HUGEN 2029: Introduction to Gene Mapping

3 credits / Fall Term 2021 / Mondays and Wednesdays, 9:30-10:55 am / 2121C Public Health

- Brenda Diergaarde, PhD; Associate Professor of Human Genetics, Pitt Public Health, and UPMC Hillman Cancer Center

Contact information:
Brenda Diergaarde, Ph.D.
UPCI Cancer Pavilion, Suite 4C
412-623-5891 (office)
bbd3@pitt.edu
Office Hours: By appointment

Faculty Availability: I welcome your questions and suggestions. Please feel free to set up an appointment. Also, if you are having any problems with the course, please contact me as soon as possible. E-mail is a good way to reach me. However, since I get many e-mails, please use an informative subject line starting with "HUGEN 2029".

Course Description: This course presents a literature-based approach to understanding and interpreting results from gene mapping papers in the field of human genetics. Traditional and state-of-the-art genetic mapping methodologies will be explored.

COVID: If you are required to isolate or quarantine, become sick, or are unable to come to class, contact me as soon as possible to discuss arrangements.

Course Overview and Learning Objectives:

This course covers the (currently) most commonly-used technologies and methodologies for discovering and exploring genotype-phenotype associations. Each methodology will be covered in one or two didactic class sessions, and then participants will read, critique, and present papers that apply the methodology.

At the end of this course, participants should be able to:

- Describe the mathematical and scientific underpinnings of each methodology
- Discuss how the choice of study design influences the choice of methodology (and vice versa)
- Discuss the strengths and limitations of each methodology
- Evaluate gene mapping results in the current literature
- Critique the study design and methodology choices in published gene-mapping studies

Texts/assigned materials:

Participants will need to read and be prepared to discuss assigned materials that will be posted for personal use on Canvas. There is no required textbook for this course.

Exams and Assignments:

- Exams: There will be two (take-home) exams, a mid-term exam and a final exam, to assess students' ability to understand and critique gene mapping methods. The format will be open-ended questions.

- Student presentations: Each student will review, present and discuss a (recently) published paper in class. Presentations should be short, approximately 20 minutes, and primarily be an introduction to the topic and briefly highlight the paper. The student will subsequently lead a discussion of the paper. The goal of these presentations is to learn to critically review papers, evaluate the strengths and weaknesses of the papers, and to gain experience in public speaking. Everyone is expected to have read the selected papers in advance and come prepared to discuss.

- Discussions: In addition to the student presentations, the class as a whole will review and discuss several papers. Everyone is expected to have read the selected papers in advance and come prepared to discuss.

Student Performance Evaluation:

All course requirements must be completed to receive credit for the course. Evaluation will be based on the following components:

- Attendance and Quality of Contribution to Discussion (25% of final grade)
Attendance, active participation in class discussions, and evidence of being prepared for class (including having read the assigned readings and completion of assignments) are expected. While cell phones and laptops/tablets may be used to access slides or assigned readings, take notes, etc., please do not use them during class time for non-class purposes. If you will miss a class, please let me know in advance (bbd3@pitt.edu).

- Student Presentations (25% of final grade)

- Mid-term Exam (25% of final grade)
Take-home mid-term exam will be posted on Canvas on October 20. The mid-term exam is due on October 25 (by midnight; please submit to: bbd3@pitt.edu).

- Final Exam (25% of final grade)
Take-home final exam will be posted on CANVAS on December 8. The final exam is due on December 15 (by midnight; please submit to: bbd3@pitt.edu).

Grade option:	**Grading scale:**		
Letter grade	97 - 100% A+	77 – 79.9% C+	< 60% F
	93 – 96.9% A	73 – 76.9% C	
	90 – 92.9% A-	70 – 72.9% C-	
	87 – 89.9% B+	67 – 69.9% D+	
	83 – 86.9% B	63 – 66.9% D	
	80 – 82.9% B-	60 – 62.9% D-	

Canvas:

The University's Canvas will be used to post announcements, assignments, and readings for personal use.

Course Policies:

The Internet generally should not be accessed during class, except to access course slides or assigned readings, help resolve a disputed point in discussion or otherwise enhance discussion. Students should check their email regularly to ensure that they receive relevant communication regarding the course.

