

 Open access • Posted Content • DOI:10.1101/2020.11.02.2022232

Advancing Human Genetics Research and Drug Discovery through Exome Sequencing of the UK Biobank — [Source link](#)

Joseph D. Szustakowski, Suganthi Balasubramanian, Ariella Sasson, Shareef Khalid ...+31 more authors

Institutions: [Bristol-Myers Squibb](#), [Regeneron](#), [Biogen Idec](#), [Takeda Pharmaceutical Company](#) ...+4 more institutions

Published on: 04 Nov 2020 - [medRxiv](#) (Cold Spring Harbor Laboratory Press)

Topics: [Biobank](#), [Exome](#), [Exome sequencing](#) and [Human genetics](#)

Related papers:

- [The UK Biobank resource with deep phenotyping and genomic data](#)
- [The mutational constraint spectrum quantified from variation in 141,456 humans](#)
- [The Ensembl Variant Effect Predictor.](#)
- [UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age](#)
- [Exome sequencing and characterization of 49,960 individuals in the UK Biobank.](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/advancing-human-genetics-research-and-drug-discovery-through-swbbk6j8u8>

Advancing Human Genetics Research and Drug Discovery through Exome Sequencing of the UK Biobank

Joseph D. Szustakowski¹, Suganthi Balasubramanian², Ariella Sasson¹, Shareef Khalid², Paola G. Bronson³, Erika Kvikstad¹, Emily Wong⁴, Daren Liu², J. Wade Davis⁵, Carolina Haefliger⁷, A. Katrina Loomis⁸, Rajesh Mikkilineni⁴, Hyun Ji Noh⁵, Samir Wadhawan¹, Xiaodong Bai², Alicia Hawes², Olga Krasheninina², Ricardo Ulloa², Alex Lopez², Erin N. Smith⁴, Jeff Waring⁵, Christopher D. Whelan³, Ellen A. Tsai³, John Overton², William Salerno², Howard Jacob⁵, Sandor Szalma⁴, Heiko Runz³, Greg Hinkle⁶, Paul Nioi⁶, Slavé Petrovski⁷, Melissa R. Miller⁸, Aris Baras², Lyndon Mitnaul², and Jeffrey G. Reid² on behalf of the UKB-ESC Research Team.

¹Bristol Myers Squibb, Route 206 and Province Line Road, Princeton, NJ 08543

²Regeneron Pharmaceuticals Inc., 777 Old Saw Mill River Road, Tarrytown, New York 10591

³Biogen Inc., 225 Binney Street, Cambridge, MA 02139

⁴Takeda Pharmaceutical Company Ltd, 1-1, Nihonbashi-Honcho 2-chome, Chuo-ku, Tokyo, Japan

⁵Abbvie Inc., 1 N. Waukegan Rd, North Chicago, IL 60064

⁶Alnylam Pharmaceuticals, 675 West Kendall St, Cambridge, MA 02142

⁷AstraZeneca Centre for Genomics Research, Discovery Sciences, BioPharmaceuticals R&D, Cambridge, UK

⁸Pfizer, Inc., 1 Portland St., Cambridge, MA 02139

Abstract

The UK Biobank Exome Sequencing Consortium (UKB-ESC) is a unique private/public partnership between the UK Biobank and eight biopharma companies that will sequence the exomes of all ~500,000 UK Biobank participants. Here we describe early results from the exome sequence data generated by this consortium for the first ~200,000 UKB subjects and the key features of this project that enabled the UKB-ESC to come together and generate this data.

Exome sequencing data from the first 200,643 UKB enrollees are now accessible to the research community. Approximately 10M variants were observed within the targeted regions, including: 8,086,176 SNPs, 370,958 indels and 1,596,984 multi-allelic variants. Of the ~8M variants observed, 84.5% are coding variants and include 2,139,318 (25.3%) synonymous, 4,549,694 (53.8%) missense, 453,733 (5.4%) predicted loss-of-function (LOF) variants (initiation codon loss, premature stop codons, stop codon loss, splicing and frameshift variants) affecting at least one coding transcript. This open access data provides a rich resource of coding variants for rare variant genetic studies, and is particularly valuable for drug discovery efforts that utilize rare, functionally consequential variants.

Over the past decade, the biopharma industry has increasingly leveraged human genetics as part of their drug discovery and development strategies. This shift was motivated by technical advances that enabled cost-effective human genetics research at scale, the emergence of electronic health records and biobanks, and a maturing understanding of how human genetics can increase the probability of successful drug development. Recognizing the need for large-scale human genetics data to drive drug discovery, and the unique value of the open data access policies and contribution terms of the UK Biobank, the UKB-ESC was formed. This precompetitive collaboration has further strengthened the ties between academia and industry and provided teams an unprecedented opportunity to interact with and learn from the wider research community.

