Daniel C. Koboldt

The Institute for Genomic Medicine at Nationwide Children's Hospital • 700 Children's Drive, Columbus, OH 43205
Phone: (614) 722-0520 • E-Mail: Daniel.Koboldt@nationwidechildrens.org

Summary

Dan Koboldt is a human geneticist and bioinformatician with expertise in next-generation sequencing (NGS) data analysis. His group at the Institute for Genomic Medicine at Nationwide Children's Hospital analyzes exome and whole-genome sequence data to uncover the genetic basis of rare diseases and pediatric cancer.

He is the lead developer of VarScan, an open-source tool for germline and somatic variant calling in NGS data, and MendelScan, an open-source tool for prioritizing variants in family-based sequencing studies.

He is also the author of MassGenomics (http://massgenomics.org), a blog about genetics, genomics, and DNA sequencing in the post-genome era; KidsGenomics (http://kidsgenomics.org), a blog about the genomics of pediatric disease; as well as three novels published by HarperCollins.

Education

M.S., Biology	Washington University in St. Louis, 2007
B.S., Computer Science	University of Missouri, Columbia, 2001
B.A., French Language	University of Missouri, Columbia, 2001

Experience

Principal Investigator II, Nationwide Children's Hospital	2016-present
Research Assistant Professor of Pediatrics, The Ohio State University	2016-present
Manager, Human Genetics Group, The Genome Institute at Washington University Supervisor: Rick Wilson	2012-2016
Staff Scientist, The Genome Institute at Washington University Supervisor: Dave Larson	2011-2012
Programmer/Analyst, Medical Genomics Group at Washington University Supervisor: Li Ding	2007-2011
Bioinformatics Programmer, Department of Genetics at Washington University Supervisor: Ray Miller	2003-2007

Professional Memberships

- American Society of Human Genetics (ASHG), 2012-present
- Editorial Board, *Front Line Genomics Magazine*, 2016-present

Technical Skills

- **Programming**: Perl, Java, PHP, R, Komodo IDE, Eclipse IDE, version control (git & svn)
- **Operating Systems**: Linux/UNIX, Mac OSX
- **Bioinformatics Software**: BWA, SAMtools, Picard, VCFtools, GATK, PLINK, BEAGLE
- **Web Development**: HTML, CSS, Wordpress, phpList, phpMyAdmin, Google Analytics
- **Desktop Software**: Microsoft Office, EndNote, Adobe Photoshop, Macromedia Dreamweaver

Research Expertise

- **Next-generation sequencing (NGS) data:** Whole-genome, exome, and targeted resequencing datasets for case-control studies, family studies, tumor-normal pairs, and clinical cases.
- **NGS bioinformatics:** Alignment, de-duplication, SAM/BAM manipulation, coverage analysis, variant calling, somatic mutation detection, annotation with external databases (e.g. Ensembl, dbSNP, VEP)
- **Case-control studies:** principal components analysis (MDS), and analysis of rare variants
- **Rare/Mendelian disease studies:** variant annotation and prioritization, NGS-based linkage mapping, clinical variant interpretation/reporting.
- **Cancer genomics:** mutation detection, significantly mutated gene (SMG) testing, mutational proximity and recurrence, germline susceptibility, copy number and LOH analysis.

Editorial and Peer Review Service

- *Frontiers in Genetics* (Topic Editor)
- *American Journal of Human Genetics*
- *Bioinformatics*
- *Genome Biology*
- *Genome Medicine*
- *Nucleic Acids Research*
- *Human Mutation*
- *Nature Biotechnology*
- *PLoS Genetics*
- *PLoS Computational Biology*

Community Outreach

- National DNA Day Ambassador Program 2009-present
- ASHG DNA Day Essay Contest Judge 2014-present
- Author, MassGenomics (http://massgenomics.org) 2008-2018
- Author, KidsGenomics (http://kidsgenomics.org) 2018-present

