and as the sole instructor

- STAT 946 - Statistical Genetics
(Innovative teaching via video conferencing)
Department of Statistics and Actuarial Science, University of Waterloo
Winter 2004
Role: proposed and designed the course, proposed the idea and as the sole instructor

- CHL 5250 - Special Topics in Biostatistics Division of Biostatistics, Department of PHS, University of Toronto
Winter 2003; Fall 2002
Role: designed the course and as the sole instructor

- CHL 7001 - Statistical Methods in Genetics
Division of Biostatistics, Department of PHS, University of Toronto
Winter 2002
Role: proposed and designed the course and as the sole instructor

Undergraduate Courses

- STA480 - Fundamentals of Statistical Genetics
Department of Statistical Sciences, University of Toronto
Winter 2017; Fall 2015
Role: proposed and designed the course and as the sole instructor

- STAT220 - Statistical Methods and Their Applications
Department of Statistics, University of Chicago
Spring 2000; Spring 1999; Spring 1998
Role: as the sole instructor with some contribution to the course design; mostly relying on the existing syllabus

Training of Highly Qualified Personnel (HQP)

PhD Students and Post-Doctoral Fellows (PDF)

- Lin Zhang, PhD in Statistics (06/2016-; Primary Supervisor). Statistical Genetics.

- Linda Kachuri, PhD in Epidemiology (10/2014-; Statistical Genetics Mentor for STAGE trainees). Comprehensive investigation of telomere length and its genetic determinants in cancer risk and clinical outcomes.

- Wei Deng, PhD in Statistics (10/2014-; Primary Supervisor). "X-inclusion: integrating the X chromosome in genome-wide study of variance heterogeneity to identify interaction effects".

- Emery Goossens, PhD in Statistics (08/2014-; Primary Supervisor). "Subset analysis accounting for post-selection bias".

- Bo Chen, PhD in Statistics (09/2013-; Co-Primary Supervisor). "Statistical insights on the association analysis of X chromosome: X-inactivation, heterogeneity and interaction effects".

- David Soave, PhD in Biostatistics (07/2012-11/2016; Co-Primary Supervisor). "Statistical methods for marker discovery and disease risk prediction amid unobserved genetic interactions".
  The paper on joint location-scale association test was Featured in the ''Trainee Paper Spotlight" that highlights outstanding publications and research done by trainees, by the American Society of Human Genetics' Training and Development Committee.

- Andriy Derkach, PhD in Statistics (06/2010-06/2014; Co-Primary Supervisor). "Statistical methodologies for genetic association studies with rare variants".
  Recipient of the Department of Statistics 2012-2013 Doctoral Award, given to a PhD Student for excellence in research.
  One of the three finalists for the 2012 Williams Award of the International Genetic Epidemiology Society.

- Marc Woodbury-Smith, PDF in Genetics (10/2011-08/2013; Statistical Genetics Mentor for STAGE trainees). "Genetic studies of autism spectrum disorder".

- Vanessa F. Gonçalves, PDF in Genetics (03/2011-08/2013; Statistical Genetics Mentor for STAGE trainees). Genetic studies of behavioural phenotypes.

- Elif Acar, PhD in Statistics (04/2009-11/2011; Primary Supervisor). "A generalized Kruskal-Wallis test incorporating group uncertainty with application to genetic association studies."

- Laura Faye, PhD in Biostatistics (09/2007-06/2013; Co-Primary Supervisor). "Identification of candidate causal variants and estimation of genetic associations in GWAS and post-GWAS studies."
  Winner of the 2012 Williams Award of the International Genetic Epidemiology Society.

- Yun Joo Yoo, PDF in Statistical Genetics (10/2007-06/2009; Co-Primary Supervisor). "Design and analysis of genome-wide studies of complex diseases and traits".

- Lizhen Xu, PhD in Statistics (09/2006-11/2012; Co-Supervisor). "Bayesian methods for genetic association studies."
  Recipient of the Department of Statistics 2011-2012 Doctoral Award, given to a PhD Student for excellence in research.

- Sophia Lee, PhD in Biostatistics (09/2003-11/2007; Co-Primary Supervisor). "Random forests for multi-locus quantitative trait linkage analysis".