Students should familiarize themselves also with the following policies:

- Academic Integrity Policy:
All individuals (students, faculty, post-doctoral researchers, and staff) at Pitt Public Health abide by the University's policy on academic integrity. In accordance with this policy, the school maintains an outline of the procedural sequence of events to occur when violations of academic integrity are brought to the attention of administrative leaders. The full policy is available in the Academic Handbook.

All students are expected to adhere to the school's standards of academic honesty. Cheating/plagiarism will not be tolerated. The Graduate School of Public Health's policy on academic integrity, which is based on the University policy, is available online in the Pitt Public Health Academic Handbook www.publichealth.pitt.edu/home/academics/academic-requirements. The policy includes obligations for faculty and students, procedures for adjudicating violations, and other critical information. Please take the time to read this policy.

Students should be especially mindful of guidelines on academic integrity and take care to avoid plagiarizing the work - including the ideas or words - of their colleagues (fellow course participants) or other authors. Students are encouraged to discuss their ideas and work together; however, a citation to a fellow student should be provided when appropriate.

- Diversity and Academic Civility Statement:
In this course, students, faculty and guests represent a diversity of individual perspectives, backgrounds, and experiences, which enriches our classes. We urge all to be respectful of others.

The University of Pittsburgh Graduate School of Public Health considers the diversity of its students, faculty, and staff to be a strength and critical to its educational mission. Pitt Public Health is committed to creating and fostering inclusive learning environments that value human dignity and equity and promote social justice. Every member of our community is expected to be respectful of the individual perspectives, experiences, behaviors, worldviews, and backgrounds of others. While intellectual disagreement may be constructive, no derogatory statements, or demeaning or discriminatory behavior will be permitted.

If you feel uncomfortable or would like to discuss a situation, please contact any of the following:
• the course instructor (bbd3@pitt.edu or see phone number above);
• the Pitt Public Health Associate Dean responsible for diversity and inclusion (Dr. Tiffany Gary-Webb: tgary@pitt.edu // 412-624-3131);
• the University's Office of Diversity and Inclusion at 412-648-7860 or https://www.diversity.pitt.edu/make-report/report-form (anonymous reporting form)

- Accommodation for Students with Disabilities:
If you have any disability for which you may require accommodation, you are encouraged to notify both your instructor (bbd3@pitt.edu) and the Office of Disability Resources and Services (DRS), 140 William Pitt Union (Voice or TTD 412-648-7890), http://www.studentaffairs.pitt.edu/drs/, drsrecep@pitt.edu, as early as possible in the term.

- Copyright Notice:
Course materials may be protected by copyright. United States copyright law, 17 USC section 101, et seq., in addition to University policy and procedures, prohibit unauthorized duplication or retransmission of course materials. See Library of Congress Copyright Office and the University Copyright Policy.

- Classroom Recording:
To ensure the free and open discussion of ideas, students may not record classroom lectures, discussion and/or activities without the advance permission of the instructor, and any such recording properly approved in advance can be used solely for the student's own private use or for all students enrolled in this class only but may not be further copied, distributed, published, or otherwise used for any other purpose without the express written consent of the course instructors. Any student who records a class session must provide a copy of the recording to the instructors if requested to do so.

- Sexual Misconduct, Required Reporting, and Title IX:
The University is committed to combatting sexual misconduct. As a result, you should know that University faculty and staff members are required to report any instances of sexual misconduct, including harassment and sexual violence, to the University's Title IX office so that the victim may be provided appropriate resources and support options. What this means is that as your professor, I am required to report any incidents of sexual misconduct that are directly reported to me, or of which I am somehow made aware.

There are two important exceptions to this requirement about which you should be aware:

(1) A list of the designated University employees who, as counselors and medical professionals, do not have this reporting responsibility and can maintain confidentiality, can be found here: https://www.diversity.pitt.edu/civil-rights-title-ix/make-report/report-form
(2) An important exception to the reporting requirement exists for academic work. Disclosures about sexual misconduct that are shared as part of an academic project, classroom discussion, or course assignment, are not required to be disclosed to the University's Title IX office

If you are the victim of sexual misconduct, Pitt encourages you to reach out to these resources:

• Title IX Office: 412-648-7860 (https://www.titleix.pitt.edu/)
• SHARE @ the University Counseling Center: 412-648-7930 (8:30 A.M. TO 5 P.M. M-F) and 412-648-7856 (AFTER BUSINESS HOURS)
If you have a safety concern, please contact the University of Pittsburgh Police, 412-624-2121.