Presentations at Professional Meetings and Workshops

2013 ASHG Annual Meeting, Platform Speaker, Boston, MA

2014 ASHG Annual Meeting, Platform Speaker, San Diego, CA

2015 ASHG Annual Meeting, Invited Session Proposer/Moderator, Baltimore, MD

2015 EMBO Workshop: Modern DNA concepts, Invited Speaker, Evry, France

2015 Molecular and Cellular Biosciences Retreat, Keynote Speaker, Portland, OR

2016 Alzheimer's Disease Sequencing Project Workshop, Invited Participant, Washington, DC.

2016 ASHG Annual Meeting, Invited Session Proposer/Moderator, Vancouver, BC, Canada

2017 ASHG Annual Meeting, Invited Session Proposer/Moderator, Orlando, FL

2017 AMP Annual Meeting, Invited Speaker, Salt Lake City

Peer-Reviewed Publications

1. Miller RD, Phillips MS, Jo I, Donaldson MA, Studebaker JF, *et al*: High-density single-nucleotide polymorphism maps of the human genome. *Genomics* 2005, 86(2):117-126.
2. The International HapMap Consortium: A haplotype map of the human genome. *Nature* 2005, 437(7063):1299-1320.
3. **Koboldt DC**, Miller RD, Kwok PY: Distribution of human SNPs and its effect on high-throughput genotyping. *Hum Mutat* 2006, 27(3):249-254.
4. Hillier LW, Miller RD, Baird SE, Chinwalla A, Fulton LA, **Koboldt DC**, Waterston RH: Comparison of *C. elegans* and *C. briggsae* genome sequences reveals extensive conservation of chromosome organization and synteny. *PLoS Biol* 2007, 5(7):e167.
5. Stanley SL, Jr., Frey SE, Taillon-Miller P, Guo J, Miller RD, **Koboldt DC**, et al.: The immunogenetics of smallpox vaccination. *J Infect Dis* 2007, 196(2):212-219.
6. The International HapMap Consortium: A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007, 449(7164):851-861.
7. Sabeti PC, Varilly P, Fry B, *et al*: Genome-wide detection and characterization of positive selection in human populations. *Nature* 2007, 449(7164):913-918.
8. Ley TJ, Mardis ER, Ding L, *et al*: DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature* 2008, 456(7218):66-72.
9. The Cancer Genome Atlas Consortium: Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 2008, 455(7216):1061-1068.
10. Ding L, Getz G, Wheeler DA, Mardis ER, *et al*: Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* 2008, 455(7216):1069-1075.
11. Zhang Q, Ding L, Larson DE, **Koboldt DC**, et al: CMDS: a population-based method for identifying recurrent DNA copy number aberrations in cancer from high-resolution data. *Bioinformatics* 2009, 26(4):464-469.
12. **Koboldt DC**, Chen K, Wylie T, et al.: VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics* 2009, 25(17):2283-2285.

