

REVIEW

Open Access

Genomics-driven drug discovery based on disease-susceptibility genes



Kyuto Sonehara¹ and Yukinori Okada^{1,2,3*}

Abstract

Genome-wide association studies have identified numerous disease-susceptibility genes. As knowledge of gene-disease associations accumulates, it is becoming increasingly important to translate this knowledge into clinical practice. This challenge involves finding effective drug targets and estimating their potential side effects, which often results in failure of promising clinical trials. Here, we review recent advances and future perspectives in genetics-led drug discovery, with a focus on drug repurposing, Mendelian randomization, and the use of multifaceted omics data.

Background

Since the first completion of human genome sequencing in 2003 [1], many more attempts have been made to elucidate the relationships between human genotypes and phenotypes. One of the approaches that has been widely adopted for this purpose is the genome-wide association study (GWAS) [2, 3]. GWAS is an observational study that is designed to statistically assess associations between traits and tens of millions of genome-wide genetic variants from population samples. Due to the advancement of genotyping technology using single-nucleotide polymorphisms (SNP) microarray, more than 4000 GWASs have been reported globally at the time of this writing [4]. With the increase in the number of studies, the number of samples in each study has also increased, reaching hundreds of thousands of samples in recent years [5–9]. Although these GWASs have identified numerous trait-associated genomic loci, it is still challenging to translate these findings into clinical practice. In this review, we summarize recent advances in disease-susceptibility genes for drug discovery applications.

Main text

Significance of the genetic evidence for drug discovery

Despite the tremendous effort and substantial resources dedicated to biomedical research, only a handful of promising academic discoveries have led to new treatments [10]. Such a gap between basic research and clinical practice is a challenge for the entire field of biomedical research and is often referred to as the “valley of death.” One of the causes of this gap is the biological differences between human and other model organisms [11–13]. Validation in other organisms, such as mice, does not necessarily mean that the results will be replicated in humans. In addition, although validation using human samples is preferable, experiments using cell lines do not reflect systemic effects [14], and interventional clinical trials are at times ethically unfeasible. Investigating the impact of human genetic variation on phenotypes can provide insight into pathophysiology in the human body, which will lead to the discovery of true drug targets. Actually, it is known that drug targets with genetic evidence are more likely to be passed into the Phase III trial or market [15].

One effective approach for enhancing clinical practice is drug repurposing. Drug repurposing is a strategy for finding novel indications for existing approved drugs or drugs in clinical trials [16]. If the safety of the drug has already been confirmed in early-stage clinical trials for its original purpose, repurposing existing drugs requires

* Correspondence: yokada@sg.med.osaka-u.ac.jp

¹Department of Statistical Genetics, Osaka University Graduate School of Medicine, 2-2 Yamadaoka, Suita 565-0871, Japan

²Laboratory of Statistical Immunology, Immunology Frontier Research Center (WPI-IFReC), Osaka University, Suita 565-0871, Japan

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

less cost for testing the safety than developing and implementing novel drugs. The information on disease-susceptibility genes for drug repurposing has been successfully exploited. By utilizing databases of existing approved drug-target genes and protein–protein interactions (PPIs), Okada et al. demonstrated that GWAS-identified rheumatoid arthritis (RA)-susceptibility genes were significantly correlated with the targets of known RA drugs, such as TNF-inhibitors [17] and JAK inhibitors [18], via the PPI networks [19]. This study further revealed that CDK4 and CDK6, which are targets of approved cancer drugs, are potential therapeutic targets for RA. The efficacy of CDK4/6 inhibitors was experimentally validated in animal models of RA [20, 21]. In another example, Imamura et al. demonstrated a significant association in the connectivity of existing drug-target genes and biological type 2 diabetes (T2D) risk genes in PPI networks [22]. They identified KIF11 inhibitor (originally indicated for several types of cancers), GSK3B inhibitor (originally indicated for several types of cancers), and AP-1 inhibitor (originally indicated for RA) as potential candidates for a repurposed treatment of T2D.

