

Review

Cite this article: Alamad B, Elliott K and Knight JC (2024). Cross-population applications of genomics to understand the risk of multifactorial traits involving inflammation and immunity. *Cambridge Prisms: Precision Medicine*, 2, e3, 1–11
<https://doi.org/10.1017/pcm.2023.25>

Received: 18 August 2023

Revised: 15 November 2023

Accepted: 18 December 2023

Keywords:

genomics; immune; personalized medicine; disease risk; non-Caucasian

Corresponding author:

Julian C. Knight;

Email: julian.knight@well.ox.ac.uk

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.



Cross-population applications of genomics to understand the risk of multifactorial traits involving inflammation and immunity

Bana Alamad¹ , Kate Elliott¹ and Julian C. Knight^{1,2}

¹Wellcome Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, Oxford, UK and ²Chinese Academy of Medical Science Oxford Institute, Nuffield Department of Medicine, University of Oxford, Oxford, UK

Abstract

The interplay between genetic and environmental factors plays a significant role in inter-individual variation in immune and inflammatory responses. The availability of high-throughput low-cost genotyping and next-generation sequencing has revolutionized our ability to identify human genetic variation and understand how this varies within and between populations, and the relationship with disease. In this review, we explore the potential of genomics for patient benefit, specifically in the diagnosis, prognosis and treatment of inflammatory and immune-related diseases. We summarize the knowledge arising from genetic and functional genomic approaches, and the opportunity for personalized medicine. The review covers applications in infectious diseases, rare immunodeficiencies and autoimmune diseases, illustrating advances in diagnosis and understanding risk including use of polygenic risk scores. We further explore the application for patient stratification and drug target prioritization. The review highlights a key challenge to the field arising from the lack of sufficient representation of genetically diverse populations in genomic studies. This currently limits the clinical utility of genetic-based diagnostic and risk-based applications in non-Caucasian populations. We highlight current genome projects, initiatives and biobanks from diverse populations and how this is being used to improve healthcare globally by improving our understanding of genetic susceptibility to diseases and regional pathogens such as malaria and tuberculosis. Future directions and opportunities for personalized medicine and wider application of genomics in health care are described, for the benefit of individual patients and populations worldwide.

Impact statement

This review provides a comprehensive overview of advances in genomics within the context of immune-related diseases. It critically examines the interplay between genetic and environmental factors in determining the risk of susceptibility to immune diseases, and emphasizes the importance of considering genetic variation across diverse populations to enhance our understanding of disease etiology and allow for global application of personalized medicine. The review highlights the potential applications of genetics and wider functional genomic approaches in enhancing diagnosis, disease risk prediction, patient stratification and prioritizing drug targets, using a wide range of inflammatory and immune-related diseases as examples. It further addresses specific challenges associated with current genomic-based approaches for personalized medicine. It discusses the opportunity and limitations of polygenic risk scores, emphasizing the need for comprehensive research that encompasses genetic variation across different ethnicities and geographical regions. By critically evaluating current translational applications of genomics, the review identifies future priorities for utilizing genomic medicine for patient benefit and sheds light on ways to foster a more comprehensive and diverse scientific approach that can enable future clinical application to individuals from all ethnic backgrounds and geographical locations.

Introduction

Heritable factors play a significant role in interindividual variation in immune and inflammatory responses. These contribute to susceptibility to disease, from very rare highly penetrant germline sequence variants causing monogenic primary immunodeficiencies (PIDs), to more common variants contributing to polygenic traits such as seen in autoimmunity or infectious diseases (Figure 1). Malaria and other pathogens have been a major selective pressure on human genetic architecture, with specific alleles driven to high frequency or fixation in some populations (Kwiatkowski, 2005; Kwok et al., 2021). While this may serve to protect individuals

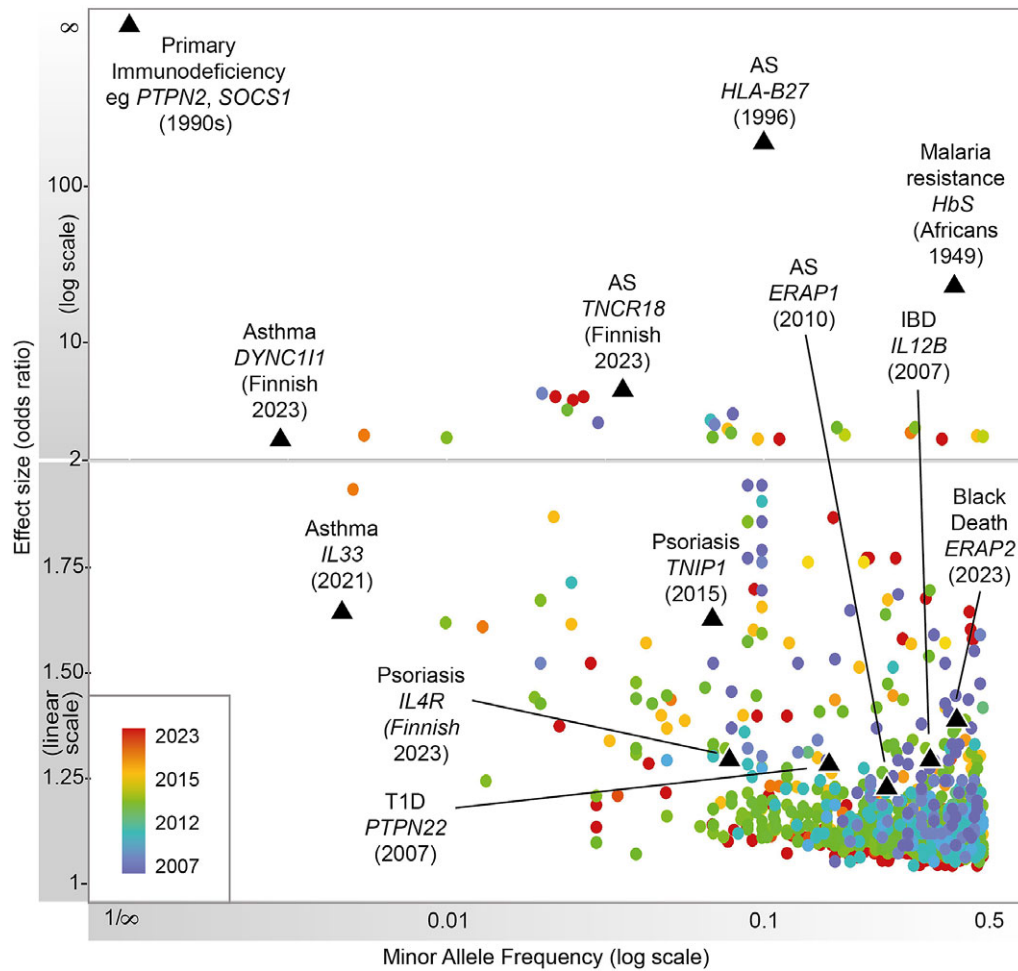


Figure 1. Progression of immune-related disease gene discovery over time, showing the increased power to discover variants of lower effect size with the use of larger cohorts. Black triangles refer to variants described in this review.

in high exposure environments, there may be a cost in terms of autoimmune risk such as seen with some human leukocyte antigen (HLA) alleles (Dendrou et al., 2018). The availability of high-throughput low-cost genotyping, and more recently next-generation sequencing (NGS), has revolutionized our ability to catalog such human genetic variation and understand the relationship with disease risk (Lappalainen et al., 2019). Genotyping more common single nucleotide variants has enabled genome-wide association studies (GWASs) with the identification of disease risk loci providing the opportunity for new insights into disease mechanisms, validation of drug targets and generation of polygenic risk scores (PRSs) (Sudlow et al., 2015; Tam et al., 2019; Kurki et al., 2023). NGS has significantly increased the number of known PID genes and substantially reduced the diagnostic odyssey for individual patients, while long-read technologies are allowing HLA typing at unprecedented resolution (Zhang et al., 2021; Redmond et al., 2022).

To date, our knowledge of disease genetics primarily arises from European populations (Phillips et al., 2021). High costs in establishing lab and computational infrastructure, coupled with limited expertise and training opportunities in genomics analysis, contribute to this disparity in developing countries. Budget priorities in many regions typically favor primary healthcare over research. Nevertheless, the paucity of genetic analysis in other

populations, particularly those with a higher burden of infectious disease, may lead to failure to identify risk loci and their underlying disease mechanisms, or appropriately interpret previously reported risk variants in different populations or geographical contexts, and contribute to regional inequality in the potential for genomic medicine (Figure 2). There is a global imperative to increase the representation of ethnically diverse populations in genetic studies to address knowledge disparities, reveal new disease mechanisms and ensure that genomic knowledge can be applied in an equitable and applicable way for all (Stuart et al., 2022). Mapping genetic associations in diverse populations with different genetic architectures addresses the constraints of linkage disequilibrium, which may limit the ability to fine-map specific causal alleles due to the coinheritance of variants. This allows the full spectrum of disease alleles to be defined, including where these may only be at a detectable frequency in specific populations, and enabling the identification of population-specific risk alleles, ensuring that diagnostic NGS panels and PRSs can be meaningfully applied.