13. Mardis ER, Ding L, Dooling DJ, Larson DE, McLellan MD, Chen K, **Koboldt DC**, *et al*: Recurring mutations found by sequencing an acute myeloid leukemia genome. *N Engl J Med* 2009, 361(11):1058-1066.
14. **Koboldt DC**, Staisch J, Thillainathan B, Haines K, Baird SE, Chamberlin HM, *et al.*: A toolkit for rapid gene mapping in the nematode *Caenorhabditis briggsae*. *BMC Genomics* 2010, 11:236.
15. Voora D, **Koboldt DC**, *et al.*: A polymorphism in the VKORC1 regulator calumenin predicts higher warfarin dose requirements in African Americans. *Clin Pharma Ther* 2010, 87(4):445-451.
16. Ding L, Ellis MJ, Li S, *et al*: Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* 2010, 464(7291):999-1005.
17. Ramsingh G*****, **Koboldt DC***, Trissal M, *et al.*: Complete characterization of the microRNAome in a patient with acute myeloid leukemia. *Blood* 2010, 116(24):5316-5326.
18. Ley TJ, Ding L, Walter MJ, *et al.*: DNMT3A mutations in acute myeloid leukemia. *N Engl J Med* 2010, 363(25):2424-2433.
19. Bowne SJ*, Sullivan LS*****, **Koboldt DC***, *et al*.: Identification of disease-causing mutations in autosomal dominant retinitis pigmentosa (adRP) using next-generation DNA sequencing. *Invest Ophthalmol Vis Sci* 2010, 52(1):494-503.
20. Fehniger TA, Wylie T, Germino E, *et al.*: Next-generation sequencing identifies the natural killer cell microRNA transcriptome. *Genome Res* 2010, 20(11):1590-1604.
21. The 1000 Genomes Project Consortium: A map of human genome variation from population-scale sequencing. *Nature* 2010, 467(7319):1061-1073.
22. Sudmant PH, Kitzman JO, Antonacci F, Alkan C, Malig M, Tsalenko A, Sampas N, Bruhn L, Shendure J, 1000 Genomes Project, & Eichler EE (2010). Diversity of human copy number variation and multicopy genes. *Science* (New York, N.Y.), 330 (6004), 641-6 PMID: 21030649
23. Marth GT, Yu F, Indap AR, Garimella K, *et al*: The functional spectrum of low-frequency coding variation. *Genome Biol* 2011, 12(9):R84.
24. Wartman LD, Larson DE, *et al*: Sequencing a mouse acute promyelocytic leukemia genome reveals genetic events relevant for disease progression. *J Clin Invest* 2011, 121(4):1445-1455.
25. Link DC, Schuettpelz LG, Shen D, *et al*: Identification of a novel TP53 cancer susceptibility mutation through whole-genome sequencing of a patient with therapy-related AML. *JAMA* 2011, 305(15):1568-1576.
26. Larson DE, Harris CC, Chen K, **Koboldt DC**, *et al*.: SomaticSniper: identification of somatic point mutations in whole genome sequencing data. *Bioinformatics* 2011, 28(3):311-317.
27. Graubert TA, Shen D, Ding L, Okeyo-Owuor T, Lunn CL, Shao J, Krysiak K, Harris CC, **Koboldt DC**, *et al*: Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. *Nat Genet* 2011, 44(1):53-57.

28. Ross JA, **Koboldt DC**, Staisch JE, Chamberlin HM, Gupta BP, Miller RD, Baird SE, Haag ES: *Caenorhabditis briggsae* recombinant inbred line genotypes reveal inter-strain incompatibility and the evolution of recombination. *PLoS Genet* 2011, 7(7):e1002174.
29. The Cancer Genome Atlas Research Network: Integrated genomic analyses of ovarian carcinoma. *Nature* 2011, 474(7353):609-615.
30. Bowne SJ, Humphries MM, Sullivan LS, *et al.*: A dominant mutation in RPE65 identified by whole-exome sequencing causes retinitis pigmentosa with choroidal involvement. *Eur J Hum Genet* 2011, 19(10):1074-1081.
31. Dees ND, Zhang Q, Kandoth C, Wendl MC, Schierding W, **Koboldt DC**, *et al.*: MuSiC: identifying mutational significance in cancer genomes. *Genome Res* 2012, 22(8):1589-1598.
32. Bose R, Kavuri SM, Searleman AC, Shen W, Shen D, **Koboldt DC**, *et al.*: Activating HER2 mutations in HER2 gene amplification negative breast cancer. *Cancer Disc* 2012, 3(2):224-237.
33. Matsushita H, Vesely MD, **Koboldt DC**, *et al.*: Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. *Nature* 2012, 482(7385):400-404.
34. Ding L, Ley TJ, Larson DE, Miller CA, **Koboldt DC**, *et al.*: Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* 2012, 481(7382):506-510.
35. Welch JS, Ley TJ, Link DC, Miller CA, Larson DE, **Koboldt DC**, *et al*.: The origin and evolution of mutations in acute myeloid leukemia. *Cell* 2012, 150(2):264-278.
36. Cancer Genome Atlas Network (2012). Comprehensive molecular characterization of human colon and rectal cancer. *Nature*, 487 (7407), 330-7 PMID: 22810696
37. Ellis MJ, Ding L, Shen D, *et al*.: Whole-genome analysis informs breast cancer response to aromatase inhibition. *Nature* 2012, 486(7403):353-360.
38. **Koboldt DC**, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L, Wilson RK: VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res 2012*, 22(3):568-576.
39. Walter MJ, Shen D, Ding L, Shao J, **Koboldt DC**, Chen K, Larson DE, et al: Clonal architecture of secondary acute myeloid leukemia. *N Engl J Med* 2012, 366(12):1090-1098.
40. The 1000 Genomes Project Consortium: An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012, 491(7422):56-65.
41. The Cancer Genome Atlas Network: Comprehensive molecular portraits of human breast tumours. *Nature* 2012, 490(7418):61-70.
42. The Cancer Genome Atlas Research Network: Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med* 2013, 368(22):2059-2074.
43. Zhan X, Larson DE, Wang C, **Koboldt DC**, Sergeev YV, Fulton RS, Fulton LL, Fronick CC, Branham KE, et al. (2013). Identification of a rare coding variant in complement 3 associated with age-related macular degeneration. *Nature Genetics*, 45 (11), 1375-9 PMID: 24036949