The studies that have been described so far mainly focused on the association of genes and drugs for a single disease. The use of existing drug classification systems is a promising approach to systematically assess gene–drug associations for a wide variety of diseases. Malik et al. utilized the Anatomical Therapeutic Chemical Classification System (ATC) to extract multiple drug–disease associations [23]. ATC is a drug classification system in which drugs are classified according to the organ or system on which the active substances act and their therapeutic, pharmacological, and chemical properties. Malik et al. evaluated overlapping between GWAS-identified stroke-susceptibility genes and known drug targets, finding that stroke-risk genes are significantly targeted by drugs that are classified into ATC B: “Blood and blood forming organs,” specifically a subcategory, ATC B01: “Antithrombotic agents.” To perform the analysis as described above for a gene set, the freely available software GREP (Genome for REPositioning drugs) by Sakaue et al., is useful [24]. GREP quantifies the association of user-defined gene sets with the categories of existing drugs, such as the ATC or International Classification of Diseases diagnostic code. It further suggests drugs that have potential in repurposing to target the given gene set.

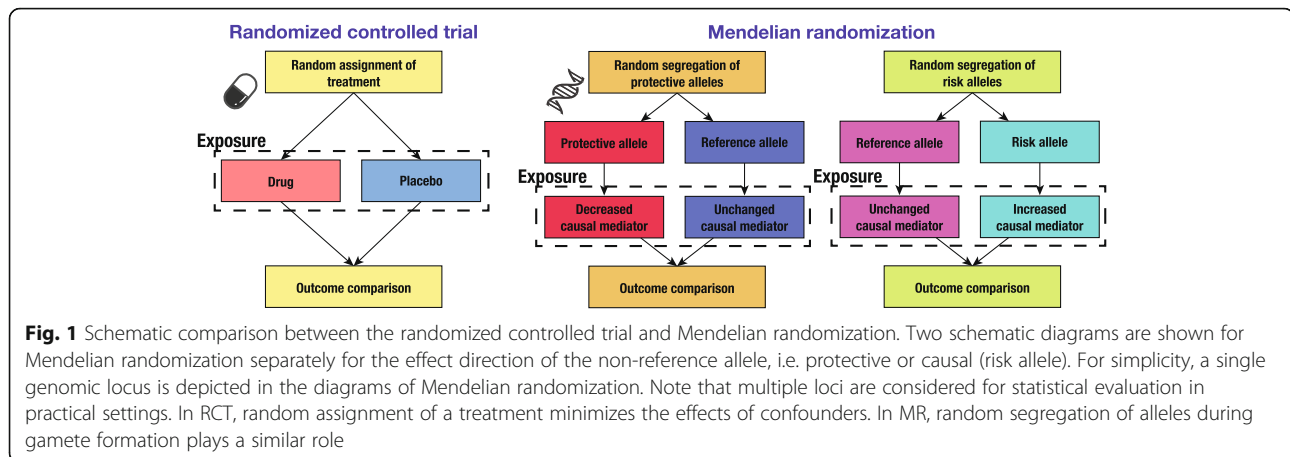
Given the need for novel drug development, prioritizing genes as therapeutic targets is a decisive step. Feng et al. devised a pipeline for this prioritization, named “priority index” (Pi), by integrating genome-scale data, disease ontologies, and PPIs [25]. To identify the genes responsible for the GWAS signals, they utilized not only

disease-associated GWAS signals but also chromatin marks and expression quantitative trait locus (eQTL) signals. eQTLs are genomic loci that are associated with mRNA expression levels. Then, the list of candidate risk genes was extended to the genes that interact with the directly observed disease-susceptible genes via the PPI networks. Feng et al. applied a Pi pipeline to 16 immunologic traits, finding that 15 of these analytic results were significantly enriched in the targets of approved medications for the corresponding traits. These pioneering studies demonstrate that insight into disease-susceptibility genes are a powerful resource for more efficient drug discovery, and that integration of other biological data is also a key to drug discovery.