- Longyang Wu, PDF in Statistical Genetics (02/2003-08/2006; Co-Primary Supervisor). "Statistical methods to improve the reliability of results from genome-wide studies of complex disease and quantitative traits."

- Lucia Mirea, PhD in Biostatistics (09/2005-03/2011; Co-Supervisor). "Approaches incorporating evidence for population stratification bias in genetic association analyses combining individual and family data".

- Joanna Biernacka, PhD in Biostatistics (09/2001-11/2004; Co-Primary Supervisor). "Statistical methods for studying two linked disease genes."

Research Associates/Assistants

The Master programs of both biostatistics and statistics are mostly course-based programs, so with a few exceptions, the supervision of MSc students in this section is research project-based unless specified otherwise.

- Apostolos Dimitromanolakis, Research Associate (01/2017-; Primary Supervisor). "Exploratory data analysis and data visualization".

- Ting Zhang, (thesis-based) MSc in Biostatistics (12/2016-; Primary Supervisor). "Statistical genetics".

- Scott Mastromatteo, MSc in Applied Computer Science with concentration in Data Science (10/2016-; Primary Supervisor). "Ranking enriched motifs from Chip-seq data".

- Rose Garrett, MSc in Biostatistics (09/2016-; Primary Supervisor). "Statistical methods for gene expression data of C.elegans".

- Emery Goossens, MSc in Statistics (01/2014-08/2014; Primary Supervisor). "Genetic model selection using a non-nested linear model comparison framework".

- Lauren Erdman, MSc in Biostatistics (12/2013-10/2014; Primary Supervisor). "Combining association evidence from different data resources".

- Sayem Borhan, MSc in Biostatistics (11/2012-08/2013; Primary Supervisor). "X-chromosome analysis".

- Apostolos Dimitromanolakis, Research Assistant (06/2012-03/2013; Primary Supervisor). "PREST-plus identifies pedigree errors and cryptic relatedness in the GAW18 sample using genome-wide SNP data".

- Jingxiong (Amelia) Xu, MSc in Biostatistics (01/2012-04/2012; Primary Supervisor). "Multiple hypothesis testing - robust methods incorporating prior information".

- Paul Popadiuk, MSc in Biostatistics (09/2011-04/2012; Primary Supervisor). "Interpreting the statistical evidence for association analysis of SNPs in the context of genotype uncertainty".

- Julia Taleban, Programmer (10/2010-09/2012; Co-Supervisor). "A general procedure to address the winner's curse in genetic association studies: bias-reduction in analysis of time-to-event traits".

- Haiyan Jiang, MSc in Biostatistics (10/2010-06/2011; Co-Primary Supervisor)."Genome-wide linkage and association study of APCR trait".

- Apostolos Dimitromanolakis, Research Associate (09/2010-09/2011; Co-Primary Supervisor). "Genetic studies of venous thromboembolism".

- Junming Yang, MSc in Biostatistics (02/2010-11/2010; Primary Supervisor). "Joint analysis of multiple correlated phenotypes and SNPs".

- Apostolos Dimitromanolakis, MSc in Biostatistics (11/2008-11/2010; Primary Supervisor). "Developing PREST-plus for detection of pedigree errors and cryptic relatedness using high-throughput genotype data".

- Andreea Cojocaru, Research Assistant (10/2008-09/2009; Co-Supervisor). "Association analysis for Meconium Ileus incorporating phenotype misclassification."

- Apostolos Dimitromanolakis, Programmer (10/2008-10/2009; Co-Primary Supervisor). "Implementation of bootstrap bias reduction methods for genome-wide association data."

- Theodore (Ted) Chiang, Programmer (02/2008-07/2008; Primary Supervisor). "Upgrading PREST for fast and efficient detection of pedigree errors and cryptic relatedness using high-throughput genotype data."

- Lam Opal Huang, MSc in Biostatistics (09/2007-09/2008; Primary Supervisor). "Some statistical issues in genome-wide association of common alleles with long-term diabetic complications."

- Baisong Huang, PhD in Biostatistics (05/2006-06/2007; Primary Supervisor). Genetic Analysis Workshop (GAW) 15: "The multiplicity problem in linkage analysis of gene expression data - the power of differentiating *cis* and *trans* regulators."