44. Service SK, Teslovich TM, Fuchsberger C, Ramensky V, Yajnik P, **Koboldt DC**, Larson DE, Zhang Q, *et al.*: Re-sequencing expands our understanding of the phenotypic impact of variants at GWAS loci. *PLoS Genetics* 2014 Jan 30;10(1):e1004147. PMID: 24497850

45. Kanchi KL, Johnson KJ, Lu C, McLellan MD, Leiserson MD, Wendl MC, Zhang Q, **Koboldt DC**, *et al.*: Integrated analysis of germline and somatic variants in ovarian cancer. 2014. *Nature Comm*. 5:3156.

46. **Koboldt DC**, Larson DE, Sullivan LS, Bowne SJ, Steinberg KM, *et al.*: Exome-based mapping and variant prioritization for inherited Mendelian disorders. *Am J Hum Genet*. 2014 Mar 6;94(3):373-84. doi: 10.1016/j.ajhg.2014.01.016. PMID: 24560519

47. Colonna V, Ayub Q, Chen Y, Pagani L, Luisi P, Pybus M, Garrison E, Xue Y, Tyler-Smith C, 1000 Genomes Project Consortium, et al. (2014). Human genomic regions with exceptionally high levels of population differentiation identified from 911 whole-genome sequences. *Genome Biology*, 15 (6) PMID: 24980144

48. Sullivan LS, Koboldt DC, Bowne SJ, Lang S, Blanton SH, Cadena EL, Avery CE, Lewis RA, Webb-Jones K, Wheaton DK, Birch DG, Coussa RG, Ren H, Lopez I, Chakarova CF, Koenekoop R, Garcia CA, Fulton RS, Wilson RK, Weinstock GM, & Daiger SP (2014). A dominant mutation in hexokinase 1 (HK1) causes retinitis pigmentosa. *Investigative Ophthalmology & Visual Science* PMID: 25190649

49. Zhang Q, Wang L, Koboldt D, Boreki IB, & Province MA (2014). Adjusting family relatedness in data-driven burden test of rare variants. *Genetic epidemiology* PMID: 25169066

50. Montague MJ, Li G, Gandolfi B, Khan R, Aken BL, Searle SM, Minx P, Hillier LW, Koboldt DC, et al. (2014). Comparative analysis of the domestic cat genome reveals genetic signatures underlying feline biology and domestication. *PNAS* USA. PMID: 25385592

51. Elizabeth J. Leslie, Margaret A. Taub, Huan Liu, Karyn Meltz Steinberg, Daniel C. Koboldt, et al. Identification of functional variants for cleft lip with or without cleft palate in or near PAX7, FGFR2, and NOG by targeted sequencing of GWAS loci. *Am J Hum Genet*. 2015 Mar 5;96(3):397-411. PMID: 25704602