Mendelian randomization for identifying drug targets

Suitable targets for therapeutic medications are not limited to genes. Other substances, such as modified proteins or metabolites, are also related to disease states [26], and such substances are called biomarkers. However, biomarkers do not necessarily play a causal role in the disease pathology, as they can be influenced by the disease states or other causes that induce disease states. One approach to solving such a causality problem with the help of genetics is Mendelian randomization (MR) [27, 28]. MR is a genetic epidemiological framework for causal inference between an exposure (i.e., biomarker) and an outcome (i.e., disease state), as if a randomized controlled trial (RCT) had been conducted [29]. Since genotypes are assigned almost independently of environment when they are inherited from parents, those who have genotypes that increase exposure are, in effect, assigned a high dosage of the exposure, independent of other confounding factors. This situation is analogous to that of an RCT. MR provides virtual RCT opportunities without actual intervention (Fig. 1).

Sjaarda et al. conducted a systematic MR analysis of 237 biomarkers to identify the causal mediators of coronary artery disease (CAD) [30]. They found six biomarkers that are suspected to increase the risk of CAD (lipoprotein[a], apolipoprotein E, interleukin-6 receptor, stromal cell–derived factor 1 [CXCL12], apolipoprotein C3, and macrophage colony-stimulating factor 1 [CSF1]). Of these, CXCL12 and CSF1 were novel findings, and higher levels of both biomarkers were linked to an increased risk of CAD. They further utilized MR to estimate whether these candidate causal mediators affect other biomarkers of CAD risk factors, revealing an increasing effect of CSF1 on C-reactive protein levels. They inspected a causal effect of interleukin-1 beta (IL-1 β) on CXCL12 and CSF1, indicating that IL-1 β is causally related to CSF1 levels. As this study shows, MR analysis facilitates identification of causal relationships among various biomarkers as well as genetic variation.



Chong et al. performed a systematic MR analysis of the human proteome to identify novel causal mediators of stroke [31]. They screened 653 circulating proteins, identifying 7 potentially causal biomarkers (histo-blood group ABO system transferase, coagulation factor XI, scavenger receptor class A5 [SCARA5], tumor necrosis factor-like weak inducer of apoptosis [TNFSF12], cluster of differentiation 40, apolipoprotein[a], and matrix metalloproteinase-12). SCARA5 and TNFSF12 had an especially protective effect on cardioembolic stroke. To assess whether these two potential drug targets for stroke adversely affect other traits, Chong et al. further performed a phenome-wide MR analysis of 679 disease traits. TNFSF12 was revealed to be deleterious for four circulatory system phenotypes, three digestive phenotypes, and one injuries and poisonings phenotype, which suggests that TNFSF12-targeted treatment may cause such diseases. In contrast, SCARA5 had no significant associations with those phenotypes other than having a protective effect on subarachnoid hemorrhage. They reported SCARA5 as a promising target for the treatment of cardioembolic stroke. This study demonstrates the capability of MR to reveal novel therapeutic targets and also elucidate probable side effects.

Future perspectives

As discussed, insight into disease-susceptibility genes are translated into clinical practice more effectively when combined with other biological resources, such as the PPI network, transcriptome, and proteome. This is because genetic information provides clues about the causal relationships among multiple traits, which can result in distinct correlations discovered from simple observational studies [32]. Hence, enhancement of multifaceted biological resources will lead to further advances in genetics-led drug discovery. For example, expansion of metabolome studies may reveal disease-

causal metabolites through MR analysis, which could expand the range of candidate therapeutic targets [33]. Another potential therapeutic target is microbiota [34, 35]. The interaction between human organs and their microbial composition has received increasing attention [36]. Linking human microbial knowledge to GWAS insight is opening up new perspectives [37–39]. As previous studies show, investigating tissue-specific gene functions is an essential approach for the development of therapeutic targets [40, 41]. Integration of insight into disease-susceptibility genes and tissue-specific biological features will lead to precise strategies for treatment [42–46]. Recent advances in single-cell analysis will further reveal cell type-specific effects of genetic variation [47–49] and provide precise descriptions of relationships between genotypes and phenotypes.