Here, we will explore the potential for genomics and personalized medicine in the diagnosis, prognosis and treatment of inflammatory and immune-related diseases. We summarize the knowledge accumulated from genetic and multi-omic studies of a broad range of diseases, including immune-mediated traits and

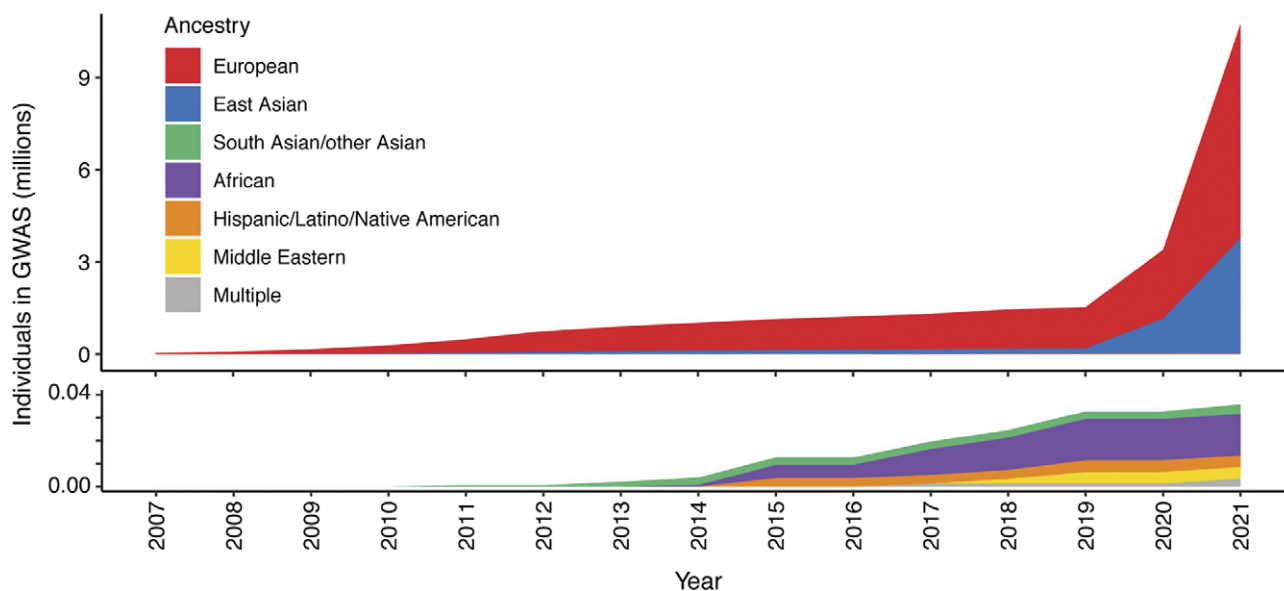


Figure 2. Ancestry composition of GWAS studies for 16 autoimmune traits according to GWAS catalog, per year since 2007. Disparities in the representation of global populations in GWAS show the focus on populations of European descent, with studies of other populations significantly under-represented until relatively recently. The lower panel zooms in on the y-axis scale to show the disproportionately low number of ethnically diverse GWAS studies. Figure adapted from Khunsiraksakul et al. (2022) licensed under CC BY 4.0.

infectious diseases across diverse populations, and how this is being exploited to improve healthcare globally.

Genome variation

Large-scale population biobanks, including clinical, pathological, molecular and radiological records, have been valuable for the dissection of immune-mediated diseases, which often show overlap of clinical features and genetic pleiotropy. Globally, over 120 biobanks and population-scale datasets have now been assembled, including UK Biobank (UKB, $n = 500,000$) (Sudlow et al., 2015), Genomics England ($n = 100,000$) in the United Kingdom, FinnGen ($n = 400,000+$) in Finland (Kurki et al., 2023), deCODE genetics ($n = 500,000$) in Iceland (Gulcher & Stefansson, 1999), China Kadoorie Biobank ($n = >510,000$) (Chen et al., 2005) and Biobank Japan ($n = 260,000$) (Kubo, 2017). In addition to common variant association, the large size of these cohorts has allowed the identification of rarer, higher impact variants, such as the identification of an association between a rare splice in *IL33* and asthma in UKB exome data (Backman et al., 2021).

Isolated populations that experienced recent bottlenecks, like the Finnish population, provide the opportunity to identify deleterious variants, too rare to be detected in other populations. Using FinnGen data (<http://www.finnngen.fi/en>), for 15 common diseases, 30 new disease associations were identified that were mostly low-frequency variants enriched in the Finnish population (Kurki et al., 2023). Of these, immune trait-associated variants were found near *TNCR18*, associated with inflammatory bowel disease (IBD) and ankylosing spondylitis (AS), *IL4R* associated with psoriasis (PSO) and asthma and *TNCR1* associated with PSO. This demonstrates the power of bottlenecked populations to uncover previously unknown biology of immune diseases.

Disease-specific biobanks that include tissue banking provide opportunities to generate functional genomic data at high resolution to complement genetic mapping efforts. For example, the

COVID-19 autopsy biobank (Delorey et al., 2021) has been used to study the effect of severe SARS-CoV-2 infection on COVID-19 pathogenesis in different organs using multi-omic information. Here, 420 autopsy tissue samples from 11 organs were used to generate a single-cell atlas of heart, liver, kidney and lung from donors who died of COVID-19. This atlas along with lung spatial atlases presented in other studies (Melms et al., 2021; Wang et al., 2021) revealed changes in the transcriptional landscapes and cell type proportions throughout the course of infection, providing information about severe disease pathogenesis.

Capturing global population genetic variation is critical to understanding genetic susceptibility to pathogens such as malaria and tuberculosis where the burden of disease is greatest in low- and middle-income countries. Established in 2002, the International HapMap project was one of the first endeavors to catalog human genetic variation (The International HapMap Project, 2003). The aim was to map the most prevalent patterns of DNA variation (haplotypes) across the human genome by genotyping 1.3 million SNPs in 270 individuals from Africa, Europe and Asia. Subsequently, the 1,000 Genome Project (1000G), which whole-genome sequenced 2,504 individuals from worldwide populations has become a key reference resource for global human genetic variation (1000 Genomes Project Consortium, 2010). Many other population-specific genome projects have now been established including Qatar (Zayed, 2016; Mbarek et al., 2022), Saudi Arabia (Team, 2015), Iceland (Gudbjartsson et al., 2015), Uganda (Gurdasani et al., 2019), Singapore (Wu et al., 2019), Brazil (Naslavsky et al., 2022) and Japan (Okada et al., 2018).

Widely recognized as the origin of modern humans, Africans represent the most genetically diverse of human populations, as well as carrying the greatest burden of infectious disease mortality. The Human Heredity and Health in Africa (H3Africa) consortium (<https://h3africa.org>) (Rotimi et al., 2014) is facilitating African-led research specifically studying African genome architecture to provide insights into causal mutations for monogenic diseases and genetic and environmental risk contributions for multifactorial

diseases and traits. H3ABioNet (Mulder et al., 2016), an initiative arising from H3Africa, is addressing bioinformatics challenges in African genomics research. By promoting access and research into genetic susceptibility to prevalent infectious diseases in Africa, such initiatives are key to addressing healthcare disparities and informing global disease knowledge (Pereira et al., 2021) (Fatumo et al., 2022).

The increasing numbers of studies of common disease genetics in non-European populations has identified new loci predisposing to disease. Disease alleles of moderate or low effect size that are rare in Europeans can be detected if they exist at greater frequencies in other populations. An example is a large meta-analysis GWAS of type 1 diabetes (T1D) adding Finnish, African and East Asian samples to European samples, identifying 36 new loci (Robertson et al., 2021). In another example, in a transethnic meta-analysis, the addition of South Asian samples to European samples identified two novel psoriasis loci (Stuart et al., 2022). Inclusion of African ancestry individuals, in general, provides better opportunities for fine mapping of genetic associations as co-inherited blocks of genetic variants are typically shorter with greater genetic diversity, by contrast to European haplotypes (Ge et al., 2022; Mahajan et al., 2022; Ruan et al., 2022). Such studies have the potential to identify new disease mechanisms and provide a greater understanding of disease etiology which can be used for drug development with a greater global scope.

In summary, utilizing large-scale biobanks, dedicated population genome projects and founding more regional genomics consortia remains crucial in the ongoing identification of both common and rare highly penetrant variants across diverse populations. Region-specific capacity building is therefore crucial. H3Africa and the Malaria Genomic Epidemiology Network (MalariaGEN) exemplified effective capacity building through aiding malaria-endemic countries in designing and implementing ethical research, establishing standardized methods and conducting genetic data analysis (Wonkam & Mayosi, 2014). More initiatives like this are essential to advance genomic research and impact clinical medicine.

Infectious disease

Infectious diseases have been a strong evolutionary selective pressure on observed human genetic variation, which in some instances also modulates other immune-mediated inflammatory disease (IMID) risk. An example of this is the *ERAP2* gene where protective alleles for the Black Death caused by the bacteria *Yersinia pestis* (Klunk et al., 2022) have been found in addition to risk variants for IMIDs including AS (International Genetics of Ankylosing Spondylitis Consortium, 2013), IBD (Liu et al., 2015) and psoriasis (Tsoi et al., 2012). The Black Death is the most lethal pandemic recorded in human history and killed 75–200 million people, up to 30–50% of populations in Afro-Eurasia. Recently, susceptibility variants have been identified in DNA extracted from 516 samples from ancient burial sites in London and Denmark, genotyped at targeted immune genes. Comparison of allele frequencies before and after the Black Death identified multiple variants under selection. The most significant change was seen at an SNP, rs2549794, modulating full-length vs. truncated transcription of the *ERAP2* gene. Individuals homozygous for the protective allele were 40% more likely to survive the Black Death than those homozygous for the deleterious allele (Klunk et al., 2022). It was also demonstrated that the protective allele is associated with increased expression and production

of full-length ERAP2 protein, and macrophages harvested from individuals carrying the selected allele inhibit *Y. pestis* replication, providing a possible mechanism of resistance to the Black Death.

The sickle hemoglobin mutation (HbS), encoded by rs334, and variation in *Plasmodium falciparum* that causes malaria, also exhibit a complex balance of selective pressures between host and parasite. Homozygosity of HbS causes sickle cell anemia, whereas heterozygosity confers tolerance to *P. falciparum* infection (Kariuki & Williams, 2020; Band et al., 2022). Recently, Band et al. identified that protection against malaria conferred by HbS is dependent on parasite genotype. In an analysis of host and parasite variation in Gambian and Kenyan children with severe malaria, variation at 3 loci (Ferreira et al., 2011) in the parasite genome was found to be associated with HbS (Band et al., 2022). These HbS-associated loci include the chr2: 631,190 T > A variant within *PfACS8*, the chr2: 814,288 C > T variant within *Pf3D7_0220300*, and the chr11: 1,058,035 T > A variant within *Pf3D7_1127000*, referred to as *Pfsa1*, *Pfsa2* and *Pfsa3*, respectively. The frequency of these parasitic variants was found to correlate with HbS frequency across populations, being most frequent in Africa, where HbS is the most prevalent. This demonstrates that genetic differences between human populations can lead to different advantageous interactions with infectious parasites.

A further example involving selection is recent work which identified immune genes involved in positive and negative selection in post-Neolithic Europe, demonstrating that the historical resistance to infection and adaptation to pathogens has increased present-day inflammatory disease risk (Kerner et al., 2023). Genomic analysis has also shown how rare variants can contribute to extreme phenotypes. For example, Zhang et al. demonstrated how highly penetrant monogenic inborn errors of *TLR3* and *IRF7* are associated with severe life-threatening COVID-19 (Zhang et al., 2020). This study compared genomic sequencing data from patients diagnosed with life-threatening COVID-19 pneumonia (n = 659) and those with asymptomatic/ benign infection (n = 534). The analysis identified enrichment for rare variants in 13 gene loci involved in *TLR3*- and *IRF7*-dependent immunity, with experimental validation including evidence for a role for autosomal recessive AR deficiencies of *IRF7* and *IFNAR1* in severe COVID-19 (Zhang et al., 2020).