52. Steinberg KM, Yu B, Koboldt DC, Mardis ER, Pamphlett R. Exome sequencing of case-unaffected-parents trios reveals recessive and de novo genetic variants in sporadic ALS. *Sci Rep*. 2015 Mar 16;5:9124. PMID: 25773295

53. Meltz Steinberg K, Nicholas TJ, Koboldt DC, Yu B, Mardis E, Pamphlett R. Whole genome analyses reveal no pathogenetic single nucleotide or structural differences between monozygotic twins discordant for amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Frontotemporal Degener*. 2015 May 11:1-8. PMID: 25960086

54. Malachi Griffith, Obi L. Griffith, Scott M. Smith, et al. Genome Modeling System: A Knowledge Management Platform for Genomics. *PLoS Comput Biol*. 2015 Jul 9;11(7):e1004274. PMID: 26158448

55. Li Tai Fang, Pegah Tootoonchi Afshar, Aparna Chhibber, et al. An ensemble approach to accurately detect somatic mutations using SomaticSeq. *Genome Biology*. 2015 Sep 17;16(1):197. doi: 10.1186/s13059-015-0758-2. PMID: 26381235

56. Leslie EJ, Koboldt DC, Kang CJ, Ma L, Hecht JT, Wehby GL, Christensen K, Czeizel AE, Deleyiannis FW, Fulton RS, Wilson RK, Beaty TH, Schutte BC, Murray JC, Marazita ML. IRF6 mutation screening in nonsyndromic orofacial clefting: analysis of 1521 families. *Clin Genet.* 2015 Sep 8. doi: 10.1111/cge.12675. PMID: 26346622

57. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature*. 2015 Oct 1;526(7571):68-74. doi: 10.1038/nature15393. PMID: 26432245.

58. Jane Churpek, Khateriaa Pyrtel, Krishna Kanchi, et al. Genomic analysis of germ line and somatic variants in familial myelodysplasia/acute myeloid leukemia. *Blood*. 2015 Nov 26;126(22):2484-90. doi: 10.1182/blood-2015-04-641100. PMID: 26492932.

59. Lu C, Xie M, Wendl MC, Wang J, McLellan MD, Leiserson MD, Huang KL, Wyczalkowski MA, Jayasinghe R, Banerjee T, Ning J, Tripathi P, Zhang Q, Niu B, Ye K, Schmidt HK, Fulton RS, McMichael JF, Batra P, Kandoth C, Bharadwaj M, Koboldt DC, et al. Patterns and functional implications of rare germline variants across 12 cancer types. *Nat Commun*. 2015 Dec 22;6:10086. doi: 10.1038/ncomms10086. PMID: 26689913

60. Troester MA, Hoadley KA, D'Arcy M, Cherniack AD, Stewart C, Koboldt DC, et al. DNA defects, epigenetics, and gene expression in cancer-adjacent breast: a study from The Cancer Genome Atlas. *NPJ Breast Cancer* (2016) 2, 16007. PMID: 28721375

61. Pfefferle AD, Agrawal YN, Koboldt DC, Kanchi KL, Herschkowitz JI, Mardis ER, et al. Genomic profiling of murine mammary tumors identifies potential personalized drug targets for p53 deficient mammary cancers. *Disease Models & Mechanisms*. 2016 Jul 1;9(7):749-57. doi: 10.1242/dmm.025239. PMID: 27149990.

62. Koboldt DC, Kanchi KL, Gui B, et al. Rare Variation in TET2 Is Associated with Clinically Relevant Prostate Carcinoma in African-Americans. *Cancer Epidemiol Biomarkers Prev*. 2016 Aug 2. pii: cebp.0373.2016. PMID: 27486019

63. Bowne SJ, Sullivan LS, Wheaton DK, Locke KG, Jones KD, Koboldt DC, et al. North Carolina macular dystrophy (MCDR1) caused by a novel tandem duplication of the PRDM13 gene. *Mol Vis*. 2016 Oct 17;22:1239-1247. PMID: 27777503

