

# References

- [1] Acuna-Hidalgo, R., Veltman, J. A., and Hoischen, A. (2016). New insights into the generation and role of de novo mutations in health and disease. *Genome Biology*, 17(1):241.
- [2] Adewoye, A. B., Lindsay, S. J., Dubrova, Y. E., and Hurles, M. E. (2015). The genome-wide effects of ionizing radiation on mutation induction in the mammalian germline. *Nature Communications*, 6(1):1–8.
- [3] Aggarwala, V. and Voight, B. F. (2016). An expanded sequence context model broadly explains variability in polymorphism levels across the human genome. *Nature Genetics*, 48(4):349.
- [4] Alexandrov, L. B., Kim, J., Haradhvala, N. J., Huang, M. N., Ng, A. W. T., Wu, Y., Boot, A., Covington, K. R., Gordenin, D. A., Bergstrom, E. N., et al. (2020). The repertoire of mutational signatures in human cancer. *Nature*, 578(7793):94–101.
- [5] Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A., Behjati, S., Biankin, A. V., Bignell, G. R., Bolli, N., Borg, A., Børresen-Dale, A.-L., et al. (2013). Signatures of mutational processes in human cancer. *Nature*, 500(7463):415.
- [6] Allen, A. S., Berkovic, S. F., Cossette, P., Delanty, N., Dlugos, D., Eichler, E. E., Epstein, M. P., Glauser, T., Goldstein, D. B., Han, Y., et al. (2013). De novo mutations in epileptic encephalopathies. *Nature*, 501(7466):217–221.
- [7] Amirnovin, R. (1997). An analysis of the metabolic theory of the origin of the genetic code. *Journal of Molecular Evolution*, 44(5):473–476.
- [8] Amos, W. (2010). Even small snp clusters are non-randomly distributed: is this evidence of mutational non-independence? *Proceedings of the Royal Society B: Biological Sciences*, 277(1686):1443–1449.
- [9] Amos, W. (2019). Flanking heterozygosity influences the relative probability of different base substitutions in humans. *Royal Society open science*, 6(9):191018.
- [10] Arnheim, N. and Calabrese, P. (2009). Understanding what determines the frequency and pattern of human germline mutations. *Nature Reviews Genetics*, 10(7):478–488.
- [11] Beal, M. A., Yauk, C. L., and Marchetti, F. (2017). From sperm to offspring: Assessing the heritable genetic consequences of paternal smoking and potential public health impacts. *Mutation Research/Reviews in Mutation Research*, 773:26–50.

- [12] Bergstrom, E. N., Huang, M. N., Mahto, U., Barnes, M., Stratton, M. R., Rozen, S. G., and Alexandrov, L. B. (2019). Sigproflernmatrixgenerator: a tool for visualizing and exploring patterns of small mutational events. *BMC Genomics*, 20(1):1–12.
- [13] Bertl, J., Guo, Q., Juul, M., Besenbacher, S., Nielsen, M. M., Hornshøj, H., Pedersen, J. S., and Hobolth, A. (2018). A site specific model and analysis of the neutral somatic mutation rate in whole-genome cancer data. *BMC Bioinformatics*, 19(1):1–15.
- [14] Besenbacher, S., Sulem, P., Helgason, A., Helgason, H., Kristjansson, H., Jonasdottir, A., Jonasdottir, A., Magnusson, O. T., Thorsteinsdottir, U., Masson, G., et al. (2016). Multi-nucleotide de novo mutations in humans. *PLoS Genetics*, 12(11):e1006315.
- [15] Bhérer, C., Campbell, C. L., and Auton, A. (2017). Refined genetic maps reveal sexual dimorphism in human meiotic recombination at multiple scales. *Nature Communications*, 8(1):1–9.
- [16] Blake, R., Hess, S. T., and Nicholson-Tuell, J. (1992). The influence of nearest neighbors on the rate and pattern of spontaneous point mutations. *Journal of Molecular Evolution*, 34(3):189–200.
- [17] Botstein, D. and Risch, N. (2003). Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nature Genetics*, 33(3):228–237.
- [18] Boycott, K. M., Vanstone, M. R., Bulman, D. E., and MacKenzie, A. E. (2013). Rare-disease genetics in the era of next-generation sequencing: discovery to translation. *Nature Reviews Genetics*, 14(10):681–691.
- [19] BrainSpan (2016). BrainSpan: Atlas of the developing human brain.
- [20] Campbell, I. M., Yuan, B., Robberecht, C., Pfundt, R., Szafranski, P., McEntagart, M. E., Nagamani, S. C., Erez, A., Bartnik, M., Wiśniowiecka-Kowalnik, B., et al. (2014). Parental somatic mosaicism is underrecognized and influences recurrence risk of genomic disorders. *The American Journal of Human Genetics*, 95(2):173–182.
- [21] Campos, D., Matos, S., and Oliveira, J. L. (2013). A modular framework for biomedical concept recognition. *BMC Bioinformatics*, 14(1):1–21.
- [22] Carlson, J., Locke, A. E., Flickinger, M., Zawistowski, M., Levy, S., Myers, R. M., Boehnke, M., Kang, H. M., Scott, L. J., Li, J. Z., et al. (2018). Extremely rare variants reveal patterns of germline mutation rate heterogeneity in humans. *Nature Communications*, 9(1):1–13.
- [23] Carter, T., Lyon, M. F., and Phillips, R. J. (1956). Induction of mutations in mice by chronic gamma irradiation; interim report. *The British Journal of Radiology*, 29(338):106–108.
- [24] Cassa, C. A., Weghorn, D., Balick, D. J., Jordan, D. M., Nusinow, D., Samocha, K. E., O’Donnell-Luria, A., MacArthur, D. G., Daly, M. J., Beier, D. R., et al. (2017). Estimating the selective effects of heterozygous protein-truncating variants from human exome data. *Nature Genetics*, 49(5):806–810.

- [25] Cawthon, R. M., Meeks, H. D., Sasani, T. A., Smith, K. R., Kerber, R. A., O'Brien, E., Baird, L., Dixon, M. M., Peiffer, A. P., Leppert, M. F., et al. (2020). Germline mutation rates in young adults predict longevity and reproductive lifespan. *medRxiv*, page 19004184.
- [26] Chance, P. F., Alderson, M. K., Leppig, K. A., Lensch, M. W., Matsunami, N., Smith, B., Swanson, P. D., Odelberg, S. J., Disteche, C. M., and Bird, T. D. (1993). DNA deletion associated with hereditary neuropathy with liability to pressure palsies. *Cell*, 72(1):143–151.
- [27] Chen, J.-M., Férec, C., and Cooper, D. N. (2015). Complex multiple-nucleotide substitution mutations causing human inherited disease reveal novel insights into the action of translesion synthesis DNA polymerases. *Human Mutation*, 36(11):1034–1038.
- [28] Ciriello, G., Miller, M. L., Aksoy, B. A., Senbabaoglu, Y., Schultz, N., and Sander, C. (2013). Emerging landscape of oncogenic signatures across human cancers. *Nature Genetics*, 45(10):1127.
- [29] Cleaver, J. E., Lam, E. T., and Revet, I. (2009). Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. *Nature Reviews Genetics*, 10(11):756–768.
- [30] Clement, Y. and Arndt, P. F. (2011). Substitution patterns are under different influences in primates and rodents. *Genome Biology and Evolution*, 3:236–245.
- [31] Coe, B. P., Stessman, H. A., Sulovari, A., Geisheker, M. R., Bakken, T. E., Lake, A. M., Dougherty, J. D., Lein, E. S., Hormozdiari, F., Bernier, R. A., et al. (2019). Neurodevelopmental disease genes implicated by de novo mutation and copy number variation morbidity. *Nature Genetics*, 51(1):106–116.
- [32] Coe, B. P., Witherspoon, K., Rosenfeld, J. A., Van Bon, B. W., Vulto-van Silfhout, A. T., Bosco, P., Friend, K. L., Baker, C., Buono, S., Vissers, L. E., et al. (2014). Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nature Genetics*, 46(10):1063.
- [33] Conrad, D. F., Keebler, J. E., DePristo, M. A., Lindsay, S. J., Zhang, Y., Casals, F., Idaghdour, Y., Hartl, C. L., Torroja, C., Garimella, K. V., et al. (2011). Variation in genome-wide mutation rates within and between human families. *Nature Genetics*, 43(7):712.
- [34] Consortium, . G. P. et al. (2015). A global reference for human genetic variation. *Nature*, 526(7571):68–74.
- [35] Consortium, M. G. S., Waterston, R., Lindblad-Toh, K., Birney, E., and Rogers, J. (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature*, 420(6915):520–562.
- [36] Cooper, G. M., Coe, B. P., Girirajan, S., Rosenfeld, J. A., Vu, T. H., Baker, C., Williams, C., Stalker, H., Hamid, R., Hannig, V., et al. (2011). A copy number variation morbidity map of developmental delay. *Nature Genetics*, 43(9):838.