HUGEN 2029: Introduction to Gene Mapping
Fall Term 2021 Schedule (version: 8/29/21)
(Mondays and Wednesdays, 9:30-10:55 am / 2121C Public Health)

Date	Topics	
Monday August 30	Course introduction	
Wednesday September 1	cancelled	
Monday September 6	**LABOR DAY HOLIDAY – No class (university closed)**	
Wednesday September 8	Study design, study populations, diversity	
September 10-12	**Human Genetics retreat**	
Monday September 13	cancelled	
Wednesday September 15	Candidate gene association studies / discussion	
Monday September 20	Genome-wide association studies (GWAS)	
Wednesday September 22	GWAS – imputation and combining datasets	
Monday September 27	Polygenic risk scores (and what about the environment?)	
Wednesday September 29	GWAS discussion	
Monday October 4	GWAS follow-up; eQTL analysis, gene set analysis	
Wednesday October 6	GWAS follow-up discussion	
Monday October 11	Sequencing - biochemistry, rare variants, and cancer	
Wednesday October 13	Sequencing - association studies	
Monday October 18	Sequencing discussion	
October 18-22	**ASHG virtual meeting: https://www.ashg.org/meetings/2021meeting/**	
Wednesday October 20	**No class – mid-term exam posted on Canvas**	
Monday October 25	**No class – mid-term exam due (by midnight)**	
Wednesday October 27	Family-based designs, linkage analysis	
Monday November 1	Family-based studies discussion	
Wednesday November 3	Epigenetics, methylation	
Monday November 8	Methylation studies discussion	
Wednesday November 10	Expression data, transcriptome-wide association studies, RNA-seq, single cell RNA-seq	
Monday November 15	Expression discussion	
Wednesday November 17	Copy number variants (CNVs) / role of CNVs in disease	
Monday November 22	**THANKSGIVING HOLIDAY – No class**	
Wednesday November 24	**THANKSGIVING HOLIDAY – No class**	
Monday November 29	Post-transcriptional regulation	
Wednesday December 1	Other genomes: the microbiome and mitochondria	
Monday December 6	Multi-omics approaches	
Wednesday December 8	Mendelian randomization	
Monday December 13	Topic tbd – **final exam posted on Canvas**	
Wednesday December 15	**No class – final exam due (by midnight)**	

General Outline for Student Presentations – HUGEN 2029

- Paper title, journal, authors (1 slide)

- Introduction: brief description of the central question addressed in the paper and its significance (1-2 slides).
  What is the goal of the described study? What is the hypothesis?

- Previous studies: summarize previous studies relevant to the paper (1-2 slides).

- Study design: give overview of study design used.
  Anything in previous studies or background that made going with this design obvious or that argues against this type of design? What are the strengths and limitations of this design? Would you have used this design? Discuss why or why not?

- Methods: describe what methods are used in the paper.
  What are the strengths and limitations of the methods used? Do the methods employed make sense given the study design? Would you have used these/similar methods? Discuss why or why not?

- Results: show and discuss the results.
  Do the results make sense? Do you agree with the authors' interpretation of the results? Discuss why or why not.

- Summary/Conclusions: briefly reiterate key findings and the strengths and limitations of the study in particular in relation to study design and methods employed.
  Did the authors use appropriate design and methods? Do you agree with the interpretation of the results? Discuss why or why not.

Rubric for Student Presentations – HUGEN 2029

Category	Elements	Weight
Knowledge and explanation of topic, discussion:	• Conveys understanding • Presents the essential information • Accurate description of methodology, study design, goals and hypotheses, etc. • Good discussion of strengths and limitations	**75**
Overall organization of section/talk	• Content introduced in logical, easy-to-follow sequence • Main points emphasized, repeated • Use of transition statements	**10**
Overall effectiveness of slides (text and visuals) and delivery	• Good balance of text & figures/tables • Text/figures/tables large enough to be seen • Not too many or too few slides • Confident, enthusiastic delivery • Eye contact • Get to main points quickly	**15**