Genomic evidence suggests that pathogen-driven selection targeted immune-related genes contributes to inflammatory disorders (Barreiro et al., 2008; Barreiro & Quintana-Murci, 2010; Matzaraki et al., 2017; Pankratov et al., 2022). This selective pressure has led to the advantageous evolution of host defense genes, contributing to the heightened polymorphism observed in the major histocompatibility complex (MHC) (Leslie et al., 2010). In HIV-1 infection, heterozygosity at MHC class I loci confers an advantage, resulting in a slower progression to AIDS (Carrington et al., 1999). Similarly, in hepatitis B virus (HBV) infection, heterozygotes at MHC class II loci show an increased likelihood of clearing the infection (Thursz et al., 1997).

Overall, the interplay between evolutionary pressures and immune responses underscores the ongoing impact of infectious diseases on genetic landscapes shaping disease susceptibility today.

Application of genomics for immune disease diagnosis and risk

Genomics has proven to be a powerful tool in the diagnosis of rare and complex diseases that pose diagnostic challenges. The use of

genomic sequencing for diagnosis and direct management has been used in many diseases including musculoskeletal, neurological, metabolic and complex syndromes affecting multiple tissues and processes (Ko et al., 2018; Symonds et al., 2021). Some, such as mutations in the *CFTR* gene underlying cystic fibrosis, may be tested for prenatally.

PIDs, typically resulting from highly penetrant rare mutations, are characterized by recurrent potentially life-threatening infections. To date, over 300 monogenic mutations have been identified to cause PID (GEL PanelApp) (Thaventhiran et al., 2020). Phenotypic heterogeneity in PID can make genetic diagnosis challenging (Pan-Hammarström et al., 2007; Lenardo et al., 2016; Thaventhiran et al., 2020), with only 29% of PID patients having a genetic cause of disease identified (Edgar et al., 2014). Sporadic (nonfamilial) PID presenting during adulthood can be particularly difficult to diagnose since they tend to be less severe. Suspected cases will typically undergo whole exome (WES) or genome sequencing (WGS). Non-coding variants in regulatory regions can also contribute to disease phenotypes, with, for example, colocalized novel high-penetrance monogenic variants and common variants (at the *PTPN2* and *SOCS1* loci) reported (Thaventhiran et al., 2020).

Common variable immunodeficiency disorders (CVIDs) are a heterogeneous group of PIDs. Genetic causes have been identified in 5% of CVID patients, with genotypic and phenotypic overlap between CVID and other immunological conditions reflecting the pathophysiological diversity of this disease (Peng et al., 2023). Genomic studies have shown that the majority of CVID cases are polygenic with multiple common low-penetrant variants, while a small subset of CVID patients including early onset monogenic antibody or immune deficiency, are caused by rare highly penetrant monogenic variants (Kienzler et al., 2017). *NFKB1* haploinsufficiency has been reported as the most common monogenic cause of CVID (Tuijnburg et al., 2018), which was confirmed by Thaventhiran et al. (2020).

In terms of IMID risk, GWAS has successfully identified multiple predisposing risk variants and demonstrated a complex polygenic architecture (Lewis & Vassos, 2020). By aggregating these genetic variants, a cumulative PRS can predict disease susceptibility and progression (Khunsriraksakul et al., 2022). PRS is calculated by summing the odds ratios of each risk allele in an individual (Lewis & Vassos, 2020; Khunsriraksakul et al., 2022) assuming an additive genetic architecture and independence between variants, which might not always be precise (Lewis & Vassos, 2020). The increasing number of studies presenting a highly significant association between PRS and disease status supports the potential for utilizing PRS as a clinical instrument, but this clinical function is still not yet robustly established.

For example, AS is a debilitating chronic inflammatory spinal arthritis affecting 20,000 people in the United Kingdom, usually presenting in young adults (Mauro et al., 2021). A hallmark of AS is the involvement of the sacroiliac joints and 85% of cases carry HLA-B27. It is one of the most heritable IMIDs (sibling relative risk ~60), of which HLA-B27 contributes ~30% of genetic risk (Khunsriraksakul et al., 2022). To date >100 loci harboring common variants have been implicated in AS risk, such as in the *ERAP1* gene (Brown et al., 2016). GWAS in AS are now sufficiently powered that PRS can be calculated that can be used to predict individual disease risk. With 78.2% positive and 100% negative predictive power, PRS has been reported to have a higher diagnostic capacity for AS than the traditional diagnostic tools, including clinical features, sacroiliac imaging, and HLA-B27 (Li et al., 2021). PRSs have also been developed for Systemic Lupus

Erythematosus (SLE), a multi-organ autoimmune disease (Rönnblom & Leonard, 2019). PRS has been used to predict late-onset SLE using renal disease as a proxy for severity in both European and Chinese populations (Chen et al., 2020). Another study showed that a high genetics-based risk score is associated with an increased risk of organ damage, renal dysfunction and all-cause mortality, indicating the utility of using such scores to predict disease complications, patient deterioration, and mortality (Reid et al., 2020).

While the use of genetic information to predict risk of disease or potential drug effects can be very beneficial for personalized medicine, the lack of sufficient representation of genetically diverse populations in genomic literature limits the clinical utility of genetic-based prediction approaches like PRS in non-Caucasian populations. The predictive performance of the current PRS is lower in non-European populations. Analysis of PRS studies 2008–2017 showed that the majority of PRS studies (67%) involved European ancestry participants, with only 19% and 3.8% targeting East Asian and African/Hispanic/Indigenous populations, respectively (Duncan et al., 2019). Predictive performance was found to be lower for PRS derived from European ancestry individuals when tested in people of non-European ancestry (Duncan et al., 2019). Consistent with this, PRS derived from African American individuals greatly enhanced the predictive performance of PRS in African populations compared to European ancestry (Kamiza et al., 2022). However, the heterogeneity of African populations revealed some subpopulations performed better than others, for example, African American-derived PRS for lipid traits had a much greater predictive value in South African Zulus compared to Ugandans (Kamiza et al., 2022). The recognition of the value of data from other ethnicities is leading to an increase in data generation that will in turn improve the transferability of PRS globally.

Genetic heterogeneity within the same population can also limit PRS predictive power if the fine-scale structure within this population is not well documented at high resolution. In a Japanese cohort, genome-wide PRS analyses on 67 complex traits were performed. Differences in PRS between the subpopulations within this Japanese cohort did not agree with the observed phenotypes for each subpopulation, suggesting that PRS differences reflect biases due to the uncorrected structure causing prediction biases in a trait-dependent manner and limiting clinical utility in non-Caucasians (Sakaue et al., 2020). Some studies also found individual-level uncertainty in PRS estimates in addition to cohort-level biases (Ding et al., 2022). A recent study by Hingorani et al. (2023) calculated the informative performance metrics for 926 PRSs across 310 diseases from the Polygenic Score Catalog. The analysis reported that PRS has a poor screening performance for individual risk prediction and population risk stratification, and indicated that the high expectations surrounding the potential of PRS for personalized medicine may be overstated.

To address the shortcomings of PRS in diverse populations, Weissbrod et al. proposed a method called PolyPred that improves PRS accuracy across populations by accounting for differences in linkage disequilibrium (Weissbrod et al., 2022). This method improved accuracy for 49 diseases/traits in some UKB populations including South Asians and Africans (Weissbrod et al., 2022). To further improve the transferability of tools like PRS, ethnically diverse genomic datasets need to be used in the development and optimization processes.

Overall, genomics stands as a robust diagnostic tool for PIDs and other rare immune- and inflammation-related diseases. While GWAS studies reveal polygenic influences on more common traits

involving such processes, the predictive potential of PRS is hindered by underrepresentation in non-Caucasian populations. Ongoing efforts are needed to generate comprehensive genomic datasets and optimized methodologies to enhance PRS clinical translation across diverse populations.

Patient stratification and drug target prioritization

Within IMIDs, significant disease heterogeneity is recognized with the opportunity for better stratification using genetics and functional genomics. For example, multiple sclerosis (MS) is a debilitating chronic inflammatory disease affecting the central nervous system leading to irreversible neurological damage including long-term functional impairment (Pinto *et al.*, 2020). Prediction of disease progression is currently difficult. Recent work by the International MS Genetics Consortium and MultipleMS Consortium have identified rs10191329 in the *DYSF-ZNF638* locus (which includes genes involved in damage repair and control of viral infections) as associated with disease severity and specifically shortening of time to requiring a walking aid together with brain stem and cortical changes (International Multiple Sclerosis Genetics Consortium & MultipleMS Consortium, 2023).

Psoriatic arthritis (PsA) is a chronic inflammatory musculoskeletal condition that arises in about 30% of psoriasis patients (Ogdie, 2017). PsA is known to significantly reduce the patient's quality of life and increase mortality, it has various clinical presentations and often goes undiagnosed in psoriasis patients. Single-cell transcriptomics and cell-surface protein expression have been used to compare immune cell populations between PsA and PsC patients, and expression differences were successfully used as markers for psoriasis subtyping through a machine-learning model (Liu *et al.*, 2022). Such studies demonstrate the potential utility of omics techniques in the implementation of personalized medicine.

Functional genomics can be used to stratify patients based on the nature of immune response to infection and the extent to which this is maladaptive. For example, sepsis is a highly heterogeneous clinical syndrome defined as life-threatening organ dysfunction caused by a dysregulated host response to infection (Singer *et al.*, 2016). Transcriptomic profiling of white blood cells has revealed distinct sepsis response signatures (SRSSs) that stratify a poor outcome immune-suppressed group of patients that have underlying differences in neutrophil function and emergency granulopoiesis (Davenport *et al.*, 2016; Kwok *et al.*, 2023). There is initial evidence that SRS is informative to define differential response to steroids (Antcliffe *et al.*, 2019) and is applicable as a quantitative likelihood score across different infectious etiologies (Cano-Gamez *et al.*, 2022) with the opportunity for future point-of-care testing.