- [37] Coulondre, C., Miller, J. H., Farabaugh, P. J., and Gilbert, W. (1978). Molecular basis of base substitution hotspots in *Escherichia coli*. *Nature*, 274(5673):775–780.
- [38] Croft, D., Mundo, A. F., Haw, R., Milacic, M., Weiser, J., Wu, G., Caudy, M., Garapati, P., Gillespie, M., Kamdar, M. R., et al. (2014). The Reactome pathway knowledgebase. *Nucleic Acids Research*, 42(D1):D472–D477.
- [39] De Ligt, J., Willemsen, M. H., Van Bon, B. W., Kleefstra, T., Yntema, H. G., Kroes, T., Vulto-van Silfhout, A. T., Koolen, D. A., De Vries, P., Gilissen, C., et al. (2012). Diagnostic exome sequencing in persons with severe intellectual disability. *New England Journal of Medicine*, 367(20):1921–1929.
- [40] de Vries, B. B., Pfundt, R., Leisink, M., Koolen, D. A., Vissers, L. E., Janssen, I. M., van Reijmersdal, S., Nillesen, W. M., Huys, E. H., de Leeuw, N., et al. (2005). Diagnostic genome profiling in mental retardation. *The American Journal of Human Genetics*, 77(4):606–616.
- [41] Deciphering Developmental Disorders Study (2017). Prevalence and architecture of de novo mutations in developmental disorders. *Nature*, 542(7642):433.
- [42] Deelen, P., van Dam, S., Herkert, J. C., Karjalainen, J. M., Brugge, H., Abbott, K. M., van Diemen, C. C., van der Zwaag, P. A., Gerkes, E. H., Zonneveld-Huijssoon, E., et al. (2019). Improving the diagnostic yield of exome-sequencing by predicting gene-phenotype associations using large-scale gene expression analysis. *Nature Communications*, 10(1):1–13.
- [43] DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A., Del Angel, G., Rivas, M. A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, 43(5):491.
- [44] Diaz, J., Berger, S., and Leon, E. (2019). TFE3-associated neurodevelopmental disorder: A distinct recognizable syndrome. *American Journal of Medical Genetics Part A*.
- [45] Duncan, B. K. and Miller, J. H. (1980). Mutagenic deamination of cytosine residues in DNA. *Nature*, 287(5782):560–561.
- [46] Dutil, J., Teer, J. K., Golubeva, V., Yoder, S., Tong, W. L., Arroyo, N., Karam, R., Echenique, M., Matta, J. L., and Monteiro, A. N. (2019). Germline variants in cancer genes in high-risk non-brca patients from Puerto Rico. *Scientific Reports*, 9(1):1–11.
- [47] Elder, R. H., Jansen, J. G., Weeks, R. J., Willington, M. A., Deans, B., Watson, A. J., Mynett, K. J., Bailey, J. A., Cooper, D. P., Rafferty, J. A., et al. (1998). Alkylpurine-dna-n-glycosylase knockout mice show increased susceptibility to induction of mutations by methyl methanesulfonate. *Molecular and Cellular Biology*, 18(10):5828–5837.
- [48] Eyler, D. E., Burnham, K. A., Wilson, T. E., and O'Brien, P. J. (2017). Mechanisms of glycosylase induced genomic instability. *PloS One*, 12(3).
- [49] Fan, H. and Chu, J.-Y. (2007). A brief review of short tandem repeat mutation. *Genomics, Proteomics & Bioinformatics*, 5(1):7–14.

- [50] Firth, H. V. and Wright, C. F. (2011). The Deciphering Developmental Disorders (DDD) study. *Developmental Medicine & Child Neurology*, 53(8):702–703.
- [51] Forster, L., Forster, P., Lutz-Bonengel, S., Willkomm, H., and Brinkmann, B. (2002). Natural radioactivity and human mitochondrial DNA mutations. *Proceedings of the National Academy of Sciences*, 99(21):13950–13954.
- [52] Francioli, L. C., Polak, P. P., Koren, A., Menelaou, A., Chun, S., Renkens, I., Van Duijn, C. M., Swertz, M., Wijmenga, C., Van Ommen, G., et al. (2015). Genome-wide patterns and properties of de novo mutations in humans. *Nature Genetics*, 47(7):822.
- [53] Fu, W., Zhang, F., Wang, Y., Gu, X., and Jin, L. (2010). Identification of copy number variation hotspots in human populations. *The American Journal of Human Genetics*, 87(4):494–504.
- [54] Gao, Z., Moorjani, P., Sasani, T. A., Pedersen, B. S., Quinlan, A. R., Jorde, L. B., Amster, G., and Przeworski, M. (2019). Overlooked roles of DNA damage and maternal age in generating human germline mutations. *Proceedings of the National Academy of Sciences*, 116(19):9491–9500.
- [55] Garcia-Diaz, M. and Kunkel, T. A. (2006). Mechanism of a genetic glissando\*: structural biology of indel mutations. *Trends in Biochemical Sciences*, 31(4):206–214.
- [56] Garrison, E. and Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv preprint arXiv:1207.3907*.
- [57] Gerstung, M., Papaemmanuil, E., and Campbell, P. J. (2014). Subclonal variant calling with multiple samples and prior knowledge. *Bioinformatics*, 30(9):1198–1204.
- [58] Gilissen, C., Hehir-Kwa, J. Y., Thung, D. T., van de Vorst, M., van Bon, B. W., Willemsen, M. H., Kwint, M., Janssen, I. M., Hoischen, A., Schenck, A., et al. (2014). Genome sequencing identifies major causes of severe intellectual disability. *Nature*, 511(7509):344–347.
- [59] Gilissen, C., Hoischen, A., Brunner, H. G., and Veltman, J. A. (2012). Disease gene identification strategies for exome sequencing. *European Journal of Human Genetics*, 20(5):490–497.
- [60] Girard, S. L., Bourassa, C. V., Lemieux Perreault, L.-P., Legault, M.-A., Barhdadi, A., Ambalavanan, A., Brendgen, M., Vitaro, F., Noreau, A., Dionne, G., et al. (2016). Paternal age explains a major portion of de novo germline mutation rate variability in healthy individuals. *PLoS One*, 11(10):e0164212.
- [61] Glassner, B. J., Rasmussen, L. J., Najarian, M. T., Posnick, L. M., and Samson, L. D. (1998). Generation of a strong mutator phenotype in yeast by imbalanced base excision repair. *Proceedings of the National Academy of Sciences*, 95(17):9997–10002.
- [62] Goldmann, J. M., Wong, W. S., Pinelli, M., Farrah, T., Bodian, D., Stittrich, A. B., Glusman, G., Vissers, L. E., Hoischen, A., Roach, J. C., et al. (2016). Parent-of-origin-specific signatures of de novo mutations. *Nature Genetics*, 48(8):935.