Another influential factor that will provide advancements in this field is the decreasing cost of whole-genome sequencing. Nowadays, we can sequence an individual's whole genome for less than 1000 U.S. dollars [50], which enables us to investigate the effect of rare variants on phenotypes, whereas most GWASs focus mainly on common variants. Because functional variants are subject to purifying selection, such variants tend to be rare in most populations. In other words, rare variants are more likely to be functional than common variants [51, 52]. By collecting such functional rare variants and phenotypes of carriers of those variants, we can more clearly grasp the functional effect of genotypes on phenotypes [53]. A prominent example is “human knockouts” [54, 55]. If individuals who have a homozygous loss-of-function variant within a gene are found in a population, inspecting their phenotypes closely will reveal how an individual is affected by the inhibited gene. Such observations will lead to the discovery of novel drug targets and provide an estimation of its side effects [56–58].

Conclusions

Advances in genotyping technologies, including SNP microarray and next-generation sequencing, have yielded numerous studies concerning the relationships between the genome and a wide range of traits. The next goal for genetics is translating these insights into clinical practice. The increasing number of attempts to achieve this goal includes drug repurposing, prioritization of candidate target genes, and MR-based causal inference. Future discoveries through these efforts will lead to solutions to the present problems that challenge drug development.

Abbreviations

SNP: Single-nucleotide polymorphism; PPI: Protein-protein interaction; ATC: Anatomical therapeutic chemical classification system; MR: Mendelian randomization

Acknowledgements

Not applicable

Authors' contributions

KS wrote the manuscript. YO revised the manuscript. All authors read and approved the final manuscript.

Funding

KS was supported by the Takeda Science Foundation and Integrated Frontier Research for Medical Science Division, Institute for Open and Transdisciplinary Research Initiatives, Osaka University.

YO was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (19H01021, 20K21834), and AMED (JP20km0405211, JP20ek0109413, JP20ek0410075, JP20gm4010006, and JP20km0405217), the Takeda Science Foundation, and Bioinformatics Initiative of Osaka University Graduate School of Medicine, Osaka University.

Availability of data and materials

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Statistical Genetics, Osaka University Graduate School of Medicine, 2-2 Yamadaoka, Suita 565-0871, Japan. ²Laboratory of Statistical Immunology, Immunology Frontier Research Center (WPI-IFReC), Osaka University, Suita 565-0871, Japan. ³Integrated Frontier Research for Medical Science Division, Institute for Open and Transdisciplinary Research Initiatives, Osaka University, Suita 565-0871, Japan.