Improving the efficiency of drug target selection for development is critical, given the time and cost (\$2b+) of taking a drug through to approval. Attrition is high due to safety concerns or lack of efficacy (DiMasi *et al.*, 2016), but it is recognized that genetic support increases the likelihood of success at least twofold (Nelson *et al.*, 2015). For rare highly penetrant mutations, the modulated gene can be identified with confidence while for GWAS, typically noncoding variants are implicated, and the modulated gene may not be the nearest gene to the associated variant given the complex three-dimensional conformation of the human genome and processes of gene regulation and other mechanisms whereby a variant may exert a functional effect. The Open Targets Initiative (<http://genetics.opentargets.org>) integrates

GWAS data, including from UKB and FinnGen, with transcriptomic, proteomic and epigenetic data from multiple tissues and cell types using fine mapping (Mountjoy *et al.*, 2021). Out of 133,441 published GWAS loci, this pipeline successfully identified 729 loci fine-mapped to a single-coding causal variant, and further drug target prioritization, performed by training a machine-learning model using functional validation data, improved precision and resulted in a 58% reduction in the number of false positives detected in the prioritized loci (Mountjoy *et al.*, 2021). A study contributing data to Open Targets (Soskic *et al.*, 2022) investigated T cell regulation during immune disease by performing single-cell transcriptomic profiling of 655,349 CD4+ T cells in healthy and activated states. Out of 6,407 genes whose expression was correlated with genetic variation between T cell states, 2,265 genes (35%) were dynamically regulated during T cell activation in immune disease providing evidence of the key genes and mechanisms underlying genetic susceptibility (Soskic *et al.*, 2022). Other studies out of the open target initiative include single-cell genomics in asthma (Vieira Braga *et al.*, 2019) and epigenomics identifying changes in T-cell states in immune diseases (Soskic *et al.*, 2019).

Priority Index is a further approach to drug target prioritization for IMIDs from GWAS that provides weighting based on functional genomic evidence such as chromosomal conformation and expression quantitative trait loci in immune cells, and takes account of network connectivity based on high confidence protein-protein interactions with genetically prioritized genes to produce a predictor matrix, with prioritization scores calculated for a given gene for a particular disease where GWAS data have been inputted (Fang *et al.*, 2019). Pathway crosstalk maximizing numbers of highly prioritized genes further identifies potential nodal genes for intervention. The ability to identify currently approved drug targets is enhanced with disease-specific functional genomic annotators, as recently demonstrated for AS (Brown *et al.*, 2023).

Genomics plays a pivotal role in navigating the complexities of IMIDs especially through utilizing functional genomics for nuanced patient stratification. Additionally, it can support efficient drug target selection by integrating GWAS data and fine mapping. Genomics can also play an important role in vaccinology by identifying potential protective antigens and characterizing the interaction between the host, vector and pathogen (de la Fuente & Contreras, 2021). Historically, analysis of well-studied vaccines including HBV (Desombere *et al.*, 1998), rubella (Lambert *et al.*, 2015) and measles (Jacobson *et al.*, 2011) vaccines have presented evidence showing genetic associations between increased antibody responses and the MHC, specifically different alleles involving HLA (Kwok *et al.*, 2021). Genetic factors, specifically HLA type, contribute to interindividual variation in COVID-19 vaccine response and to the risk of breakthrough infection (Mentzer *et al.*, 2023). Higher levels of antibodies against the SARS-CoV-2 spike receptor-binding domain were associated with the carriage of HLA-DQB1*06. Individuals with HLA-DQB1*06 alleles were also less likely to experience PCR-confirmed breakthrough infection (Mentzer *et al.*, 2022). Such information may inform targeting booster vaccinations to the most vulnerable and vaccine design. Similar genetic factors may contribute to such a response in other infectious diseases, providing an opportunity to stratify individuals into groups based on their genetic information to effectively allocate vaccine distribution.

Following the identification of genetic variants, determining their pathogenicity poses challenges, especially for novel variants in disease-associated genes. To address this, computational tools

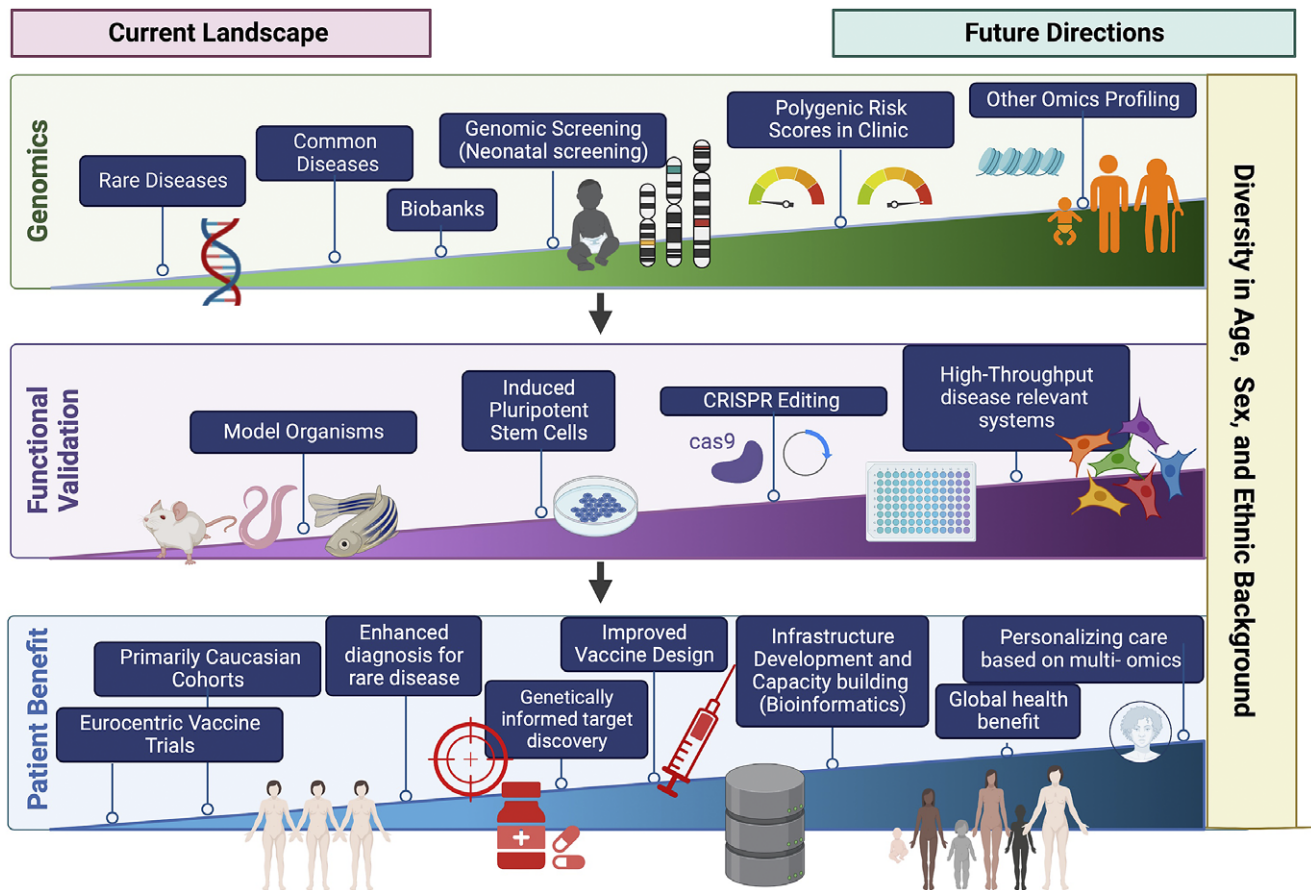


Figure 3. Current landscape and future directions of global genomics applications in personalized medicine highlighting key future areas in diversifying genomic medicine. Figure created with BioRender.com.

like REVEL (Ioannidis et al., 2016), AlphaMissense (Cheng et al., 2023), spliceAI (Jaganathan et al., 2019) and LoFTK (Alasiri et al., 2023) predict pathogenicity for missense, splice and loss-of-function variants while tools like DDMut (Zhou et al., 2023) predict variant effects on protein stability. With the increase in data availability and the development of machine learning and AI-based approaches, the assessment of pathogenicity for novel and known variants will improve. In addressing the challenge of clinical translation in genomics, the NHS Genomic Medicine Service, leveraging the 100,000 Genomes Project (1000 Genomes Project Consortium, 2010), stands as an exemplar to deliver genetic testing for inherited rare diseases and cancer through a National genomic test directory, supra-regional Genomic Laboratory Hubs and associated service alliances in the United Kingdom (NHS England, 2021). This infrastructure enhances the translation of genomic insights into clinical decisions for improved patient outcomes.

Concluding remarks

Over the past few decades, advances in genotyping and sequencing technologies have allowed genomics to become an increasingly important tool in clinical medicine (Figure 3). This has expanded from the genomic diagnosis of rare diseases to the opportunity to apply PRS in predicting disease risk and outcome. However, the relative paucity of genetic studies in non-European populations limits wide applicability, including PRS. Additionally, cost and

infrastructure in many low-income countries remains a barrier to the development of genomic medicine. Advances in genomics are shaping the future of medicine, particularly contributing to the shift from reactive to proactive medicine that focuses on preventing disease rather than treating it. We need to ensure that equitable application to benefit all underpins both discovery and translational genomics with study designs involving diverse population groups worldwide, paired with appropriate knowledge sharing and capacity building.

Open peer review. To view the open peer review materials for this article, please visit <http://doi.org/10.1017/pcm.2023.25>.

Data availability statement. Data availability is not applicable to this article as no new data were created or analyzed in this study.

Acknowledgments. The authors acknowledge Dr. Calliope Dendrou, Dr. Smita Patel and Dr. Paul Wordsworth for engaging in scientific discussions about the topic of the review and providing clinical examples from their respective specialties.

Author contribution. B.A. wrote the first draft of the manuscript, incorporated edits and generated the final manuscript. K.E. edited and redrafted sections of the manuscript. J.C.K. edited the manuscript and oversaw the review. All authors contributed to the conception and structure of the review, read and approved the final manuscript.

Financial support. This work was supported by the Wellcome Trust (WT Reference Number: 204969/Z/16/Z), Chinese Academy of Medical Sciences (CAMS) Innovation 537 Fund for Medical Science (2018-I2M-2-002), the

Medical Research Council (MR/V002503/1) and the NIHR Oxford Biomedical Research Centre.

Competing interest. The authors declare none.