- [63] Goodman, M. (1963). Serological analysis of the systematics of recent hominoids. *Human Biology*, 35(3):377–436.
- [64] Goriely, A., Hansen, R. M., Taylor, I. B., Olesen, I. A., Jacobsen, G. K., McGowan, S. J., Pfeifer, S. P., McVean, G. A., Rajpert-De Meyts, E., and Wilkie, A. O. (2009). Activating mutations in FGFR3 and hras reveal a shared genetic origin for congenital disorders and testicular tumors. *Nature Genetics*, 41(11):1247.
- [65] Goriely, A., McGrath, J. J., Hultman, C. M., Wilkie, A. O., and Malaspina, D. (2013). “selfish spermatogonial selection”: a novel mechanism for the association between advanced paternal age and neurodevelopmental disorders. *American Journal of Psychiatry*, 170(6):599–608.
- [66] Goriely, A. and Wilkie, A. O. (2010). Missing heritability: paternal age effect mutations and selfish spermatogonia. *Nature Reviews Genetics*, 11(8):589–589.
- [67] Goriely, A. and Wilkie, A. O. (2012). Paternal age effect mutations and selfish spermatogonial selection: causes and consequences for human disease. *The American Journal of Human Genetics*, 90(2):175–200.
- [68] Green, P., Ewing, B., Miller, W., Thomas, P. J., and Green, E. D. (2003). Transcription-associated mutational asymmetry in mammalian evolution. *Nature Genetics*, 33(4):514–517.
- [69] Guo, H., Wang, T., Wu, H., Long, M., Coe, B. P., Li, H., Xun, G., Ou, J., Chen, B., Duan, G., et al. (2018). Inherited and multiple de novo mutations in autism/developmental delay risk genes suggest a multifactorial model. *Molecular Autism*, 9(1):64.
- [70] Gymrek, M., Willems, T., Reich, D., and Erlich, Y. (2017). Interpreting short tandem repeat variations in humans using mutational constraint. *Nature Genetics*, 49(10):1495.
- [71] Haldane, J. (1946). The mutation rate of the gene for haemophilia, and its segregation ratios in males and females. *Annals of Eugenics*, 13(1):262–271.
- [72] Haldane, J. B. (1935). The rate of spontaneous mutation of a human gene. *Journal of Genetics*, 31(3):317.
- [73] Halldorsson, B. V., Palsson, G., Stefansson, O. A., Jonsson, H., Hardarson, M. T., Eggertsson, H. P., Gunnarsson, B., Oddsson, A., Halldorsson, G. H., Zink, F., et al. (2019). Characterizing mutagenic effects of recombination through a sequence-level genetic map. *Science*, 363(6425):eaau1043.
- [74] Happle, R. (1987). Lethal genes surviving by mosaicism: a possible explanation for sporadic birth defects involving the skin. *Journal of the American Academy of Dermatology*, 16(4):899–906.
- [75] Harris, K. (2015). Evidence for recent, population-specific evolution of the human mutation rate. *Proceedings of the National Academy of Sciences*, 112(11):3439–3444.
- [76] Harris, K. and Nielsen, R. (2014). Error-prone polymerase activity causes multinucleotide mutations in humans. *Genome Research*, 24(9):1445–1454.

- [77] Harris, R. S. (2013). Cancer mutation signatures, DNA damage mechanisms, and potential clinical implications. *Genome Medicine*, 5(9):87.
- [78] He, X., Sanders, S. J., Liu, L., De Rubeis, S., Lim, E. T., Sutcliffe, J. S., Schellenberg, G. D., Gibbs, R. A., Daly, M. J., Buxbaum, J. D., et al. (2013). Integrated model of de novo and inherited genetic variants yields greater power to identify risk genes. *PLoS Genetics*, 9(8).
- [79] Hendrich, B., Hardeland, U., Ng, H.-H., Jiricny, J., and Bird, A. (1999). The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature*, 401(6750):301–304.
- [80] Hodgkinson, A. and Eyre-Walker, A. (2011). Variation in the mutation rate across mammalian genomes. *Nature Reviews Genetics*, 12(11):756–766.
- [81] Hu, C., Hart, S. N., Polley, E. C., Gnanaolivu, R., Shimelis, H., Lee, K. Y., Lilyquist, J., Na, J., Moore, R., Antwi, S. O., et al. (2018). Association between inherited germline mutations in cancer predisposition genes and risk of pancreatic cancer. *JAMA*, 319(23):2401–2409.
- [82] Huang, K.-I., Mashl, R. J., Wu, Y., Ritter, D. I., Wang, J., Oh, C., Paczkowska, M., Reynolds, S., Wyczalkowski, M. A., Oak, N., et al. (2018). Pathogenic germline variants in 10,389 adult cancers. *Cell*, 173(2):355–370.
- [83] Huang, N., Lee, I., Marcotte, E. M., and Hurles, M. E. (2010). Characterising and predicting haploinsufficiency in the human genome. *PLoS Genetics*, 6(10).
- [84] Huang, Q.-Y., Xu, F.-H., Shen, H., Deng, H.-Y., Liu, Y.-J., Liu, Y.-Z., Li, J.-L., Recker, R. R., and Deng, H.-W. (2002). Mutation patterns at dinucleotide microsatellite loci in humans. *The American Journal of Human Genetics*, 70(3):625–634.
- [85] Huang, Y.-F., Gulko, B., and Siepel, A. (2017). Fast, scalable prediction of deleterious noncoding variants from functional and population genomic data. *Nature Genetics*, 49(4):618–624.
- [86] Hwang, D. G. and Green, P. (2004). Bayesian markov chain monte carlo sequence analysis reveals varying neutral substitution patterns in mammalian evolution. *Proceedings of the National Academy of Sciences*, 101(39):13994–14001.
- [87] Iafrate, A. J., Feuk, L., Rivera, M. N., Listewnik, M. L., Donahoe, P. K., Qi, Y., Scherer, S. W., and Lee, C. (2004). Detection of large-scale variation in the human genome. *Nature Genetics*, 36(9):949–951.
- [88] Ionita-Laza, I., McCallum, K., Xu, B., and Buxbaum, J. D. (2016). A spectral approach integrating functional genomic annotations for coding and noncoding variants. *Nature Genetics*, 48(2):214.
- [89] Iossifov, I., O’roak, B. J., Sanders, S. J., Ronemus, M., Krumm, N., Levy, D., Stessman, H. A., Witherspoon, K. T., Vives, L., Patterson, K. E., et al. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature*, 515(7526):216–221.