64. Xiao Y, Taub MA, Ruczinski I, et al. Evidence for SNP-SNP interaction identified through targeted sequencing of cleft case-parent trios. *Genet Epidemiol.* 2016 Dec 26. doi: 10.1002/gepi.22023. PMID: 28019042

65. Sullivan LS, Bowne SJ, Koboldt DC, et al. A Novel Dominant Mutation in SAG, the Arrestin-1 Gene, Is a Common Cause of Retinitis Pigmentosa in Hispanic Families in the Southwestern United States. *Invest Ophthalmol Vis Sci*. 2017 May 1;58(5):2774-2784. PMID: 28549094

66. Mashl RJ, Scott AD, Huang KL, Wyczalkowski MA, Yoon CJ, Niu B, DeNardo E, Yellapantula VD, Handsaker RE, Chen K, Koboldt DC, Ye K, Fenyö D, Raphael BJ, Wendl MC, Ding L.

GenomeVIP: a cloud platform for genomic variant discovery and interpretation. *Genome Res*. 2017 Aug;27(8):1450-1459. doi: 10.1101/gr.211656.116. Epub 2017 May 18. PMID: 28522612

67. Beecham GW, Bis JC, Martin ER, Choi SH, DeStefano AL, van Duijn CM, Fornage M, Gabriel SB, Koboldt DC, et al. The Alzheimer's Disease Sequencing Project: Study design and sample selection. *Neurol Genet*. 2017 Oct 13;3(5):e194. PMID: 29184913
68. Koboldt DC, Mihalic Mosher T, Kelly BJ, Sites E, Bartholomew D, Hickey SE, McBride K, Wilson RK, White P. A de novo nonsense mutation in ASXL3 shared by siblings with Bainbridge-Ropers syndrome. *Cold Spring Harb Mol Case Stud*. 2018 Jan 5. PMID: 29305346
69. Miller KE, Kelly B, Fitch J, Ross N, Avenarius MR, Varga E, Koboldt DC, et al. Genome sequencing identifies somatic BRAF duplication c.1794_1796dupTAC;p.Thr599dup in pediatric patient with low-grade ganglioglioma. *Cold Spring Harb Mol Case Stud*. 2018 Feb 6. PMID: 29434027
70. Kalscheuer V, Frints S, Ozanturk A, et al. Pathogenic variants in E3 ubiquitin ligase RLIM/RNF12 lead to a syndromic X-linked intellectual disability and behavior disorder. *Molecular Psychiatry*. 2018 (in press)

Book Chapters and Review Articles

1. **Koboldt DC**, Miller RD: Identification of polymorphic markers for genetic mapping. *Genomics: Essential Methods* 2010.
2. Ding L, Wendl MC, **Koboldt DC**, Mardis ER: Analysis of next-generation genomic data in cancer: accomplishments and challenges. *Hum Mol Genet 2010*, 19(R2):R188-196.
3. **Koboldt DC**, Ding L, Mardis ER, Wilson RK: Challenges of sequencing human genomes. *Brief Bioinform* 2010, 11(5):484-498.
4. **Koboldt DC**, Larson DE, Chen K, Ding L, Wilson RK: Massively parallel sequencing approaches for characterization of structural variation. *Methods Mol Biol* 2012, 838:369-384.
5. **Koboldt DC**, Steinberg KM, Larson DE, Wilson RK, & Mardis ER (2013). The next-generation sequencing revolution and its impact on genomics. *Cell*, 155 (1), 27-38 PMID: 24074859
6. Daiger SP, Bowne SJ, Sullivan LS, Blanton SH, Weinstock GM, **Koboldt DC**, et al (2014). Application of next-generation sequencing to identify genes and mutations causing autosomal dominant retinitis pigmentosa (adRP). Advances in experimental medicine and biology, 801, 123-9 PMID: 24664689
7. Del-Aguila JL, **Koboldt DC**, Black K, Chasse R, Norton J, Wilson RK, Cruchaga C. Alzheimer's disease: rare variants with large effect sizes. Curr Opin Genet Dev. 2015 Aug 22;33:49-55. doi: 10.1016/j.gde.2015.07.008. PMID: 26311074