References

- 1000 Genomes Project Consortium** (2010) A map of human genome variation from population scale sequencing. *Nature* **467**(7319), 1061–1073. <https://doi.org/10.1038/nature09534>.
- Alasiri A, Karczewski KJ, Cole B, Loza B-L, Moore JH, van der Laan SW, Asselbergs FW, Keating BJ and van Setten J** (2023) LoFTK: A framework for fully automated calculation of predicted loss-of-function variants and genes. *Biodata Mining* **16**(1), 3. <https://doi.org/10.1186/s13040-023-00321-5>.
- Antcliffe DB, Burnham KL, Al-Beidh F, Santhakumaran S, Brett SJ, Hinds CJ, Ashby D, Knight JC and Gordon AC** (2019) Transcriptomic signatures in sepsis and a differential response to steroids. From the VANISH randomized trial. *American Journal of Respiratory and Critical Care Medicine* **199**(8), 980–986. <https://doi.org/10.1164/rccm.201807-1419OC>.
- Backman JD, Li AH, Marcketta A, Sun D, Mbatouch J, Kessler MD, Benner C, Liu D, Locke AE and Balasubramanian S** (2021) Exome sequencing and analysis of 454,787 UK biobank participants. *Nature* **599**(7886), 628–634. <https://doi.org/10.1038/s41586-021-04103-z>.
- Band G, Leffler EM, Jallow M, Sisay-Joof F, Ndila CM, Macharia AW, Hubbard C, Jeffreys AE, Rowlands K, Nguyen T, Gonçalves S, Ariani CV, Stalker J, Pearson RD, Amato R, Drury E, Sirugo G, d'Alessandro U, Bojang KA, Marsh K, Peshu N, Saelens JW, Diakité M, Taylor SM, Conway DJ, Williams TN, Rockett KA and Kwiatkowski DP** (2022) Malaria protection due to sickle haemoglobin depends on parasite genotype. *Nature* **602**(7895), 106–111. <https://doi.org/10.1038/s41586-021-04288-3>.
- Barreiro LB, Laval G, Quach H, Patin E and Quintana-Murci L** (2008) Natural selection has driven population differentiation in modern humans. *Nature Genetics* **40**(3), 340–345. <https://doi.org/10.1038/ng.78>.
- Barreiro LB and Quintana-Murci L** (2010) From evolutionary genetics to human immunology: How selection shapes host defence genes. *Nature Reviews Genetics* **11**(1), 17–30. <https://doi.org/10.1038/nrg2698>.
- Brown AC, Cohen CJ, Mielczarek O, Migliorini G, Costantino F, Allcock A, Davidson C, Elliott KS, Fang H, Lledó Lara A, Martin AC, Osgood JA, Sanniti A, Scozzafava G, Vecellio M, Zhang P, Black MH, Li S, Truong D, Molineros J, Howe T, Wordsworth BP, Bowness P and Knight JC** (2023) Comprehensive epigenomic profiling reveals the extent of disease-specific chromatin states and informs target discovery in ankylosing spondylitis. *Cell Genomics* **3**(6), 100306. <https://doi.org/10.1016/j.xgen.2023.100306>.
- Brown MA, Kenna T and Wordsworth BP** (2016) Genetics of ankylosing spondylitis—Insights into pathogenesis. *Nature Reviews Rheumatology* **12**(2), 81–91. <https://doi.org/10.1038/nrrheum.2015.133>.
- Cano-Gomez E, Burnham KL, Goh C, Allcock A, Malick ZH, Overend L, Kwok A, Smith DA, Peters-Sengers H, Antcliffe D, McKechnie S, Scicluna BP, van der Poll T, Gordon AC, Hinds CJ, Davenport EE, Knight JC, Webster N, Galley H, Taylor J, Hall S, Addison J, Roughton S, Tennant H, Guleri A, Waddington N, Arawawala D, Durcan J, Short A, Swan K, Williams S, Smolen S, Mitchell-Inwag C, Gordon T, Errington E, Templeton M, Venatesh P, Ward G, McCauley M, Baudouin S, Higham C, Soar J, Grier S, Hall E, Brett S, Kitson D, Wilson R, Mountford L, Moreno J, Hall P, Hewlett J, McKechnie S, Garrard C, Millo J, Young D, Hutton P, Parsons P, Smiths A, Faras-Arraya R, Soar J, Raymode P, Thompson J, Bowrey S, Kazembe S, Rich N, Andreou P, Hales D, Roberts E, Fletcher S, Rosbergen M, Glistler G, Cuesta JM, Bion J, Millar J, Perry EJ, Willis H, Mitchell N, Ruel S, Carrera R, Wilde J, Nilson A, Lees S, Kapila A, Jacques N, Atkinson J, Brown A, Prowse H, Krige A, Bland M, Bullock L, Harrison D, Mills G, Humphreys J, Armitage K, Laha S, Baldwin J, Walsh A, Doherty N, Drage S, Ortiz-Ruiz de Gordo L, Lowes S, Higham C, Walsh H, Calder V, Swan C, Payne H, Higgins D, Andrews S, Mappleback S, Hind C, Garrard C, Watson D, McLees E, Purdy A, Stotz M, Ochelli-Okpue A, Bonner S, Whitehead I, Hugil K, Goodridge V, Cawthor L, Kuper M, Pahary S, Bellingan G, Marshall R, Montgomery H, Ryu JH, Bercades G, Boluda S, Bentley A, McCalman K, Jefferies F, Knight J, Davenport E, Burnham K, Maugeri N, Radhakrishnan J, Mi Y, Allcock A and Goh C** (2022) An immune dysfunction score for stratification of patients with acute infection based on whole-blood gene expression. *Science Translational Medicine* **14**(669), eabq4433. <https://doi.org/10.1126/scitranslmed.abq4433>.
- Carrington M, Nelson GW, Martin MP, Kissner T, Vlahov D, Goedert JJ, Kaslow R, Buchbinder S, Hoots K and O'Brien SJ** (1999) HLA and HIV-1: Heterozygote advantage and B* 35-Cw* 04 disadvantage. *Science* **283**(5408), 1748–1752. <https://doi.org/10.1126/science.283.5408.1748>.
- Chen Z, Lee L, Chen J, Collins R, Wu F, Guo Y, Linksted P and Peto R** (2005) Cohort profile: The Kadoorie study of chronic disease in China (KSCDC). *International Journal of Epidemiology* **34**(6), 1243–1249. <https://doi.org/10.1093/ije/dyi174>.
- Chen L, Wang Y-F, Liu L, Bielowska A, Ahmed R, Zhang H, Tombleson P, Roberts AL, Odhams CA and Cunninghame Graham DS** (2020) Genome-wide assessment of genetic risk for systemic lupus erythematosus and disease severity. *Human Molecular Genetics* **29**(10), 1745–1756. <https://doi.org/10.1093/hmg/ddaa030>.
- Cheng J, Novati G, Pan J, Bycroft C, Žemgulytė A, Applebaum T, Pritzel A, Wong LH, Zielinski M, Sargeant T, Schneider RG, Senior AW, Jumper J, Hassabis D, Kohli P and Avsec Ž** (2023) Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science* **381**(6664), eadg7492. <https://doi.org/10.1126/science.adg7492>.
- Davenport EE, Burnham KL, Radhakrishnan J, Humburg P, Hutton P, Mills TC, Rautanen A, Gordon AC, Garrard C, Hill AV, Hinds CJ and Knight JC** (2016) Genomic landscape of the individual host response and outcomes in sepsis: A prospective cohort study. *The Lancet Respiratory Medicine* **4**(4), 259–271. [https://doi.org/10.1016/S2213-2600\(16\)00046-1](https://doi.org/10.1016/S2213-2600(16)00046-1).
- de la Fuente J and Contreras M** (2021) Vaccinomics: A future avenue for vaccine development against emerging pathogens. *Expert Review of Vaccines* **20**(12), 1561–1569. <https://doi.org/10.1080/14760584.2021.1987222>.
- Delorey TM, Ziegler CG, Heimberg G, Normand R, Yang Y, Segerstolpe Å, Abbondanza D, Fleming SJ, Subramanian A and Montoro DT** (2021) COVID-19 tissue atlases reveal SARS-CoV-2 pathology and cellular targets. *Nature* **595**(7865), 107–113. <https://doi.org/10.1038/s41586-021-03570-8>.
- Dendrou CA, Petersen J, Rossjohn J and Fugger L** (2018) HLA variation and disease. *Nature Reviews Immunology* **18**(5), 325–339. <https://doi.org/10.1038/nri.2017.143>.
- Desombere I, Willems A and Leroux-Roels G** (1998) Response to hepatitis B vaccine: Multiple HLA genes are involved. *Tissue Antigens* **51**(6), 593–604. <https://doi.org/10.1111/j.1399-0039.1998.tb03001.x>.
- DiMasi JA, Grabowski HG and Hansen RW** (2016) Innovation in the pharmaceutical industry: New estimates of R&D costs. *J Health Econ* **47**, 20–33. <https://doi.org/10.1016/j.jhealeco.2016.01.012>.
- Ding Y, Hou K, Burch KS, Lapinska S, Privé F, Vilhjálmsson B, Sankaraman S and Pasiunic B** (2022) Large uncertainty in individual polygenic risk score estimation impacts PRS-based risk stratification. *Nature Genetics* **54**(1), 30–39. <https://doi.org/10.1038/s41588-021-00961-5>.
- Duncan L, Shen H, Gelaye B, Meijssen J, Ressler K, Feldman M, Peterson R and Domingue B** (2019) Analysis of polygenic risk score usage and performance in diverse human populations. *Nature Communications* **10**(1), 3328. <https://doi.org/10.1038/s41467-019-11112-0>.
- Edgar J, Buckland M, Guzman D, Conlon N, Knerr V, Bangs C, Reiser V, Panahloo Z, Workman S and Slatter M** (2014) The United Kingdom primary immune deficiency (UKPID) registry: Report of the first 4 years' activity 2008–2012. *Clinical & Experimental Immunology* **175**(1), 68–78. <https://doi.org/10.1111/cei.12172>.
- Fang H, De Wolf H, Knezevic B, Burnham KL, Osgood J, Sanniti A, Lledó Lara A, Kasela S, De Cesco S, Wegner JK, Handunnethi L, McCann FE, Chen L, Sekine T, Brennan PE, Marsden BD, Damerell D, O'Callaghan CA, Bountra C, Bowness P, Sundström Y, Milani L, Berg L, Göhlmann HW, Peeters PJ, Fairfax BP, Sundström M and Knight JC** (2019) A genetics-led approach defines the drug target landscape of 30 immune-related traits. *Nat Genet* **51**(7), 1082–1091. <https://doi.org/10.1038/s41588-019-0456-1>.
- Fatumo S, Chikowore T, Choudhury A, Ayub M, Martin AR and Kuchenbaecker K** (2022) A roadmap to increase diversity in genomic studies. *Nature Medicine* **28**(2), 243–250. <https://doi.org/10.1038/s41591-021-01672-4>.