- [90] Itsara, A., Wu, H., Smith, J. D., Nickerson, D. A., Romieu, I., London, S. J., and Eichler, E. E. (2010). De novo rates and selection of large copy number variation. *Genome Research*, 20(11):1469–1481.
- [91] Jaganathan, K., Panagiotopoulou, S. K., McRae, J. F., Darbandi, S. F., Knowles, D., Li, Y. I., Kosmicki, J. A., Arbelaez, J., Cui, W., Schwartz, G. B., et al. (2019). Predicting splicing from primary sequence with deep learning. *Cell*, 176(3):535–548.
- [92] Jager, M., Blokzijl, F., Kuijk, E., Bertl, J., Vougioukalaki, M., Janssen, R., Besselink, N., Boymans, S., de Ligt, J., Pedersen, J. S., et al. (2019). Deficiency of nucleotide excision repair is associated with mutational signature observed in cancer. *Genome Research*, 29(7):1067–1077.
- [93] Jeffreys, A. J., Royle, N. J., Wilson, V., and Wong, Z. (1988). Spontaneous mutation rates to new length alleles at tandem-repetitive hypervariable loci in human DNA. *Nature*, 332(6161):278–281.
- [94] Jeffreys, A. J., Tamaki, K., MacLeod, A., Monckton, D. G., Neil, D. L., and Armour, J. A. (1994). Complex gene conversion events in germline mutation at human minisatellites. *Nature Genetics*, 6(2):136–145.
- [95] Jeffreys, A. J., Wilson, V., and Thein, S. L. (1985). Hypervariable ‘minisatellite’ regions in human DNA. *Nature*, 314(6006):67–73.
- [96] Johnson, R. E., Washington, M. T., Prakash, S., and Prakash, L. (2000). Fidelity of human DNA polymerase  $\eta$ . *Journal of Biological Chemistry*, 275(11):7447–7450.
- [97] Jónsson, H., Sulem, P., Arnadottir, G. A., Pálsson, G., Eggertsson, H. P., Kristmundsdottir, S., Zink, F., Kehr, B., Hjorleifsson, K. E., Jensson, B. Ö., et al. (2018). Multiple transmissions of de novo mutations in families. *Nature Genetics*, 50(12):1674–1680.
- [98] Jónsson, H., Sulem, P., Kehr, B., Kristmundsdottir, S., Zink, F., Hjartarson, E., Hardarson, M. T., Hjorleifsson, K. E., Eggertsson, H. P., Gudjonsson, S. A., et al. (2017). Parental influence on human germline de novo mutations in 1,548 trios from iceland. *Nature*, 549(7673):519–522.
- [99] Ju, Y. S., Martincorena, I., Gerstung, M., Petljak, M., Alexandrov, L. B., Rahbari, R., Wedge, D. C., Davies, H. R., Ramakrishna, M., Fullam, A., et al. (2017). Somatic mutations reveal asymmetric cellular dynamics in the early human embryo. *Nature*, 543(7647):714–718.
- [100] Kaplanis, J., Akawi, N., Gallone, G., McRae, J. F., Prigmore, E., Wright, C. F., Fitzpatrick, D. R., Firth, H. V., Barrett, J. C., Hurles, M. E., et al. (2019). Exome-wide assessment of the functional impact and pathogenicity of multinucleotide mutations. *Genome Research*, 29(7):1047–1056.
- [101] Kaplanis, J., Samocho, K. E., Wiel, L., Zhang, Z., Arvai, K. J., Eberhardt, R. Y., Gallone, G., Lelieveld, S. H., Martin, H. C., McRae, J. F., et al. (2020). Integrating healthcare and research genetic data empowers the discovery of 28 novel developmental disorders. *BioRxiv*, page 797787.



- [102] Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., et al. (2019). Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *BioRxiv*, page 531210.
- [103] Kato, T., Inagaki, H., Yamada, K., Kogo, H., Ohye, T., Kowa, H., Nagaoka, K., Taniguchi, M., Emanuel, B. S., and Kurahashi, H. (2006). Genetic variation affects de novo translocation frequency. *Science*, 311(5763):971–971.
- [104] Keshava Prasad, T., Goel, R., Kandasamy, K., Keerthikumar, S., Kumar, S., Mathivanan, S., Telikicherla, D., Raju, R., Shafreen, B., Venugopal, A., et al. (2009). Human protein reference database—2009 update. *Nucleic Acids Research*, 37(suppl\_1):D767–D772.
- [105] Kilpivaara, O. and Aaltonen, L. (2013). Diagnostic cancer genome sequencing and the contribution of germline variants. *Science*, 339(6127):1559–1562.
- [106] Kimura, M. (1968). Evolutionary rate at the molecular level. *Nature*, 217(5129):624–626.
- [107] Kircher, M., Witten, D. M., Jain, P., O’Roak, B. J., Cooper, G. M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics*, 46(3):310.
- [108] Klein, H. L. (2017). Genome instabilities arising from ribonucleotides in DNA. *DNA Repair*, 56:26–32.
- [109] Kondrashov, A. S. (2003). Direct estimates of human per nucleotide mutation rates at 20 loci causing Mendelian diseases. *Human Mutation*, 21(1):12–27.
- [110] Kondrashov, A. S. and Crow, J. F. (1993). A molecular approach to estimating the human deleterious mutation rate. *Human mutation*, 2(3):229–234.
- [111] Kong, A., Frigge, M. L., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., Gudjonsson, S. A., Sigurdsson, A., Jonasdottir, A., Jonasdottir, A., et al. (2012). Rate of de novo mutations and the importance of father’s age to disease risk. *Nature*, 488(7412):471–475.
- [112] Kosmicki, J. A., Samocha, K. E., Howrigan, D. P., Sanders, S. J., Slowikowski, K., Lek, M., Karczewski, K. J., Cutler, D. J., Devlin, B., Roeder, K., et al. (2017). Refining the role of de novo protein-truncating variants in neurodevelopmental disorders by using population reference samples. *Nature Genetics*, 49(4):504.
- [113] Kurek, K. C., Luks, V. L., Ayturk, U. M., Alomari, A. I., Fishman, S. J., Spencer, S. A., Mulliken, J. B., Bowen, M. E., Yamamoto, G. L., Kozakewich, H. P., et al. (2012). Somatic mosaic activating mutations in PIK3CA cause CLOVES syndrome. *The American Journal of Human Genetics*, 90(6):1108–1115.
- [114] Lai, Z., Markovets, A., Ahdesmaki, M., Chapman, B., Hofmann, O., McEwen, R., Johnson, J., Dougherty, B., Barrett, J. C., and Dry, J. R. (2016). VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. *Nucleic Acids Research*, 44(11):e108–e108.

- [115] Landrum, M. J., Lee, J. M., Riley, G. R., Jang, W., Rubinstein, W. S., Church, D. M., and Maglott, D. R. (2013). ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Research*, 42(D1):D980–D985.
- [116] Lange, S. S., Takata, K.-i., and Wood, R. D. (2011). DNA polymerases and cancer. *Nature Reviews Cancer*, 11(2):96–110.
- [117] Laskowski, R. A., Stephenson, J. D., Sillitoe, I., Orengo, C. A., and Thornton, J. M. (2020). VarSite: Disease variants and protein structure. *Protein Science*, 29(1):111–119.
- [118] Lau, A. Y., Wyatt, M. D., Glassner, B. J., Samson, L. D., and Ellenberger, T. (2000). Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, aag. *Proceedings of the National Academy of Sciences*, 97(25):13573–13578.
- [119] Lawrence, M. S., Stojanov, P., Mermel, C. H., Robinson, J. T., Garraway, L. A., Golub, T. R., Meyerson, M., Gabriel, S. B., Lander, E. S., and Getz, G. (2014). Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature*, 505(7484):495–501.
- [120] Lawrence, M. S., Stojanov, P., Polak, P., Kryukov, G. V., Cibulskis, K., Sivachenko, A., Carter, S. L., Stewart, C., Mermel, C. H., Roberts, S. A., et al. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*, 499(7457):214–218.
- [121] Leclercq, S., Rivals, E., and Jarne, P. (2010). DNA slippage occurs at microsatellite loci without minimal threshold length in humans: a comparative genomic approach. *Genome Biology and Evolution*, 2:325–335.
- [122] Leeksa, O., De Miranda, N., and Veelken, H. (2017). Germline mutations predisposing to diffuse large b-cell lymphoma. *Blood Cancer Journal*, 7(2):e532–e532.
- [123] Lehmann, A. R., McGibbon, D., and Stefanini, M. (2011). Xeroderma pigmentosum. *Orphanet Journal of Rare Diseases*, 6(1):70.
- [124] Lejeune, J., Turpin, R., and Gautier, M. (1959). Le mongolisme, premier exemple d’aberration autosomique humaine. *Ann Genet*, 1(4):1–49.
- [125] Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., O’Donnell-Luria, A. H., Ware, J. S., Hill, A. J., Cummings, B. B., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, 536(7616):285.
- [126] Lenz, C., Haerty, W., and Golding, G. B. (2014). Increased substitution rates surrounding low-complexity regions within primate proteins. *Genome Biology and Evolution*, 6(3):655–665.
- [127] Lercher, M. J. and Hurst, L. D. (2002). Human SNP variability and mutation rate are higher in regions of high recombination. *Trends in Genetics*, 18(7):337–340.
- [128] Li, C. and Luscombe, N. M. (2019). Nucleosome positioning stability is a significant modulator of germline mutation rate variation across the human genome. *BioRxiv*.