Background/suggested readings (all posted on CANVAS)

Study design, study populations, diversity:

- Buffalo gave us spicy wings and the 'book of life.' Here's why that's undermining personalized medicine (STAT news, March 11, 2019): https://www.statnews.com/2019/03/11/human-reference-genome-shortcomings/
- Hindorff et al. Prioritizing diversity in human genomics research. *Nat Rev Genet*. 2018 Mar;19(3):175-185. https://pubmed.ncbi.nlm.nih.gov/29151588/
- Diversity matters. *Nat Rev Genet*. 2019 August (uploaded)
- Gurdasani, D. et al. Genomics of disease risk in globally diverse populations. *Nat Rev Genet*. 2019 August (uploaded)
- Ballouz et al. Is it time to change the reference genome? (uploaded)
- Too many scientists still say Caucasian. Racist ideas of categories for human identity continue to warp research and medicine. Alice B. Popejoy Too many scientists still say Caucasian (nature.com)

Candidate gene association studies:

- Paul R Burton, Martin D Tobin, John L Hopper. Key concepts in genetic epidemiology. *Lancet* 2005; 366: 941–51
- Cordell and Clayton. Genetic association studies. *Lancet* 2005; 366:1121-31
- Hattersley and McCarthy. What makes a good genetic association study? *Lancet* 2005; 366: 1115-23
- Jorgensen et al. Hypothesis-driven candidate gene association studies: practical design and analytical considerations. *Am J Epidemiol* 2009;170:986–993

Genome-wide association studies:

- Manolio T.A. Genomewide Association Studies and Assessment of the Risk of Disease.*N Engl J Med* 2010;363:166-76.
- Manolio T.A. Bringing genome-wide association findings into clinical use. *Nat Rev Genet.* 2013 Aug;14(8):549-58. doi: 10.1038/nrg3523.
-Sud A, Kinnersley B, Houlston RS. Genome-wide association studies of cancer: current insights and future perspectives. *Nat Rev Cancer*. 2017 Nov;17(11):692-704.
- Making the case for more inclusive GWAS. *Nat Rev Genetics* 20, 500–501 (2019)
- Wojcik, G. L. et al. Genetic analyses of diverse populations improves discovery for complex traits. *Nature* 570, 514–518 (2019) (attached)
- Howie et al. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nature Genetics 2012 (attached)
- Marchini and Howie. Genotype imputation for genome-wide association studies. Nature Reviews Genetics 2010 (attached)
- van Leeuwen EM, Kanterakis A, Deelen P, Kattenberg MV; Genome of the Netherlands Consortium, Slagboom PE, de Bakker PI, Wijmenga C, Swertz MA, Boomsma DI, van Duijn CM, Karssen LC, Hottenga JJ. Population-specific genotype imputations using minimac or IMPUTE2. *Nat Protoc*. 2015 Sep;10(9):1285-96.
- Benefits and limitations of genome-wide association studies. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Nat Rev Genet. 2019 Aug;20(8):467-484.

Polygenic risk scores:

- What are polygenic risk scores and why are they important? *JAMA* May 14, 2019 Volume 321, Number 18
- Martin et al. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet*. 2019 Apr;51(4):584-591
- Wan Choi et al. Tutorial: a guide to performing polygenic risk score analyses. *Nat Protoc*. 2020 Sep;15(9):2759-2772.
- Lewis and Vassos. Polygenic risk scores: from research tools to clinical instruments. *Genome Medicine* (2020)
- Ali Torkamani, Nathan E. Wineinger & Eric J. Topol. The personal and clinical utility of polygenic risk scores. *Nature Reviews Genetics* volume 19, pages581–590 (2018).
- Wand H, Lambert SA, Tamburro C, et al. Improving reporting standards for polygenic scores in risk prediction studies. *Nature.* Published online March 10, 2021. doi:10.1038/s41586-021-03243-6
- Lambert et al. The Polygenic Score Catalog as an open database for reproducibility and systematic evaluation. Nature Genetics 2021

Sequencing:

Websites:

BEST PRACTICES FOR VARIANT CALLING WITH THE GATK - Link to presentations filmed during the March 2015 GATK Workshop, part of the BroadE Workshop series. This workshop focused on the core steps involved in calling variants with Broad's Genome Analysis Toolkit, using the "Best Practices" developed by the GATK team. View the workshop materials below to gain an understanding of the rationale, theory, and real-life applications of GATK Best Practices. Learn why each step is essential to the calling process, what key operations are performed on the data at each step, and how to use the GATK tools to get the most accurate and reliable results out of your dataset. Best Practices for Variant Calling with the GATK | Broad Institute (Links to an external site.)