- Ferreira A, Marguti I, Bechmann I, Jeney V, Chora Á, Palha NR, Rebelo S, Henri A, Beuzard Y and Soares MP (2011) Sickle hemoglobin confers tolerance to plasmodium infection. *Cell* 145(3), 398–409. <https://doi.org/10.1016/j.cell.2011.03.049>.
- Ge T, Irvin MR, Patki A, Srinivasasainagendra V, Lin Y-F, Tiwari HK, Armstrong ND, Benoit B, C-Y C and Choi KW (2022) Development and validation of a trans-ancestry polygenic risk score for type 2 diabetes in diverse populations. *Genome Medicine* 14(1), 1–16. <https://doi.org/10.1186/s13073-022-01074-2>.
- Gudbjartsson DF, Helgason H, Gudjonsson SA, Zink F, Oddson A, Gylfason A, Besenbacher S, Magnusson G, Halldorsson BV and Hjartarson E (2015) Large-scale whole-genome sequencing of the Icelandic population. *Nature Genetics* 47(5), 435–444. <https://doi.org/10.1038/ng.3247>.
- Gulcher J and Stefansson K (1999) An Icelandic saga on a centralized health-care database and democratic decision making. *Nature Biotechnology* 17(7), 620–620. <https://doi.org/10.1038/10796>.
- Gurdasani D, Carstensen T, Fatumo S, Chen G, Franklin CS, Prado-Martinez J, Bouman H, Abascal F, Haber M and Tachmazidou I (2019) Uganda genome resource enables insights into population history and genomic discovery in Africa. *Cell* 179(4), 984–1002.e36. <https://doi.org/10.1016/j.cell.2019.10.004>.
- Hingorani AD, Gratton J, Finan C, Schmidt AF, Patel R, Sofat R, Kuan V, Langenberg C, Hemingway H, Morris JK and Wald NJ (2023) Performance of polygenic risk scores in screening, prediction, and risk stratification: secondary analysis of data in the Polygenic Score Catalog. *BMJ Med* 2(1), e000554. <https://doi.org/10.1136/bmjmed-2023-000554>.
- International Genetics of Ankylosing Spondylitis Consortium (2013) Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nat Genet* 45(7), 730–738. <https://doi.org/10.1038/ng.2667>.
- International Multiple Sclerosis Genetics Consortium and MultipleMS Consortium (2023) Locus for severity implicates CNS resilience in progression of multiple sclerosis. *Nature* 619(7969), 323–331. <https://doi.org/10.1038/s41586-023-06250-x>.
- Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, Musolf A, Li Q, Holzinger E and Karyadi D (2016) REVEL: An ensemble method for predicting the pathogenicity of rare missense variants. *The American Journal of Human Genetics* 99(4), 877–885. <https://doi.org/10.1016/j.ajhg.2016.08.016>.
- Jacobson RM, Ovsyannikova IG, Vierkant RA, Pankratz VS and Poland GA (2011) Human leukocyte antigen associations with humoral and cellular immunity following a second dose of measles-containing vaccine: Persistence, dampening, and extinction of associations found after a first dose. *Vaccine* 29(45), 7982–7991. <https://doi.org/10.1016/j.vaccine.2011.08.060>.
- Jaganathan K, Panagiotopoulou SK, McRae JF, Darbandi SF, Knowles D, Li YI, Kosmicki JA, Arbelaez J, Cui W and Schwartz GB (2019) Predicting splicing from primary sequence with deep learning. *Cell* 176(3), 535–548.e24. <https://doi.org/10.1016/j.cell.2018.12.015>.
- Kamiza AB, Toure SM, Vujkovic M, Machipisa T, Soremekun OS, Kintu C, Corpas M, Pirie F, Young E and Gill D (2022) Transferability of genetic risk scores in African populations. *Nature Medicine* 28(6), 1163–1166. <https://doi.org/10.1038/s41591-022-01835-x>.
- Kariuki SN and Williams TN (2020) Human genetics and malaria resistance. *Human Genetics* 139(6), 801–811. <https://doi.org/10.1007/s00439-020-02142-6>.
- Kerner G, Neehus A-L, Philippot Q, Bohlen J, Rinchai D, Kerrouche N, Puel A, Zhang S-Y, Boisson-Dupuis S and Abel L (2023) Genetic adaptation to pathogens and increased risk of inflammatory disorders in post-Neolithic Europe. *Cell Genomics* 3(2), 100248. <https://doi.org/10.1016/j.xgen.2022.100248>.
- Khunsriraksakul C, Markus H, Olsen NJ, Carrel L, Jiang B and Liu DJ (2022) Construction and application of polygenic risk scores in autoimmune diseases. *Frontiers in Immunology* 13, 889296. <https://doi.org/10.3389/fimmu.2022.889296>.
- Kienzler A, Hargreaves C and Patel S (2017) The role of genomics in common variable immunodeficiency disorders. *Clinical & Experimental Immunology* 188(3), 326–332. <https://doi.org/10.1111/cei.12947>.
- Klunk J, Vilgalys TP, Demeure CE, Cheng X, Shiratori M, Madej J, Beau R, Elli D, Patino M, Redfern R, DeWitte SN, Gamble JA, Boldsen JL, Carmichael A, Varlik N, Eaton K, Grenier JC, Golding GB, Devault A, Rouillard JM, Yotova V, Sindeaux R, Ye CJ, Bikaran M, Dumaine A, Brinkworth JF, Missiakas D, Rouleau GA, Steinrücken M, Pizarro-Cerdá J, Poinar HN and Barreiro LB (2022) Evolution of immune genes is associated with the Black Death. *Nature* 611(7935), 312–319. <https://doi.org/10.1038/s41586-022-05349-x>.
- Ko A, Jung DE, Kim SH, Kang H-C, Lee JS, Lee ST, Choi JR and Kim HD (2018) The efficacy of ketogenic diet for specific genetic mutation in developmental and epileptic encephalopathy. *Frontiers in Neurology* 530. <https://doi.org/10.3389/fneur.2018.00530>.
- Kubo M (2017) BioBank Japan project: Epidemiological study. *Journal of Epidemiology* 27(Supplement_III), S1. <https://doi.org/10.1016/j.je.2016.11.001>.
- Kurki MI, Karjalainen J, Palta P, Sipilä TP, Kristiansson K, Donner KM, Reeve MP, Laivuori H, Aavikko M and Kaunisto MA (2023) FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature* 613(7944), 508–518. <https://doi.org/10.1038/s41586-022-05473-8>.
- Kwiatkowski DP (2005) How malaria has affected the human genome and what human genetics can teach us about malaria. *Am J Hum Genet* 77(2), 171–192. <https://doi.org/10.1086/432519>.
- Kwok AJ, Mentzer A and Knight JC (2021) Host genetics and infectious disease: new tools, insights and translational opportunities. *Nat Rev Genet* 22(3), 137–153. <https://doi.org/10.1038/s41576-020-00297-6>.
- Kwok AJ, Allcock A, Ferreira RC, Cano-Gamez E, Smeets M, Burnham KL, Zurke YX, McKechnie S, Mentzer AJ, Monaco C, Udalova IA, Hinds CJ, Todd JA, Davenport EE and Knight JC (2023) Neutrophils and emergency granulopoiesis drive immune suppression and an extreme response endotype during sepsis. *Nature Immunology* 24, 767–779. <https://doi.org/10.1038/s41590-023-01490-5>.
- Lambert ND, Haralambieva IH, Kennedy RB, Ovsyannikova IG, Pankratz VS and Poland GA (2015) Polymorphisms in HLA-DPB1 are associated with differences in rubella virus-specific humoral immunity after vaccination. *The Journal of Infectious Diseases* 211(6), 898–905. <https://doi.org/10.1093/infdis/jiu553>.
- Lappalainen T, Scott AJ, Brandt M and Hall IM (2019) Genomic analysis in the age of human genome sequencing. *Cell* 177(1), 70–84. <https://doi.org/10.1016/j.cell.2019.02.032>.
- Lenardo M, Lo B and Lucas CL (2016) Genomics of immune diseases and new therapies. *Annual Review of Immunology* 34, 121–149. <https://doi.org/10.1146/annurev-immunol-041015-055620>.
- Leslie A, Matthews PC, Listgarten J, Carlson JM, Kadie C, Ndung'u T, Brander C, Coovadia H, Walker BD and Heckerman D (2010) Additive contribution of HLA class I alleles in the immune control of HIV-1 infection. *Journal of Virology* 84(19), 9879–9888. <https://doi.org/10.1128/jvi.00320-10>.
- Lewis CM and Vassos E (2020) Polygenic risk scores: From research tools to clinical instruments. *Genome Medicine* 12(1), 44. <https://doi.org/10.1186/s13073-020-00742-5>.
- Li Z, Wu X, Leo PJ, De Guzman E, Akkoc N, Breban M, Macfarlane GJ, Mahmoudi M, Marzo-Ortega H and Anderson LK (2021) Polygenic risk scores have high diagnostic capacity in ankylosing spondylitis. *Annals of the Rheumatic Diseases* 80(9), 1168–1174. <https://doi.org/10.1136/annrheumdis-2020-219446>.
- Liu J, Kumar S, Hong J, Huang Z-M, Paez D, Castillo M, Calvo M, Chang H-W, Cummins DD, Chung M, Yeroushalmi S, Bartholomew E, Hakimi M, Ye CJ, Bhutani T, Matloubian M, Gensler LS and Liao W (2022) Combined Single Cell Transcriptome and Surface Epitope Profiling Identifies Potential Biomarkers of Psoriatic Arthritis and Facilitates Diagnosis via Machine Learning. *Frontiers in Immunology* 13, 835760. <https://doi.org/10.3389/fimmu.2022.835760>.
- Liu JZ, Van Sommeren S, Huang H, Ng SC, Alberts R, Takahashi A, Ripke S, Lee JC, Jostins L and Shah T (2015) Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nature Genetics* 47(9), 979–986. <https://doi.org/10.1038/ng.3359>.
- Mahajan A, Spracklen CN, Zhang W, Ng MC, Petty LE, Kitajima H, Yu GZ, Rieger S, Speidel L and Kim YJ (2022) Multi-ancestry genetic study of type