- [129] Li, W.-H., Wu, C.-I., and Luo, C.-C. (1984). Nonrandomness of point mutation as reflected in nucleotide substitutions in pseudogenes and its evolutionary implications. *Journal of Molecular Evolution*, 21(1):58–71.
- [130] Lindsay, S. J., Rahbari, R., Kaplanis, J., Keane, T., and Hurles, M. E. (2019). Similarities and differences in patterns of germline mutation between mice and humans. *Nature Communications*, 10(1):1–12.
- [131] Liu, P., Yuan, B., Carvalho, C. M., Wuster, A., Walter, K., Zhang, L., Gambin, T., Chong, Z., Campbell, I. M., Akdemir, Z. C., et al. (2017). An organismal CNV mutator phenotype restricted to early human development. *Cell*, 168(5):830–842.
- [132] Longley, M. J., Nguyen, D., Kunkel, T. A., and Copeland, W. C. (2001). The fidelity of human DNA polymerase  $\gamma$  with and without exonucleolytic proofreading and the p55 accessory subunit. *Journal of Biological Chemistry*, 276(42):38555–38562.
- [133] Lord, J., McMullan, D. J., Eberhardt, R. Y., Rinck, G., Hamilton, S. J., Quinlan-Jones, E., Prigmore, E., Keelagher, R., Best, S. K., Carey, G. K., et al. (2019). Prenatal exome sequencing analysis in fetal structural anomalies detected by ultrasonography (PAGE): a cohort study. *The Lancet*, 393(10173):747–757.
- [134] Lupski, J. R., de Oca-Luna, R. M., Slaugenhaupt, S., Pentao, L., Guzzetta, V., Trask, B. J., Saucedo-Cardenas, O., Barker, D. F., Killian, J. M., Garcia, C. A., et al. (1991). DNA duplication associated with charcot-marie-tooth disease type 1a. *Cell*, 66(2):219–232.
- [135] Lynch, M. (2010). Evolution of the mutation rate. *TRENDS in Genetics*, 26(8):345–352.
- [136] Maher, G. J., McGowan, S. J., Giannoulatou, E., Verrill, C., Goriely, A., and Wilkie, A. O. (2016). Visualizing the origins of selfish de novo mutations in individual seminiferous tubules of human testes. *Proceedings of the National Academy of Sciences*, 113(9):2454–2459.
- [137] Maher, G. J., Ralph, H. K., Ding, Z., Koelling, N., Mlcochova, H., Giannoulatou, E., Dhami, P., Paul, D. S., Stricker, S. H., Beck, S., et al. (2018). Selfish mutations dysregulating RAS-MAPK signaling are pervasive in aged human testes. *Genome Research*, 28(12):1779–1790.
- [138] Makova, K. D. and Hardison, R. C. (2015). The effects of chromatin organization on variation in mutation rates in the genome. *Nature Reviews Genetics*, 16(4):213–223.
- [139] Marchetti, F., Rowan-Carroll, A., Williams, A., Polyzos, A., Berndt-Weis, M. L., and Yauk, C. L. (2011). Sidestream tobacco smoke is a male germ cell mutagen. *Proceedings of the National Academy of Sciences*, 108(31):12811–12814.
- [140] Martin, H. C., Jones, W. D., McIntyre, R., Sanchez-Andrade, G., Sanderson, M., Stephenson, J. D., Jones, C. P., Handsaker, J., Gallone, G., Bruntraeger, M., et al. (2018). Quantifying the contribution of recessive coding variation to developmental disorders. *Science*, 362(6419):1161–1164.

- [141] Martincorena, I., Raine, K. M., Gerstung, M., Dawson, K. J., Haase, K., Van Loo, P., Davies, H., Stratton, M. R., and Campbell, P. J. (2017). Universal patterns of selection in cancer and somatic tissues. *Cell*, 171(5):1029–1041.
- [142] Mathieson, I. and Reich, D. (2017). Differences in the rare variant spectrum among human populations. *PLoS Genetics*, 13(2):e1006581.
- [143] McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9):1297–1303.
- [144] McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R., Thormann, A., Flicek, P., and Cunningham, F. (2016). The ensembl variant effect predictor. *Genome Biology*, 17(1):122.
- [145] Miccoli, L., Burr, K. L., Hickenbotham, P., Friedberg, E. C., Angulo, J. F., and Dubrova, Y. E. (2007). The combined effects of xeroderma pigmentosum c deficiency and mutagens on mutation rates in the mouse germ line. *Cancer Research*, 67(10):4695–4699.
- [146] Michaelson, J. J., Shi, Y., Gujral, M., Zheng, H., Malhotra, D., Jin, X., Jian, M., Liu, G., Greer, D., Bhandari, A., et al. (2012). Whole-genome sequencing in autism identifies hot spots for de novo germline mutation. *Cell*, 151(7):1431–1442.
- [147] Millar, C. B., Guy, J., Sansom, O. J., Selfridge, J., MacDougall, E., Hendrich, B., Keightley, P. D., Bishop, S. M., Clarke, A. R., and Bird, A. (2002). Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. *Science*, 297(5580):403–405.
- [148] Miller, J. A., Ding, S.-L., Sunkin, S. M., Smith, K. A., Ng, L., Szafer, A., Ebbert, A., Riley, Z. L., Royall, J. J., Aiona, K., et al. (2014). Transcriptional landscape of the prenatal human brain. *Nature*, 508(7495):199–206.
- [149] Mills, R. E., Pittard, W. S., Mullaney, J. M., Farooq, U., Creasy, T. H., Mahurkar, A. A., Kemeza, D. M., Strassler, D. S., Ponting, C. P., Webber, C., et al. (2011). Natural genetic variation caused by small insertions and deletions in the human genome. *Genome Research*, 21(6):830–839.
- [150] Monckton, D. G., Neumann, R., Guram, T., Fretwell, N., Tamaki, K., MacLeod, A., and Jeffreys, A. J. (1994). Minisatellite mutation rate variation associated with a flanking DNA sequence polymorphism. *Nature Genetics*, 8(2):162–170.
- [151] Montgomery, S. B., Goode, D. L., Kvikstad, E., Albers, C. A., Zhang, Z. D., Mu, X. J., Ananda, G., Howie, B., Karczewski, K. J., Smith, K. S., et al. (2013). The origin, evolution, and functional impact of short insertion–deletion variants identified in 179 human genomes. *Genome Research*, 23(5):749–761.
- [152] Moorjani, P., Amorim, C. E. G., Arndt, P. F., and Przeworski, M. (2016). Variation in the molecular clock of primates. *Proceedings of the National Academy of Sciences*, 113(38):10607–10612.