- Jag Rangrej, Research Assistant (01/2006-04/2007; Co-Supervisor). "Optimal selection of markers for validation from genome-wide association studies"

- Xiaoquan (William) Wen, Programmer (05/2002-11/2004; Primary Supervisor). Methods implementation as software STEPC - STatistical Explanation for Positional Cloning.

Undergraduate Students

The Dalla Lana School of Public Health does not have an undergraduate program, so the undergraduate supervision started after moving to the Department of Statistical Sciences (75% FTE), Faculty of Arts and Science in 2014.

- Rose Garrett, Undergraduate Student majored in Statistics and Ecology & Evolutionary Biology (04/2016-08/2016; Primary Supervisor). "Statistical methods for genetic studies of complex human traits ". Undergraduate Student Research Award, Natural Sciences and Engineering Research Council of Canada (NSERC).

- Xiaotong Wang, Undergraduate Student majored in statistic and immunology (04/2016-07/2016; Primary Supervisor). "A statistical analysis of association between physiological trait and human sexual orientation".

Statistical Mentoring in Weekly Interdisciplinary Lab Meetings (since 2010)

Trainees in this section are MSc and PhD level graduate students, PDFs or Research Assistants/Associates of my collaborators. I provide statistical mentoring during weekly interdisciplinary lab meetings (1.5-2 hr/week/lab) and manuscripts/conference abstracts preparation period.

- Dr. Lisa Strug's Lab - Genetic studies of Cystic Fibrosis modifier genes.
  - Bowei Xiao, MSc level Biostatistician (09/2014-)
  - Gengming He, MSc level Biostatistician (09/2014-)
  - Naim Panjwani, MSc level Biostatistician (06/2014-)
  - Jiafen Gong, PhD level Biostatistician (06/2012-)
  - Weili (Liz) Li, MSc level Biostatistics (09/2009-11/2015)
  - Melissa Miller, PDF in Genetic Epidemiology (01/2012-10/2014)
  - Theodore (Ted) Chiang, Bioinformatician (10/2007-03/2012)
  - Dong Su, MSc level Biostatistician (09/2011-08/2012)
  - Xin Li, MSc level Biostatistician (09/2010-10/2011)

- Dr. Andrew Paterson's Lab - Genetic studies of type 1 diabetic complications.
  - Sareh Keshavarzi, PDF in Genetics (09/2016-)
  - Jingjing Cao, MSc level Biostatistician (09/2016-)
  - Delnaz Roshandel PDF in Genetics (09/2014-)

- Linda Hiraki, PDF in Genetics (07/2013-)
- Mohsen S. Hosseini, PDF in Genetics (10/2010-12/2015)
- Karen M. Eny, PDF in Genetics (10/2010-10/2014)
- Hoi Wong, MSc level Biostatistician (05/2012-10/2014)
- Zhijian (Charlie) Chen, PDF in Statistical Genetics (07/2010-09/2013)

Ad-hoc Mentoring

- Ryan Ma, Undergraduate Student, President's Scholar of Excellence Award (PSEP) Mentorship Match. (11/2016-)

Committee Member

- Linda Kachuri, PhD in Epidemiology (07/2013-). "Identifying genetic determinants of telomere length and evaluating telomere length as a predictor of cancer risk and clinical outcomes in lung cancer".
- Weili (Liz) Li, PhD in Biostatistics (09/2010-11/2015). "Pure likelihood-based methods for genetic association studies".
- Merav Yarkoni-Abitbul, PhD in Dentistry (01/2008-09/2014). "Expression profiling of chronic pain genes in mice using a whole genome approach".
- Yan Lu, MSc in Dentistry (05/2005-06/2006). "Polymorphisms in Catechol-O-Methyl-transferase gene affect chronic post-mastectomy pain syndrome".
- Juan Pablo Lewinger, PhD in Statistics (09/2001-02/2004). "Family-based nonparametric tests of linkage and association".

Thesis Examination Committee Member

- Daniel Pu, PhD in Statistics, Department of Mathematics and Statistics, York University. Thesis title: "Dependency network in multivariate analysis" (External thesis examiner, May 2009).