- 2 diabetes highlights the power of diverse populations for discovery and translation. *Nature Genetics* 54(5), 560–572. <https://doi.org/10.1038/s41588-022-01058-3>.
- Matzaraki V, Kumar V, Wijmenga C and Zhernakova A (2017) The MHC locus and genetic susceptibility to autoimmune and infectious diseases. *Genome Biology* 18(1), 76. <https://doi.org/10.1186/s13059-017-1207-1>.
- Mauro D, Thomas R, Guggino G, Lories R, Brown MA and Ciccia F (2021) Ankylosing spondylitis: An autoimmune or autoinflammatory disease? *Nature Reviews Rheumatology* 17(7), 387–404. <https://doi.org/10.1038/s41584-021-00625-y>.
- Mbarek H, Devadoss Gandhi G, Selvaraj S, Al-Muftah W, Badji R, Al-Sarraj Y, Saad C, Darwish D, Alvi M and Fadl T (2022) Qatar genome: Insights on genomics from the Middle East. *Human Mutation* 43(4), 499–510. <https://doi.org/10.1002/humu.24336>.
- Melms JC, Biermann J, Huang H, Wang Y, Nair A, Tagore S, Katsyv I, Rendeiro AF, Amin AD and Schapiro D (2021) A molecular single-cell lung atlas of lethal COVID-19. *Nature* 595(7865), 114–119. <https://doi.org/10.1038/s41586-021-03569-1>.
- Mentzer AJ, O'Connor D, Bibi S, Chelysheva I, Clutterbuck EA, Demissie T, Dinesh T, Edwards NJ, Felle S, Feng S, Flaxman AL, Karp-Tatham E, Li G, Liu X, Marchevsky N, Godfrey L, Makinson R, Bull MB, Fowler J, Alamad B, Malinauskas T, Chong AY, Sanders K, Shaw RH, Voysey M, Snape MD, Pollard AJ, Lambe T and Knight JC (2023) Human leukocyte antigen alleles associate with COVID-19 vaccine immunogenicity and risk of breakthrough infection. *Nature Medicine* 29(1), 147–157. <https://doi.org/10.1038/s41591-022-02078-6>.
- Mountjoy E, Schmidt EM, Carmona M, Schwartzentruber J, Peat G, Miranda A, Fumis L, Hayhurst J, Buniello A and Karim MA (2021) An open approach to systematically prioritize causal variants and genes at all published human GWAS trait-associated loci. *Nature Genetics* 53(11), 1527–1533. <https://doi.org/10.1038/s41588-021-00945-5>.
- Mulder NJ, Adebisi E, Alami R, Benkahla A, Brandful J, Doumbia S, Everett D, Fadllemola FM, Gaboun F and Gaseitsiwe S (2016) H3ABioNet, a sustainable pan-African bioinformatics network for human heredity and health in Africa. *Genome Research* 26(2), 271–277. <https://doi.org/10.1101/gr.196295.115>.
- Naslavsky MS, Scliar MO, Yamamoto GL, JYT W, Zverinova S, Karp T, Nunes K, JRM C, de Carvalho DL and da Silva Simões CE (2022) Whole-genome sequencing of 1,171 elderly admixed individuals from Brazil. *Nature Communications* 13(1), 1–11. <https://doi.org/10.1038/s41467-022-29575-z>.
- Nelson MR, Tipney H, Painter JL, Shen J, Nicoletti P, Shen Y, Floratos A, Sham PC, Li MJ, Wang J, Cardon LR, Whittaker JC and Sanson P (2015) The support of human genetic evidence for approved drug indications. *Nat Genet* 47(8), 856–860. <https://doi.org/10.1038/ng.3314>.
- NHS England (2021) *National genomic test directory. Testing Criteria for Rare and Inherited Disease*. Accessed August.
- Ogdie A (2017) The Preclinical Phase of PsA: A Challenge for the Epidemiologist. *Ann Rheum Dis* 76(9), 1481–1483. <https://doi.org/10.1136/annrheumdis-2017-211109>.
- Okada Y, Momozawa Y, Sakaue S, Kanai M, Ishigaki K, Akiyama M, Kishikawa T, Arai Y, Sasaki T and Kosaki K (2018) Deep whole-genome sequencing reveals recent selection signatures linked to evolution and disease risk of Japanese. *Nature Communications* 9(1), 1631. <https://doi.org/10.1038/s41467-018-03274-0>.
- Pan-Hammarström Q, Salzer U, Du L, Björkander J, Cunningham-Rundles C, Nelson DL, Bacchelli C, Gaspar HB, Offer S and Behrens TW (2007) Reexamining the role of TAC1 coding variants in common variable immunodeficiency and selective IgA deficiency. *Nature Genetics* 39(4), 429–430. <https://doi.org/10.1038/ng0407-429>.
- Pankratov V, Yunusbaeva M, Ryakhovskiy S, Zarodniuk M, Estonian Biobank Research Team and Yunusbayev B (2022) Prioritizing autoimmunity risk variants for functional analyses by fine-mapping mutations under natural selection. *Nature Communications* 13(1), 7069. <https://doi.org/10.1038/s41467-022-34461-9>.
- Peng XP, Caballero-Oteyza A and Grimbacher B (2023) Common variable immunodeficiency: More pathways than roads to Rome. *Annual Review of Pathology: Mechanisms of Disease* 18, 283–310.
- Pereira L, Mutesa L, Tindana P and Ramsay M (2021) African genetic diversity and adaptation inform a precision medicine agenda. *Nature Reviews Genetics* 22(5), 284–306. <https://doi.org/10.1038/s41576-020-00306-8>.
- Phillips KA, Douglas MP, Wordsworth S, Buchanan J and Marshall DA (2021) Availability and funding of clinical genomic sequencing globally. *BMJ Global Health* 6(2), e004415. <https://doi.org/10.1136/bmjgh-2020-004415>.
- Pinto MF, Oliveira H, Batista S, Cruz L, Pinto M, Correia I, Martins P and Teixeira C (2020) Prediction of disease progression and outcomes in multiple sclerosis with machine learning. *Scientific Reports* 10(1), 1–13. <https://doi.org/10.1038/s41598-020-78212-6>.
- Redmond MT, Scherzer R and Prince BT (2022) Novel genetic discoveries in primary immunodeficiency disorders. *Clinical Reviews in Allergy and Immunology* 63(1), 55–74. <https://doi.org/10.1007/s12016-021-08881-2>.
- Reid S, Alexsson A, Frodlund M, Morris D, Sandling JK, Bolin K, Svenungsson E, Jönsen A, Bengtsson C and Gunnarsson I (2020) High genetic risk score is associated with early disease onset, damage accrual and decreased survival in systemic lupus erythematosus. *Annals of the Rheumatic Diseases* 79(3), 363–369. <https://doi.org/10.1136/annrheumdis-2019-216227>.
- Robertson CC, Inshaw JR, Onengut-Gumuscu S, Chen W-M, Santa Cruz DF, Yang H, Cutler AJ, Crouch DJ, Farber E and Bridges SL, Jr (2021) Fine-mapping, trans-ancestral and genomic analyses identify causal variants, cells, genes and drug targets for type 1 diabetes. *Nature Genetics* 53(7), 962–971. <https://doi.org/10.1038/s41588-021-00880-5>.
- Rönnblom L and Leonard D (2019) Interferon pathway in SLE: One key to unlocking the mystery of the disease. *Lupus Science & Medicine* 6(1), e000270. <https://doi.org/10.1136/lupus-2018-000270>.
- Rotimi C, Abayomi A, Al A, Adabayeri VM, Adebamowo C, Adebisi E, Ademola AD, Adeyemo A, Adu D and Affolabi D (2014) Research capacity. Enabling the genomic revolution in Africa. *Science (New York, N.Y.)* 344(6190), 1346–1348. <https://doi.org/10.1126/science.1251546>.
- Ruan Y, Lin Y-F, Feng Y-CA, Chen C-Y, Lam M, Guo Z, He L, Sawa A, Martin AR and Qin S, Huang H and Ge T (2022) Improving polygenic prediction in ancestrally diverse populations. *Nature Genetics* 54(5), 573–580. <https://doi.org/10.1038/s41588-022-01054-7>.
- Sakaue S, Hirata J, Kanai M, Suzuki K, Akiyama M, Lai Too C, Arayssi T, Hammoudeh M, Al Emadi S, Masri BK, Halabi H, Badsha H, Uthman IW, Saxena R, Padyukov L, Hirata M, Matsuda K, Murakami Y, Kamatani Y and Okada Y (2020) Dimensionality reduction reveals fine-scale structure in the Japanese population with consequences for polygenic risk prediction. *Nature Communications* 11(1), 1569. <https://doi.org/10.1038/s41467-020-15194-z>.
- Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, Bellomo R, Bernard GR, Chiche JD, Coopersmith CM, Hotchkiss RS, Levy MM, Marshall JC, Martin GS, Opal SM, Rubenfeld GD, van der Poll T, Vincent JL and Angus DC (2016) The third international consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA* 315(8), 801–810. <https://doi.org/10.1001/jama.2016.0287>.
- Soskic B, Cano-Gamez E, Smyth DJ, Ambridge K, Ke Z, Matte JC, Bossini-Castillo L, Kaplanis J, Ramirez-Navarro L, Lorenc A, Nakic N, Esparza-Gordillo J, Rowan W, Wille D, Tough DF, Bronson PG and Trynka G (2022) Immune disease risk variants regulate gene expression dynamics during CD4(+) T cell activation. *Nature Genetics* 54(6), 817–826. <https://doi.org/10.1038/s41588-022-01066-3>.
- Soskic B, Cano-Gamez E, Smyth DJ, Rowan WC, Nakic N, Esparza-Gordillo J, Bossini-Castillo L, Tough DF, Larminie CG, Bronson PG, Willé D and Trynka G (2019) Chromatin activity at GWAS loci identifies T cell states driving complex immune diseases. *Nature Genetics* 51(10), 1486–1493. <https://doi.org/10.1038/s41588-019-0493-9>.
- Stuart PE, Tsoi LC, Nair RP, Ghosh M, Kabra M, Shaiq PA, Raja GK, Qamar R, Thelma B, Patrick MT, Parihar A, Singh S, Khandpur S, Kumar U, Wittig M, Degenhardt F, Tejasvi T, Voorhees JJ, Weidinger S, Franke A, Abecasis GR, Sharma VK and Elder JT (2022) Transethnic analysis of psoriasis susceptibility in South Asians and Europeans enhances fine-mapping in the MHC and genomewide. *HGG Adv* 3(1). <https://doi.org/10.1016/j.xhgg.2021.100069>.
- Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, Liu B, Matthews P, Ong G, Pell J, Silman A, Young A, Sprosen T, Peakman T and Collins R (2015) UK biobank: an open