Received: 17 December 2020 Accepted: 26 February 2021

Published online: 10 March 2021

References

- Collins FS, Morgan M, Patrinos A. The human genome project: lessons from large-scale biology. *Science*. 2003;300:286–90.
- Ozaki K, Ohnishi Y, Iida A, Sekine A, Yamada R, Tsunoda T, et al. Functional SNPs in the lymphotoxin- α gene that are associated with susceptibility to myocardial infarction. *Nat Genet*. 2002;32:650–4.
- Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, et al. 10 years of GWAS discovery: biology, function, and translation. *Am J Hum Genet*. 2017;101:5–22.
- GWAS Catalog. The European Bioinformatics Institute, Hinxton. <https://www.ebi.ac.uk/gwas/>. Accessed 10 Oct 2020.
- Marouli E, Graff M, Medina-Gomez C, Lo KS, Wood AR, Kjaer TR, et al. Rare and low-frequency coding variants alter human adult height. *Nature*. 2017; 542:186–90.
- Akiyama M, Ishigaki K, Sakaue S, Momozawa Y, Horikoshi M, Hirata M, et al. Characterizing rare and low-frequency height-associated variants in the Japanese population. *Nat Commun*. 2019;10:1–11.
- Ishigaki K, Akiyama M, Kanai M, Takahashi A, Kawakami E, Sugishita H, et al. Large-scale genome-wide association study in a Japanese population identifies novel susceptibility loci across different diseases. *Nat Genet*. 2020; 52:669–79.
- Vujkovic M, Keaton JM, Lynch JA, Miller DR, Zhou J, Tcheandjieu C, et al. Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis. *Nat Genet*. 2020;52:680–91.
- Sakaue S, Kanai M, Tanigawa Y, Karjalainen J, Kurki M, Koshihara S, et al. A global atlas of genetic associations of 220 deep phenotypes. *medRxiv*. 2020. <https://doi.org/10.1101/2020.10.23.20213652>.
- Butler D. Translational research: crossing the valley of death. *Nature*. 2008; 453:840–2.
- Seok J, Warren HS, Cuenca AG, Mindrinos MN, Baker HV, Xu W, et al. Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proc Natl Acad Sci U S A*. 2013;110:3507–12.
- Shay T, Jovic V, Zuk O, Rothamel K, Puyraimond-Zemmour D, Feng T, et al. Conservation and divergence in the transcriptional programs of the human and mouse immune systems. *Proc Natl Acad Sci*. 2013;110:2946–51.
- Mak IW, Evaniew N, Ghert M. Lost in translation: animal models and clinical trials in cancer treatment. *Am J Transl Res*. 2014;6:114–8.
- Gillet J-P, Varma S, Gottesman MM. The clinical relevance of cancer cell lines. *JNCI J Natl Cancer Inst*. 2013;105:452–8.
- Nelson MR, Tipney H, Painter JL, Shen J, Nicoletti P, Shen Y, et al. The support of human genetic evidence for approved drug indications. *Nat Genet*. 2015;47:856–60.
- Pushpakom S, Iorio F, Eyers PA, Escott KJ, Hopper S, Wells A, et al. Drug repurposing: progress, challenges and recommendations. *Nat Rev Drug Discov*. 2019;18:41–58.
- Porter D, van Melckebeke J, Dale J, Messow CM, McConnachie A, Walker A, et al. Tumour necrosis factor inhibition versus rituximab for patients with rheumatoid arthritis who require biological treatment (ORBIT): an open-label, randomised controlled, non-inferiority, trial. *The Lancet*. 2016;388:239–47.
- Taylor PC, Keystone EC, van der Heijde D, Weinblatt ME, del Carmen ML, Reyes Gonzaga J, et al. Baricitinib versus Placebo or Adalimumab in Rheumatoid Arthritis. *N Engl J Med*. 2017;376:652–62.
- Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature*. 2014;506:376–81.
- Sekine C, Sugihara T, Miyake S, Hirai H, Yoshida M, Miyasaka N, et al. Successful Treatment of Animal Models of Rheumatoid Arthritis with Small-Molecule Cyclin-Dependent Kinase Inhibitors. *J Immunol*. 2008;180:1954–61.
- Hosoya T, Iwai H, Yamaguchi Y, Kawahata K, Miyasaka N, Kohsaka H. Cell cycle regulation therapy combined with cytokine blockade enhances antiarthritic effects without increasing immune suppression. *Ann Rheum Dis*. 2016;75:253–9.
- Imamura M, Takahashi A, Yamauchi T, Hara K, Yasuda K, Grarup N, et al. Genome-wide association studies in the Japanese population identify seven novel loci for type 2 diabetes. *Nat Commun*. 2016;7:1–12.
- Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, et al. Multi-ancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet*. 2018;50:524–37.
- Sakaue S, Okada Y. GREP: genome for REPositioning drugs. *Bioinformatics*. 2019;35:3821–3.
- Fang H, Wolf HD, Knezevic B, Burnham KL, Osgood J, Sanniti A, et al. A genetics-led approach defines the drug target landscape of 30 immune-related traits. *Nat Genet*. 2019;51:1082–91.
- Tanaka K, Yamagata K, Kubo S, Nakayama S, Sakata K, Matsui T, et al. Glycolaldehyde-modified advanced glycation end-products inhibit differentiation of human monocytes into osteoclasts via upregulation of IL-10. *Bone*. 2019;128:115034.

27. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003;32:1–22.
28. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet*. 2014;23:R89–98.
29. Thanassoulis G, O'Donnell CJ. Mendelian Randomization: Nature's Randomized Trial in the Post-Genome Era. *JAMA*. 2009;301:2386.
30. Sjaarda J, Gerstein H, Chong M, Yusuf S, Meyre D, Anand SS, et al. Blood CSF1 and CXCL12 as Causal Mediators of Coronary Artery Disease. *J Am Coll Cardiol*. 2018;72:300–10.
31. Michael C, Jennifer S, Marie P, Pedrum M-S, Ricky L, Ashkan S, et al. Novel drug targets for ischemic stroke identified through mendelian randomization analysis of the blood proteome. *Circulation*. 2019;140:819–30.
32. Sakaue S, Kanai M, Karjalainen J, Akiyama M, Kurki M, Matoba N, et al. Transbiobank analysis with 676,000 individuals elucidates the association of polygenic risk scores of complex traits with human lifespan. *Nat Med*. 2020;26:542–8.
33. Hartiala JA, Wilson Tang WH, Wang Z, Crow AL, Stewart AFR, Roberts R, et al. Genome-wide association study and targeted metabolomics identifies sex-specific association of CPS1 with coronary artery disease. *Nat Commun*. 2016;7:1–10.
34. Costello SP, Hughes PA, Waters O, Bryant RV, Vincent AD, Blatchford P, et al. Effect of Fecal Microbiota Transplantation on 8-Week Remission in Patients With Ulcerative Colitis: A Randomized Clinical Trial. *JAMA*. 2019;321:156–64.
35. Morita N, Umemoto E, Fujita S, Hayashi A, Kikuta J, Kimura I, et al. GPR31-dependent dendrite protrusion of intestinal CX3CR1+ cells by bacterial metabolites. *Nature*. 2019;566:110–4.
36. Nagashima K, Sawa S, Nitta T, Tsutsumi M, Okamura T, Penninger JM, et al. Identification of subepithelial mesenchymal cells that induce IgA and diversify gut microbiota. *Nat Immunol*. 2017;18:675–82.
37. Kishikawa T, Maeda Y, Nii T, Motoooka D, Matsumoto Y, Matsushita M, et al. Metagenome-wide association study of gut microbiome revealed novel aetiology of rheumatoid arthritis in the Japanese population. *Ann Rheum Dis*. 2019. <https://doi.org/10.1136/annrheumdis-2019-215743>.
38. Kishikawa T, Maeda Y, Nii T, Okada Y. Response to: "Can sexual dimorphism in rheumatoid arthritis be attributed to the different abundance of *Gardnerella*?" by Liu et al. *Ann Rheum Dis*. 2020. <https://doi.org/10.1136/annrheumdis-2020-217264>.
39. Kishikawa T, Maeda Y, Nii T, Okada Y. The positive correlation between *Porphyromonas gingivalis* and *Prevotella* spp. Response to: "Comment on 'Metagenome-wide association study of gut microbiome revealed novel aetiology of rheumatoid arthritis in the Japanese population' by Kishikawa et al." by Kitamura et al. *Ann Rheum Dis*. 2020. <https://doi.org/10.1136/annrheumdis-2020-217897>.
40. Jin Y, Tachibana I, Takeda Y, He P, Kang S, Suzuki M, et al. Statins decrease lung inflammation in mice by upregulating tetraspanin CD9 in macrophages. *PLoS One*. 2013;8:e73706.
41. Morita K, Okamura T, Inoue M, Komai T, Teruya S, Iwasaki Y, et al. Egr2 and Egr3 in regulatory T cells cooperatively control systemic autoimmunity through Ltbp3-mediated TGF- β 3 production. *Proc Natl Acad Sci U S A*. 2016;113:E8131–40.