GATK: GATK (broadinstitute.org) (Links to an external site.)
The Broad technical workshops: https://www.broadinstitute.org/scientists/technical-workshops (Links to an external site.)

On CADD (Combined Annotation Dependent Depletion; variant annotation): https://cadd.gs.washington.edu/ (Links to an external site.)

Papers:

Jay Shendure et al DNA sequencing at 40: past, present and future. Nature volume 550, pages 345–353 (2017) Shendurenature24286.pdf

Daniel C. Koboldt Best practices for variant calling in clinical sequencing. Genome Medicine volume 12, Article number: 91 (2020) Koboldts13073-020-00791-w.pdf

Altmann et al. A beginners guide to SNP calling from high-throughput DNA-sequencing data Hum Genet 2012 Oct;131(10):1541-54. Altmann2012_Article_ABeginnersGuideToSNPCallingFro.pdf

DePristo et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature Genetics volume 43, pages491–498 (2011) Depristoetalng.806.pdf

Family-based approaches:

Ellen Wijsman. Family-based approaches: design, imputation, analysis, and beyond. BMC Genetics 2016, 17(Suppl 2):9: family-based_Wijsman-1.pdf

Ott et al. Family-based designs for genome-wide association studies. Nature Reviews, July 2011: family-based designsnrg2989.pdf

Ott et al. Genetic linkage analysis in the age of whole-genome sequencing. Nature Reviews Genetics volume 16, pages275–284 (2015): Ottsequencingnrg3908.pdf

Methylation, chromatin profiling - background papers

Methylation:

Epigenome-wide association studies for common human diseases. Vardhman K. Rakyan, Thomas A. Down, David J. Balding & Stephan Beck Nature Reviews Genetics volume 12, pages529–541 (2011) Reviewnrg3000.pdf

Statistical and integrative system-level analysis of DNA methylation data. Andrew E. Teschendorff & Caroline L. Relton Nature Reviews Genetics volume 19, pages129–147 (2018). MethylationReltonnrg.2017.86.pdf

Recommendations for the design and analysis of epigenome-wide association studies. Karin B Michels, Alexandra M Binder, Sarah Dedeurwaerder, Charles B Epstein, John M Greally, Ivo Gut, E Andres Houseman, Benedetta Izzi, Karl T Kelsey, Alexander Meissner, Aleksandar Milosavljevic, Kimberly D Siegmund, Christoph Bock & Rafael A Irizarry. Nature Methods volume 10, pages949–955 (2013) Michelsetalsnmeth.2632.pdf

Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Martin J Aryee, Andrew E Jaffe, Hector Corrada-Bravo, Christine Ladd-Acosta, Andrew P Feinberg, Kasper D Hansen, Rafael A Irizarry Bioinformatics 2014 May 15;30(10):1363-9. doi: 10.1093/bioinformatics/btu049. Epub 2014 Jan 28. Minfi_btu049.pdf

Chromatin:

Chromatin accessibility profiling methods. Nature Reviews Methods Primers volume 1, Article number: 11 (2021) . (Highlights the main chromatin accessibility profiling methods. These methods include DNase-seq, ATAC-seq, MNase-seq, and DNA methylation-based methods to assess open chromatin and regulatory elements.) Chromatinmethodss43586-020-00010-1.pdf

Liesbeth Minnoye, Georgi K. Marinov, Thomas Krausgruber, Lixia Pan, Alexandre P. Marand, Stefano Secchia, William J. Greenleaf, Eileen E. M. Furlong, Keji Zhao, Robert J. Schmitz, Christoph Bock & Stein Aerts. Chromatin accessibility profiling methods. Nature Reviews Methods Primers volume 1, Article number: 10 (2021) chromatinprofilingmethodss43586-020-00008-9.pdf