- [153] Mugal, C. F., von Grünberg, H.-H., and Peifer, M. (2009). Transcription-induced mutational strand bias and its effect on substitution rates in human genes. *Molecular Biology and Evolution*, 26(1):131–142.
- [154] Muller, H. J. (1928). The measurement of gene mutation rate in drosophila, its high variability, and its dependence upon temperature. *Genetics*, 13(4):279.
- [155] Nachman, M. W. (2004). Haldane and the first estimates of the human mutation rate. *Journal of Genetics*, 83(3):231–233.
- [156] Nachman, M. W. and Crowell, S. L. (2000). Estimate of the mutation rate per nucleotide in humans. *Genetics*, 156(1):297–304.
- [157] Neale, B. M., Kou, Y., Liu, L., Ma’Ayan, A., Samocha, K. E., Sabo, A., Lin, C.-F., Stevens, C., Wang, L.-S., Makarov, V., et al. (2012). Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature*, 485(7397):242–245.
- [158] Neveling, K., Feenstra, I., Gilissen, C., Hoefsloot, L. H., Kamsteeg, E.-J., Mensenkamp, A. R., Rodenburg, R. J., Yntema, H. G., Spruijt, L., Vermeer, S., et al. (2013). A post-hoc comparison of the utility of sanger sequencing and exome sequencing for the diagnosis of heterogeneous diseases. *Human Mutation*, 34(12):1721–1726.
- [159] Nik-Zainal, S., Alexandrov, L. B., Wedge, D. C., Van Loo, P., Greenman, C. D., Raine, K., Jones, D., Hinton, J., Marshall, J., Stebbings, L. A., et al. (2012). Mutational processes molding the genomes of 21 breast cancers. *Cell*, 149(5):979–993.
- [160] Ohashi, E., Ogi, T., Kusumoto, R., Iwai, S., Masutani, C., Hanaoka, F., and Ohmori, H. (2000). Error-prone bypass of certain DNA lesions by the human DNA polymerase  $\kappa$ . *Genes & Development*, 14(13):1589–1594.
- [161] Okajima, K., Warman, M. L., Byrne, L. C., and Kerr, D. S. (2006). Somatic mosaicism in a male with an exon skipping mutation in PDHA1 of the pyruvate dehydrogenase complex results in a milder phenotype. *Molecular Genetics and Metabolism*, 87(2):162–168.
- [162] O’roak, B., Stessman, H., Boyle, E., Witherspoon, K., Martin, B., Lee, C., Vives, L., Baker, C., Hiatt, J., Nickerson, D., et al. (2014). Recurrent de novo mutations implicate novel genes underlying simplex autism risk. *Nature Communications*, 5(1):1–6.
- [163] O’Roak, B. J., Deriziotis, P., Lee, C., Vives, L., Schwartz, J. J., Girirajan, S., Karakoc, E., MacKenzie, A. P., Ng, S. B., Baker, C., et al. (2011). Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nature Genetics*, 43(6):585.
- [164] Petljak, M., Alexandrov, L. B., Brammied, J. S., Price, S., Wedge, D. C., Grossmann, S., Dawson, K. J., Ju, Y. S., Iorio, F., Tubio, J. M., et al. (2019). Characterizing mutational signatures in human cancer cell lines reveals episodic APOBEC mutagenesis. *Cell*, 176(6):1282–1294.
- [165] Phillips, D. H. (2018). Mutational spectra and mutational signatures: Insights into cancer aetiology and mechanisms of DNA damage and repair. *DNA Repair*, 71:6–11.

- [166] Pich, O., Muiños, F., Lolkema, M. P., Steeghs, N., Gonzalez-Perez, A., and Lopez-Bigas, N. (2019). The mutational footprints of cancer therapies. *Nature Genetics*, 51(12):1732–1740.
- [167] Pino, M. S., Mino-Kenudson, M., Wildemore, B. M., Ganguly, A., Batten, J., Sperduti, I., Iafrate, A. J., and Chung, D. C. (2009). Deficient DNA mismatch repair is common in Lynch syndrome-associated colorectal adenomas. *The Journal of Molecular Diagnostics*, 11(3):238–247.
- [168] Pinto, Y., Gabay, O., Arbiza, L., Sams, A. J., Keinan, A., and Levanon, E. Y. (2016). Clustered mutations in hominid genome evolution are consistent with APOBEC3G enzymatic activity. *Genome Research*, 26(5):579–587.
- [169] Pippard, E. C., Hall, A. J., Barker, D. J., and Bridges, B. A. (1988). Cancer in homozygotes and heterozygotes of ataxia-telangiectasia and xeroderma pigmentosum in britain. *Cancer Research*, 48(10):2929–2932.
- [170] Plant, K. E., Boye, E., Green, P. M., Vetrie, D., and Flintner, F. A. (2000). Somatic mosaicism associated with a mild Alport syndrome phenotype. *Journal of Medical Genetics*, 37(3):238–239.
- [171] Polak, P., Karlić, R., Koren, A., Thurman, R., Sandstrom, R., Lawrence, M. S., Reynolds, A., Rynes, E., Vlahoviček, K., Stamatoyannopoulos, J. A., et al. (2015). Cell-of-origin chromatin organization shapes the mutational landscape of cancer. *Nature*, 518(7539):360–364.
- [172] Porta-Pardo, E., Kamburov, A., Tamborero, D., Pons, T., Grases, D., Valencia, A., Lopez-Bigas, N., Getz, G., and Godzik, A. (2017). Comparison of algorithms for the detection of cancer drivers at subgene resolution. *Nature Methods*, 14(8):782–788.
- [173] Rahbari, R., Wuster, A., Lindsay, S. J., Hardwick, R. J., Alexandrov, L. B., Al Turki, S., Dominiczak, A., Morris, A., Porteous, D., Smith, B., et al. (2016). Timing, rates and spectra of human germline mutation. *Nature Genetics*, 48(2):126.
- [174] Ramu, A., Noordam, M. J., Schwartz, R. S., Wuster, A., Hurles, M. E., Cartwright, R. A., and Conrad, D. F. (2013). DeNovoGear: de novo indel and point mutation discovery and phasing. *Nature Methods*, 10(10):985.
- [175] Redon, R., Ishikawa, S., Fitch, K. R., Feuk, L., Perry, G. H., Andrews, T. D., Fiegler, H., Shapero, M. H., Carson, A. R., Chen, W., et al. (2006). Global variation in copy number in the human genome. *Nature*, 444(7118):444–454.
- [176] Reijns, M. A., Rabe, B., Rigby, R. E., Mill, P., Astell, K. R., Lettice, L. A., Boyle, S., Leitch, A., Keighren, M., Kilanowski, F., et al. (2012). Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. *Cell*, 149(5):1008–1022.
- [177] Retterer, K., Juusola, J., Cho, M. T., Vitazka, P., Millan, F., Gibellini, F., Vertino-Bell, A., Smaoui, N., Neidich, J., Monaghan, K. G., et al. (2016). Clinical application of whole-exome sequencing across clinical indications. *Genetics in Medicine*, 18(7):696–704.

- [178] Rimmer, A., Phan, H., Mathieson, I., Iqbal, Z., Twigg, S. R., Wilkie, A. O., McVean, G., Lunter, G., Consortium, W., et al. (2014). Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nature Genetics*, 46(8):912.
- [179] Rizzato, F., Rodriguez, A., and Laio, A. (2016). Non-markovian effects on protein sequence evolution due to site dependent substitution rates. *BMC Bioinformatics*, 17(1):258.
- [180] Roach, J. C., Glusman, G., Smit, A. F., Huff, C. D., Hubley, R., Shannon, P. T., Rowen, L., Pant, K. P., Goodman, N., Bamshad, M., et al. (2010). Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science*, 328(5978):636–639.
- [181] Roberts, S. A., Lawrence, M. S., Klimczak, L. J., Grimm, S. A., Fargo, D., Stojanov, P., Kiezun, A., Kryukov, G. V., Carter, S. L., Saksena, G., et al. (2013). An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nature Genetics*, 45(9):970.
- [182] Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., and Mesirov, J. P. (2011). Integrative genomics viewer. *Nature Biotechnology*, 29(1):24–26.
- [183] Ropers, H. H. (2010). Genetics of early onset cognitive impairment. *Annual Review of Genomics and Human Genetics*, 11:161–187.
- [184] Russell, W. (1951). X-ray-induced mutations in mice. In *Cold Spring Harbor symposia on quantitative biology*, volume 16, pages 327–336. Cold Spring Harbor Laboratory Press.
- [185] Russell, W. (1989). Reminiscences of a mouse specific-locus test addict. *Environmental and Molecular Mutagenesis*, 14(S16):16–22.
- [186] Samocha, K. E., Kosmicki, J. A., Karczewski, K. J., O’Donnell-Luria, A. H., Pierce-Hoffman, E., MacArthur, D. G., Neale, B. M., and Daly, M. J. (2017). Regional missense constraint improves variant deleteriousness prediction. *BioRxiv*, page 148353.
- [187] Samocha, K. E., Robinson, E. B., Sanders, S. J., Stevens, C., Sabo, A., McGrath, L. M., Kosmicki, J. A., Rehnström, K., Mallick, S., Kirby, A., et al. (2014). A framework for the interpretation of de novo mutation in human disease. *Nature Genetics*, 46(9):944.
- [188] Sandmann, S., De Graaf, A. O., Karimi, M., Van Der Reijden, B. A., Hellström-Lindberg, E., Jansen, J. H., and Dugas, M. (2017). Evaluating variant calling tools for non-matched next-generation sequencing data. *Scientific Reports*, 7:43169.
- [189] Sasani, T. A., Pedersen, B. S., Gao, Z., Baird, L., Przeworski, M., Jorde, L. B., and Quinlan, A. R. (2019). Large, three-generation human families reveal post-zygotic mosaicism and variability in germline mutation accumulation. *elife*, 8.
- [190] Scally, A. and Durbin, R. (2012). Revising the human mutation rate: implications for understanding human evolution. *Nature Reviews Genetics*, 13(10):745–753.
- [191] Schrider, D. R., Hourmozdi, J. N., and Hahn, M. W. (2011). Pervasive multinucleotide mutational events in eukaryotes. *Current Biology*, 21(12):1051–1054.