- access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Medicine* 12(3), e1001779. <https://doi.org/10.1371/journal.pmed.1001779>.
- Symonds JD, Elliott KS, Shetty J, Armstrong M, Brunklaus A, Cutcutache I, Diver LA, Dorris L, Gardiner S, Jollands A, Joss S, Kirkpatrick M, McLellan A, MacLeod S, O'Regan M, Page M, Pilley E, Pilz DT, Stephen E, Stewart K, Ashrafian H, Knight JC and Zuberi SM (2021) Early childhood epilepsies: epidemiology, classification, aetiology, and socio-economic determinants. *Brain* 144(9), 2879–2891. <https://doi.org/10.1093/brain/awab162>.
- Tam V, Patel N, Turcotte M, Bosse Y, Pare G and Meyre D (2019) Benefits and limitations of genome-wide association studies. *Nature Reviews. Genetics* 20(8), 467–484. <https://doi.org/10.1038/s41576-019-0127-1>.
- Team SGP (2015) The Saudi human genome program: An oasis in the desert of Arab medicine is providing clues to genetic disease. *IEEE Pulse* 6(6), 22–26. <https://doi.org/10.1109/mpul.2015.2476541>.
- Thaventhiran JE, Lango Allen H, Burren OS, Rae W, Greene D, Staples E, Zhang Z, Farmery JH, Simeoni I, Rivers E, Maimaris J, Penkett CJ, Stephens J, Deevi SVV, Sanchis-Juan A, Gleadall NS, Thomas MJ, Sargur RB, Gordins P, Baxendale HE, Brown M, Tuijnburg P, Worth A, Hanson S, Linger RJ, Buckland MS, Rayner-Matthews PJ, Gilmour KC, Samarghitean C, Seneviratne SL, Sansom DM, Lynch AG, Megy K, Ellinghaus E, Ellinghaus D, Jorgensen SF, Karlsen TH, Stirrups KE, Cutler AJ, Kumararatne DS, Chandra A, Edgar JDM, Herwadkar A, Cooper N, Grigoriadou S, Huissoon AP, Goddard S, Jolles S, Schuetz C, Boschann F, Lyons PA, Hurles ME, Savic S, Burns SO, Kuijpers TW, Turro E, Ouwehand WH, Thrasher AJ and Smith KGC (2020) Whole-genome sequencing of a sporadic primary immunodeficiency cohort. *Nature* 583(7814), 90–95. <https://doi.org/10.1038/s41586-020-2265-1>.
- The International HapMap Project (2003) *Nature* 426(6968), 789–796. <https://doi.org/10.1038/nature02168>.
- Thursz MR, Thomas HC, Greenwood BM and Hill AV (1997) Heterozygote advantage for HLA class-II type in hepatitis B virus infection. *Nature Genetics* 17(1), 11–12. <https://doi.org/10.1038/ng0997-11>.
- Tsoi LC, Spain SL, Knight J, Ellinghaus E, Stuart PE, Capon F, Ding J, Li Y, Tejasvi T, Gudjonsson JE, Kang HM, Allen MH, McManus R, Novelli G, Samuelsson L, Schalkwijk J, Ståhle M, Burden AD, Smith CH, Cork MJ, Estivill X, Bowcock AM, Krueger GG, Weger W, Worthington J, Tazi-Ahni R, Nestle FO, Hayday A, Hoffmann P, Winkelmann J, Wijmenga C, Langford C, Edkins S, Andrews R, Blackburn H, Strange A, Band G, Pearson RD, Vukcevic D, Spencer CC, Deloukas P, Mrowietz U, Schreiber S, Weidinger S, Koks S, Kingo K, Esko T, Metspalu A, Lim HW, Voorhees JJ, Weichenthal M, Wichmann HE, Chandran V, Rosen CF, Rahman P, Gladman DD, Griffiths CE, Reis A, Kere J, Nair RP, Franke A, Barker JN, Abecasis GR, Elder JT and Trembath RC (2012) Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity. *Nature Genetics* 44(12), 1341–1348. <https://doi.org/10.1038/ng.2467>.
- Tuijnburg P, Allen HL, Burns SO, Greene D, Jansen MH, Staples E, Stephens J, Carss KJ, Biasci D, Baxendale H, Thomas M, Chandra A, Kiani-Alikhan S, Longhurst HJ, Seneviratne SL, Oksenhendler E, Simeoni I, de Bree GJ, Tool ATJ, van Leeuwen EMM, Ebberink E, Meijer AB, Tuna S, Whitehorn D, Brown M, Turro E, Thrasher AJ, Smith KGC, Thaventhiran JE and Kuijpers TW (2018) Loss-of-function nuclear factor κB subunit 1 (NFKB1) variants are the most common monogenic cause of common variable immunodeficiency in Europeans. *J Allergy Clin Immunol* 142(4), 1285–1296. <https://doi.org/10.1016/j.jaci.2018.01.039>.
- Vieira Braga FA, Kar G, Berg M, Carpaj OA, Polanski K, Simon LM, Brouwer S, Gomes T, Hesse L, Jiang J, Fasouli ES, Efremova M, Vento-Tormo R, Talavera-López C, Jonker MR, Affleck K, Palit S, Strzelecka PM, Firth HV, Mahbubani KT, Cvejic A, Meyer KB, Saeb-Parsy K, Luinge M, Brandsma CA, Timens W, Angelidis I, Strunz M, Koppelman GH, van Oosterhout AJ, Schiller HB, Theis FJ, van den Berge M, Nawijn MC and Teichmann SA (2019) A cellular census of human lungs identifies novel cell states in health and in asthma. *Nature Medicine* 25(7), 1153–1163. <https://doi.org/10.1038/s41591-019-0468-5>.
- Wang S, Yao X, Ma S, Ping Y, Fan Y, Sun S, He Z, Shi Y, Sun L, Xiao S, Song M, Cai J, Li J, Tang R, Zhao L, Wang C, Wang Q, Zhao L, Hu H, Liu X, Sun G, Chen L, Pan G, Chen H, Li Q, Zhang P, Xu Y, Feng H, Zhao GG, Wen T, Yang Y, Huang X, Li W, Liu Z, Wang H, Wu H, Hu B, Ren Y, Zhou Q, Qu J, Zhang W, Liu GH and Bian XW (2021) A single-cell transcriptomic landscape of the lungs of patients with COVID-19. *Nature Cell Biology* 23(12), 1314–1328. <https://doi.org/10.1038/s41556-021-00796-6>.
- Weissbrod O, Kanai M, Shi H, Gazal S, Peyrot WJ, Khara AV, Okada Y, Martin AR, Finucane HK and Price AL (2022) Leveraging fine-mapping and multipopulation training data to improve cross-population polygenic risk scores. *Nature Genetics* 54(4), 450–458. <https://doi.org/10.1038/s41588-022-01036-9>.
- Wonkam A and Mayosi BM (2014) Genomic medicine in Africa: Promise, problems and prospects. *Genome Medicine* 6(2), 11. <https://doi.org/10.1186/gm528>.
- Wu D, Dou J, Chai X, Bellis C, Wilm A, Shih CC, Soon WWJ, Bertin N, Lin CB and Khor CC (2019) Large-scale whole-genome sequencing of three diverse Asian populations in Singapore. *Cell* 179(3), 736–749.e15. <https://doi.org/10.1016/j.cell.2019.09.019>.
- Zayed H (2016) The Qatar genome project: Translation of whole-genome sequencing into clinical practice. *International Journal of Clinical Practice* 70(10), 832–834. <https://doi.org/10.1111/ijcp.12871>.
- Zhang Q, Bastard P, Liu Z, Le Pen J, Moncada-Velez M, Chen J, Ogishi M, Sabli IK, Hodeib S, Korol C, Rosain J, Bilguvar K, Ye J, Bolze A, Bigio B, Yang R, Arias AA, Zhou Q, Zhang Y, Onodi F, Korniotis S, Karpf L, Philippot Q, Chbihi M, Bonnet-Madin L, Dorgham K, Smith N, Schneider WM, Razoooky BS, Hoffmann HH, Michailidis E, Moens L, Han JE, Lorenzo L, Bizien L, Meade P, Neehus AL, Ugurbil AC, Corneau A, Kerner G, Zhang P, Rapaport F, Seeleuthner Y, Manry J, Masson C, Schmitt Y, Schlüter A, Le Voyer T, Khan T, Li J, Fellay J, Roussel L, Shahrooei M, Alosaimi MF, Mansouri D, Al-Saud H, Al-Mulla F, Almourfi F, Al-Muhesen SZ, Alshime F, Al Turki S, Hasanato R, van de Beek D, Biondi A, Bettini LR, D'Angio M, Bonfanti P, Imberti L, Sottini A, Paghera S, Quiros-Roldan E, Rossi C, Oler AJ, Tompkins MF, Alba C, Vandernoot I, Goffard JC, Smits G, Migeotte I, Haerynck F, Soler-Palacin P, Martin-Nalda A, Colobran R, Morange PE, Keles S, Çölkesen F, Özcelik T, Yasar KK, Senoglu S, Karabela Ş N, Rodríguez-Gallego C, Novelli G, Hraiech S, Tandjaoui-Lambiotte Y, Duval X, Laouénan C, Snow AL, Dalgard CL, Milner JD, Vinh DC, Mogensen TH, Marr N, Spaan AN, Boisson B, Boisson-Dupuis S, Bustamante J, Puel A, Ciancanelli MJ, Meyts I, Maniatis T, Soumelis V, Amara A, Nussenzweig M, García-Sastre A, Krammer F, Pujol A, Duffy D, Lifton RP, Zhang SY, Gorochov G, Béziat V, Jouanguy E, Sancho-Shimizu V, Rice CM, Abel H, Al-Mulla F, Cobat A, Su HC and Casanova JL (2020) Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science* 370(6515). <https://doi.org/10.1126/science.abd4570>.
- Zhang JY, Roberts H, Flores DSC, Cutler AJ, Brown AC, Whalley JP, Mielczarek O, Buck D, Lockstone H, Xella B, Oliver K, Corton C, Betteridge E, Bashford-Rogers R, Knight JC, Todd JA and Band G (2021) Using de novo assembly to identify structural variation of eight complex immune system gene regions. *PLoS Computational Biology* 17(8), e1009254. <https://doi.org/10.1371/journal.pcbi.1009254>.
- Zhou Y, Pan Q, Pires DE, Rodrigues CH and Ascher DB (2023) DDMut: Predicting effects of mutations on protein stability using deep learning. *Nucleic Acids Research* 51(W1), W122–W128. <https://doi.org/10.1093/nar/gkad472>.