From reads to insight: a hitchhiker's guide to ATAC-seq data analysis. Feng Yan, David R. Powell, David J. Curtis & Nicholas C. Wong Genome Biology volume 21, Article number: 22 (2020). ATACseqs13059-020-1929-3.pdf

Gene expression, TWAS, RNA-seq:

Wainberg M, Sinnott-Armstrong N, Mancuso N, Barbeira AN, Knowles DA, Golan D, Ermel R, Ruusalepp A, Quertermous T, Hao K, Björkegren JLM, Im HK, Pasaniuc B, Rivas MA, Kundaje A. Opportunities and challenges for transcriptome-wide association studies. Nat Genet. 2019 Apr;51(4):592-599. Wainbergs41588-019-0385-z (2).pdf

Michael D. Gallagher and Alice S. Chen-Plotkin. The Post-GWAS Era: From Association to Function. Am J Hum Genet. 2018 May 3; 102(5): 717–730. GallagherandChen_plotkin-1.pdf

Alexander Gusev, Arthur Ko, Huwenbo Shi, Gaurav Bhatia, Wonil Chung, Brenda W J H Penninx, Rick Jansen, Eco J C de Geus, Dorret I Boomsma, Fred A Wright, Patrick F Sullivan, Elina Nikkola, Marcus Alvarez, Mete Civelek, Aldons J Lusis, Terho Lehtimäki, Emma Raitoharju, Mika Kähönen, Ilkka Seppälä, Olli T Raitakari, Johanna Kuusisto, Markku Laakso, Alkes L Price, Päivi Pajukanta & Bogdan Pasaniuc. Integrative approaches for large-scale transcriptome-wide association studies. Nature Genetics volume 48, pages245–252 (2016). Gusev2016ng.3506-1.pdf

Rory Stark, Marta Grzelak & James Hadfield. RNA sequencing: the teenage years. Nature Reviews Genetics volume 20, pages631–656 (2019). RNAsequencing_theteenageyearss41576-019-0150-2.pdf

McDermaid et al. Interpretation of differential gene expression results of RNA-seq data: review and integration. Brief Bioinform 2019 Nov 27;20(6):2044-2054. doi: 10.1093/bib/bby067. McDermaidbby067.pdf

Van den Berge et al. RNA Sequencing Data: Hitchhiker's Guide to Expression Analysis. Annual Review of Biomedical Data Science. Vol. 2:139-173 (Volume publication date July 2019) https://doi.org/10.1146/annurev-biodatasci-072018-021255 (Links to an external site.) annurev-biodatasci-072018-021255.pdf

Geng Chen, Baitang Ning and Tieliu Shi. Single-Cell RNA-Seq Technologies and Related Computational Data Analysis. Front. Genet., 05 April 2019https://doi.org/10.3389/fgene.2019.00317 (Links to an external site.) fgene-10-00317.pdf

Ashraful Haque, Jessica Engel, Sarah A. Teichmann & Tapio Lönnberg. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. Genome Medicine volume 9, Article number: 75 (2017). Haqueetals13073-017-0467-4.pdf

CNV and other structural variation:

Nigel P Carter Methods and strategies for analyzing copy number variation using DNA microarrays. Nature Genetics 2007. Carterng2028.pdf

Wang et al. PennCNV: An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data.PennCNVGenome Res.-2007-Wang-1665-74.pdf

Tech note from Illumina: Interpreting Infinium® Assay Data for Whole-Genome Structural Variation technote_cytoanalysis.pdf

Lin et al. Analyzing Copy Number Variation using SNP Array Data: Protocols for Calling CNV and Association Tests. Curr Protoc Hum Genet. 2014 LinetalCNVnihms-531582.pdf

Zarrei et al. A copy number variation map of the human genome. Nature Reviews Genetics 2015 ZarreiCNVnrg3871.pdf

Lars Feuk, Andrew R. Carson & Stephen W. Scherer. Structural variation in the human genome. Nature Reviews Genetics volume 7, pages85–97 (2006) nrg1767.pdf

Spielman et al. Structural variation in the 3D genome. Nature Reviews Genetics 2018 Spielmanetals41576-018-0007-0.pdf

Collins et al. A structural variation reference for medical and population genetics. Nature volume 581, pages444–451 (2020). Collinsetals41586-020-2287-8.pdf