- [192] Searle, A. (1984). The specific locus test in the mouse. In *Handbook of mutagenicity test procedures*, pages 371–391. Elsevier Amsterdam.
- [193] Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., Yamrom, B., Yoon, S., Krasnitz, A., Kendall, J., et al. (2007). Strong association of de novo copy number mutations with autism. *Science*, 316(5823):445–449.
- [194] Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P., Månér, S., Massa, H., Walker, M., Chi, M., et al. (2004). Large-scale copy number polymorphism in the human genome. *Science*, 305(5683):525–528.
- [195] Séguirel, L., Wyman, M. J., and Przeworski, M. (2014). Determinants of mutation rate variation in the human germline. *Annual Review of Genomics and Human Genetics*, 15:47–70.
- [196] Seidman, M. M., Bredberg, A., Seetharam, S., and Kraemer, K. H. (1987). Multiple point mutations in a shuttle vector propagated in human cells: evidence for an error-prone DNA polymerase activity. *Proceedings of the National Academy of Sciences*, 84(14):4944–4948.
- [197] Seoighe, C. and Scally, A. (2017). Inference of candidate germline mutator loci in humans from genome-wide haplotype data. *PLoS Genetics*, 13(1):e1006549.
- [198] Sheridan, E., Wright, J., Small, N., Corry, P. C., Oddie, S., Whibley, C., Petherick, E. S., Malik, T., Pawson, N., McKinney, P. A., et al. (2013). Risk factors for congenital anomaly in a multiethnic birth cohort: an analysis of the born in bradford study. *The Lancet*, 382(9901):1350–1359.
- [199] Shihab, H. A., Rogers, M. F., Gough, J., Mort, M., Cooper, D. N., Day, I. N., Gaunt, T. R., and Campbell, C. (2015). An integrative approach to predicting the functional effects of non-coding and coding sequence variation. *Bioinformatics*, 31(10):1536–1543.
- [200] Shlien, A., Tabori, U., Marshall, C. R., Pienkowska, M., Feuk, L., Novokmet, A., Nanda, S., Druker, H., Scherer, S. W., and Malkin, D. (2008). Excessive genomic DNA copy number variation in the Li–Fraumeni cancer predisposition syndrome. *Proceedings of the National Academy of Sciences*, 105(32):11264–11269.
- [201] Short, P. J., McRae, J. F., Gallone, G., Sifrim, A., Won, H., Geschwind, D. H., Wright, C. F., Firth, H. V., FitzPatrick, D. R., Barrett, J. C., et al. (2018). De novo mutations in regulatory elements in neurodevelopmental disorders. *Nature*, 555(7698):611–616.
- [202] Singh, T., Walters, J. T., Johnstone, M., Curtis, D., Suvisaari, J., Torniaainen, M., Rees, E., Iyegbe, C., Blackwood, D., McIntosh, A. M., et al. (2017). The contribution of rare variants to risk of schizophrenia in individuals with and without intellectual disability. *Nature Genetics*, 49(8):1167.
- [203] Smedley, D., Schubach, M., Jacobsen, J. O., Köhler, S., Zemojtel, T., Spielmann, M., Jäger, M., Hochheiser, H., Washington, N. L., McMurry, J. A., et al. (2016). A whole-genome analysis framework for effective identification of pathogenic regulatory variants in Mendelian disease. *The American Journal of Human Genetics*, 99(3):595–606.



- [204] Smith, T. C., Arndt, P. F., and Eyre-Walker, A. (2018). Large scale variation in the rate of germ-line de novo mutation, base composition, divergence and diversity in humans. *PLoS Genetics*, 14(3):e1007254.
- [205] Soemedi, R., Wilson, I. J., Bentham, J., Darlay, R., Töpf, A., Zelenika, D., Cosgrove, C., Setchfield, K., Thornborough, C., Granados-Riveron, J., et al. (2012). Contribution of global rare copy-number variants to the risk of sporadic congenital heart disease. *The American Journal of Human Genetics*, 91(3):489–501.
- [206] Soojin, V. Y. (2013). Morris goodman’s hominoid rate slowdown: The importance of being neutral. *Molecular Phylogenetics and Evolution*, 66(2):569–574.
- [207] Stadler, L. (1932). On the genetic nature of induced mutations in plants, reprinted from the proceedings of the sixth international congress of genetics, vol. 1.
- [208] Stamatoyannopoulos, J. A., Adzhubei, I., Thurman, R. E., Kryukov, G. V., Mirkin, S. M., and Sunyaev, S. R. (2009). Human mutation rate associated with DNA replication timing. *Nature Genetics*, 41(4):393.
- [209] Stone, J. E., Lujan, S. A., and Kunkel, T. A. (2012). DNA polymerase zeta generates clustered mutations during bypass of endogenous DNA lesions in *saccharomyces cerevisiae*. *Environmental and Molecular Mutagenesis*, 53(9):777–786.
- [210] Sun, J. X., Helgason, A., Masson, G., Ebenesersdóttir, S. S., Li, H., Mallick, S., Gnerre, S., Patterson, N., Kong, A., Reich, D., et al. (2012). A direct characterization of human mutation based on microsatellites. *Nature Genetics*, 44(10):1161.
- [211] Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., et al. (2019). String v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, 47(D1):D607–D613.
- [212] Tate, J. G., Bamford, S., Jubb, H. C., Sondka, Z., Beare, D. M., Bindal, N., Boutselakis, H., Cole, C. G., Creatore, C., Dawson, E., et al. (2019). COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Research*, 47(D1):D941–D947.
- [213] Tawn, E. J. (2002). Hereditary effects of radiation: Unsear 2001 report to the general assembly, with scientific annex.
- [214] The, I., of Whole, T. P.-C. A., Consortium, G., et al. (2020). Pan-cancer analysis of whole genomes. *Nature*, 578(7793):82.
- [215] The Deciphering Developmental Disorder Study (2015). Large-scale discovery of novel genetic causes of developmental disorders. *Nature*, 519(7542):223–228.
- [216] Tissier, A., McDonald, J. P., Frank, E. G., and Woodgate, R. (2000). *polI*, a remarkably error-prone human DNA polymerase. *Genes & Development*, 14(13):1642–1650.
- [217] Tubbs, A. and Nussenzweig, A. (2017). Endogenous DNA damage as a source of genomic instability in cancer. *Cell*, 168(4):644–656.