42. Ishigaki K, Kochi Y, Suzuki A, Tsuchida Y, Tsuchiya H, Sumitomo S, et al. Polygenic burdens on cell-specific pathways underlie the risk of rheumatoid arthritis. *Nat Genet*. 2017;49:1120–5.
43. Kanai M, Akiyama M, Takahashi A, Matoba N, Momozawa Y, Ikeda M, et al. Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nat Genet*. 2018;50:390–400.
44. Sakaue S, Hirata J, Maeda Y, Kawakami E, Nii T, Kishikawa T, et al. Integration of genetics and miRNA–target gene network identified disease biology implicated in tissue specificity. *Nucleic Acids Res*. 2018;46:11898–909.
45. Ohkura N, Yasumizu Y, Kitagawa Y, Tanaka A, Nakamura Y, Motoooka D, et al. Regulatory T Cell-Specific Epigenomic Region Variants Are a Key Determinant of Susceptibility to Common Autoimmune Diseases. *Immunity*. 2020;52:1119–32 e4.
46. Koido M, Kawakami E, Fukumura J, Noguchi Y, Otori M, Nio Y, et al. Polygenic architecture informs potential vulnerability to drug-induced liver injury. *Nat Med*. 2020;26:1–8.
47. van der Wijst MGP, Brugge H, de Vries DH, Deelen P, Swertz MA, Franke L. Single-cell RNA sequencing identifies celltype-specific cis-eQTLs and co-expression QTLs. *Nat Genet*. 2018;50:493–7.
48. Chignon A, Bon-Baret V, Boulanger M-C, Li Z, Argaud D, Bossé Y, et al. Single-cell expression and Mendelian randomization analyses identify blood genes associated with lifespan and chronic diseases. *Commun Biol*. 2020;3:1–15.
49. Kim-Hellmuth S, Aguet F, Oliva M, Muñoz-Aguirre M, Kasela S, Wucher V, et al. Cell type-specific genetic regulation of gene expression across human tissues. *Science*. 2020;369. <https://doi.org/10.1126/science.aaz8528>.
50. DNA Sequencing Costs: Data. The National Human Genome Research Institute. 2020. <https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data>. Accessed 24 Feb 2020.
51. Zhu Q, Ge D, Maia JM, Zhu M, Petrovski S, Dickson SP, et al. A Genome-wide Comparison of the Functional Properties of Rare and Common Genetic Variants in Humans. *Am J Hum Genet*. 2011;88:458–68.
52. Long T, Hicks M, Yu H-C, Biggs WH, Kirkness EF, Menni C, et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. *Nat Genet*. 2017;49:568–78.
53. Whiffin N, Armean IM, Kleinman A, Marshall JL, Minikel EV, Goodrich JK, et al. The effect of LRRK2 loss-of-function variants in humans. *Nat Med*. 2020;26:869–77.
54. Perdigoto C. Dawn of the Human Knockout Project. *Nat Rev Genet*. 2017;18:328–9.
55. Saleheen D, Natarajan P, Armean IM, Zhao W, Rasheed A, Khetarpal SA, et al. Human knockouts and phenotypic analysis in a cohort with a high rate of consanguinity. *Nature*. 2017;544:235–9.
56. Khan SS, Shah SJ, Klyachko E, Baldrige AS, Eren M, Place AT, et al. A null mutation in SERPINE1 protects against biological aging in humans. *Sci Adv*. 2017;3:eaao1617.
57. McGregor TL, Hunt KA, Yee E, Mason D, Nioi P, Ticao S, et al. Characterising a healthy adult with a rare HAO1 knockout to support a therapeutic strategy for primary hyperoxaluria. *eLife*. 2020;9:e54363.
58. Minikel EV, Karczewski KJ, Martin HC, Cummings BB, Whiffin N, Rhodes D, et al. Evaluating drug targets through human loss-of-function genetic variation. *Nature*. 2020;581:459–64.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