- [218] Turner, D. J., Miretti, M., Rajan, D., Fiegler, H., Carter, N. P., Blayney, M. L., Beck, S., and Hurlles, M. E. (2008). Germline rates of de novo meiotic deletions and duplications causing several genomic disorders. *Nature Genetics*, 40(1):90.
- [219] Uchimura, A., Higuchi, M., Minakuchi, Y., Ohno, M., Toyoda, A., Fujiyama, A., Miura, I., Wakana, S., Nishino, J., and Yagi, T. (2015). Germline mutation rates and the long-term phenotypic effects of mutation accumulation in wild-type laboratory mice and mutator mice. *Genome Research*, 25(8):1125–1134.
- [220] Van Dongen, S. (2008). Graph clustering via a discrete uncoupling process. *SIAM Journal on Matrix Analysis and Applications*, 30(1):121–141.
- [221] Veltman, J. A. and Brunner, H. G. (2012). De novo mutations in human genetic disease. *Nature Reviews Genetics*, 13(8):565–575.
- [222] Venkat, A., Hahn, M. W., and Thornton, J. W. (2018). Multinucleotide mutations cause false inferences of lineage-specific positive selection. *Nature Ecology & Evolution*, 2(8):1280.
- [223] Venkatarajan, M. S. and Braun, W. (2001). New quantitative descriptors of amino acids based on multidimensional scaling of a large number of physical–chemical properties. *Molecular Modeling Annual*, 7(12):445–453.
- [224] Villegas, F., Lehalle, D., Mayer, D., Rittirsch, M., Stadler, M. B., Zinner, M., Olivieri, D., Vabres, P., Duplomb-Jego, L., De Bont, E. S., et al. (2019). Lysosomal signaling licenses embryonic stem cell differentiation via inactivation of TFE3. *Cell stem cell*, 24(2):257–270.
- [225] Vissers, L. E., van Ravenswaaij, C. M., Admiraal, R., Hurst, J. A., de Vries, B. B., Janssen, I. M., van der Vliet, W. A., Huys, E. H., de Jong, P. J., Hamel, B. C., et al. (2004). Mutations in a new member of the chromodomain gene family cause charge syndrome. *Nature Genetics*, 36(9):955–957.
- [226] Watterson, G. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, 7(2):256–276.
- [227] Webb, A. J., Berg, I. L., and Jeffreys, A. (2008). Sperm cross-over activity in regions of the human genome showing extreme breakdown of marker association. *Proceedings of the National Academy of Sciences*, 105(30):10471–10476.
- [228] Wei, L., Liu, L. T., Conroy, J. R., Hu, Q., Conroy, J. M., Morrison, C. D., Johnson, C. S., Wang, J., and Liu, S. (2015). MAC: identifying and correcting annotation for multi-nucleotide variations. *BMC Genomics*, 16(1):569.
- [229] Weinstein, L. S., Shenker, A., Gejman, P. V., Merino, M. J., Friedman, E., and Spiegel, A. M. (1991). Activating mutations of the stimulatory G protein in the McCune–Albright syndrome. *New England Journal of Medicine*, 325(24):1688–1695.
- [230] Wells, A., Heckerman, D., Torkamani, A., Yin, L., Sebat, J., Ren, B., Telenti, A., and di Iulio, J. (2019). Ranking of non-coding pathogenic variants and putative essential regions of the human genome. *Nature Communications*, 10(1):1–9.

- [231] Weren, R. D., Ligtenberg, M. J., Kets, C. M., De Voer, R. M., Verwiel, E. T., Spruijt, L., van Zelst-Stams, W. A., Jongmans, M. C., Gilissen, C., Hehir-Kwa, J. Y., et al. (2015). A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer. *Nature Genetics*, 47(6):668.
- [232] Wong, E., Yang, K., Kuraguchi, M., Werling, U., Avdievich, E., Fan, K., Fazzari, M., Jin, B., Brown, A. M., Lipkin, M., et al. (2002). MBD4 inactivation increases c>t transition mutations and promotes gastrointestinal tumor formation. *Proceedings of the National Academy of Sciences*, 99(23):14937–14942.
- [233] Wong, J. T.-F. (1975). A co-evolution theory of the genetic code. *Proceedings of the National Academy of Sciences of the United States of America*, 72(5):1909.
- [234] Wong, W. S., Solomon, B. D., Bodian, D. L., Kothiyal, P., Eley, G., Huddleston, K. C., Baker, R., Thach, D. C., Iyer, R. K., Vockley, J. G., et al. (2016). New observations on maternal age effect on germline de novo mutations. *Nature Communications*, 7(1):1–10.
- [235] Wright, C., Prigmore, E., Rajan, D., Handsaker, J., McRae, J., Kaplanis, J., Fitzgerald, T., FitzPatrick, D., Firth, H., and Hurles, M. (2019). Clinically-relevant postzygotic mosaicism in parents and children with developmental disorders in trio exome sequencing data. *Nature Communications*, 10(1):1–11.
- [236] Wright, C. F., Fitzgerald, T. W., Jones, W. D., Clayton, S., McRae, J. F., Van Kogelenberg, M., King, D. A., Ambridge, K., Barrett, D. M., Bayzeturina, T., et al. (2015). Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *The Lancet*, 385(9975):1305–1314.
- [237] Wright, C. F., FitzPatrick, D. R., and Firth, H. V. (2018). Paediatric genomics: diagnosing rare disease in children. *Nature Reviews Genetics*, 19(5):253.
- [238] Wu, X., Pang, E., Lin, K., and Pei, Z.-M. (2013). Improving the measurement of semantic similarity between gene ontology terms and gene products: insights from an edge-and ic-based hybrid method. *PloS One*, 8(5).
- [239] Xu, B., Roos, J. L., Levy, S., Van Rensburg, E., Gogos, J. A., and Karayiorgou, M. (2008). Strong association of de novo copy number mutations with sporadic schizophrenia. *Nature Genetics*, 40(7):880.
- [240] Yang, S., Wang, L., Huang, J., Zhang, X., Yuan, Y., Chen, J.-Q., Hurst, L. D., and Tian, D. (2015). Parent–progeny sequencing indicates higher mutation rates in heterozygotes. *Nature*, 523(7561):463–467.
- [241] Yang, Y., Muzny, D. M., Xia, F., Niu, Z., Person, R., Ding, Y., Ward, P., Braxton, A., Wang, M., Buhay, C., et al. (2014). Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA*, 312(18):1870–1879.
- [242] Yauk, C. L., Berndt, M. L., Williams, A., Rowan-Carroll, A., Douglas, G. R., and Stämpfli, M. R. (2007). Mainstream tobacco smoke causes paternal germ-line DNA mutation. *Cancer Research*, 67(11):5103–5106.

- [243] Yilmaz, R., Beleza-Meireles, A., Price, S., Oliveira, R., Kubisch, C., Clayton-Smith, J., Szakszon, K., and Borck, G. (2015). A recurrent synonymous KAT6B mutation causes say-barber-biesecker/young-simpson syndrome by inducing aberrant splicing. *American Journal of Medical Genetics Part A*, 167(12):3006–3010.
- [244] Young, L. C., Hartig, N., del Río, I. B., Sari, S., Ringham-Terry, B., Wainwright, J. R., Jones, G. G., McCormick, F., and Rodriguez-Viciano, P. (2018). SHOC2–MRAS–PP1 complex positively regulates RAF activity and contributes to noonan syndrome pathogenesis. *Proceedings of the National Academy of Sciences*, 115(45):E10576–E10585.
- [245] Zhang, F., Gu, W., Hurles, M. E., and Lupski, J. R. (2009). Copy number variation in human health, disease, and evolution. *Annual Review of Genomics and Human Genetics*, 10:451–481.
- [246] Zhou, J. and Troyanskaya, O. G. (2015). Predicting effects of noncoding variants with deep learning–based sequence model. *Nature Methods*, 12(10):931–934.