Why are drug developers interested in human genetics?

Developing novel therapies to address unmet medical needs is a resource-intensive challenge characterized by high attrition rates (Dimasi, Grabowski, & Hansen, 2016). While biopharmaceutical companies have recently improved overall R&D productivity and success rates for drug candidates in late stage clinical development, the probability of a drug candidate proceeding from Phase 1 clinical trials through approval and launch remains near 10%. Most clinical failures (approximately 75%) are attributable to safety concerns or lack of efficacy (Dowden & Munro, 2019).

The potential of human genetics to increase the likelihood of successful drug discovery has long been recognized, though largely in anecdotal form. Anti-*PCSK9* cholesterol-lowering drugs are a highly publicized example where human genetic evidence for a target contributed to technical and regulatory success, providing rationale for further investment (Furtado & Giugliano, 2020). Human genetics may also identify potential liability phenotypes associated with a target. For example, it is plausible that gastro-intestinal adverse events observed in clinical trials of *DGAT1* inhibitors (Karlsson, Knutsson, & Eriksson, 2013; Meyers et al., 2015) could have been predicted based on the causal link between rare, highly penetrant *DGAT1* variants and congenital diarrheal disorder (Haas et al., 2012). Genetics has also proven its value in bringing more precision to drug development, particularly in the context of large trials. Stratification of patients to enrich for signal has shown success in *PCSK9* inhibitors (Sabatine et al., 2017; Schwartz et al., 2018), providing the ability to select genetically defined patient-populations. Table S1 enumerates a wide array of approved drugs with targets supported by human genetics evidence.

Recently, several studies systematically characterized the role of human genetics in drug discovery (Cook et al., 2014; King et al., 2019; Nelson et al., 2015). These retrospective analyses of drug development successes and failures suggest that drug-target pairs with human genetics evidence are at least twice as likely to reach approval as those without. Moreover, those success rates increase with the strength of evidence. Targets whose strongest evidence is a GWAS hit demonstrate approximately twice the probability of success as do targets lacking human genetics evidence, whereas those associated with rare Mendelian disorders are at least five times as likely to be successful. Conversely, drugs whose targets lack genetic association with a relevant indication are less likely to progress to market (King et al., 2019). Furthermore, genetic association with a non-relevant phenotype increases the likelihood of corresponding adverse events (Nguyen et al., 2019).

Building on these and similar experiences, several frameworks for the systematic evaluation of genetically motivated targets have emerged. For example, the allelic series model leverages the existence of multiple, independent genetic alterations in a gene of interest to assess its

potential as a drug target. Genetic dose-response curves can be used to investigate the relationship between variants in the allelic series and phenotypes of interest to assess the potential for a tractable therapeutic window (Plenge, Scolnick, & Altshuler, 2013). In addition, several promising drug-development candidate targets with robust genetic evidence were identified based on associations between disease phenotypes, biomarker endo-phenotypes, and functionally consequential genetic variants. Recent examples include *HSD17B13* for chronic liver disease (Abul-Husn et al., 2018), *TYK2* for multiple autoimmune disorders (Dendrou et al., 2016; Diogo et al., 2015; Minegishi et al., 2006), *NRXN1* for neuropsychiatric disease (Noh et al. 2017), and *ASGR1* for cardiovascular disease (Nioi et al., 2016)

Large-scale biobanks have emerged to catalyze biomedical research, in part due to widespread adoption of electronic health records, and maturation of experimental and computational platforms capable of cost-effective population-scale sequencing and analysis. This creates a unique opportunity for the scientific community to build upon these foundations and accelerate the use of human genetics to inform drug discovery. Innovative public-private partnerships in the precompetitive space have demonstrated value for furthering this acceleration (Dolgin, 2019). Here, we describe one such effort – the UK Biobank Exome Sequencing Consortium (UKB-ESC) – focused on generating Whole Exome Sequencing (WES) data for all subjects (apprx. 500,000) in the UK Biobank (UKB).

Why are drug developers interested in UK Biobank?

Nominating, evaluating, and prioritizing potential drug targets based on human genetics is data intensive. The most interesting novel insights are likely to come from rare or private functional genetic variants leading to highly penetrant gain or loss of protein function (M. I. McCarthy et al., 2008). These are most effectively discovered through direct sequencing of large populations, or those enriched for functionally important variation. Human genetic data alone, however, are insufficient. Comprehensive, consistent, longitudinal phenotypic data that can be linked to the genetic data are needed to explore the breadth of biological consequences of genetic variation. Essential phenotypes may include accurate disease diagnosis, molecular and cellular biomarkers, treatment and outcome information, imaging endpoints, self-reported conditions, and environmental and lifestyle data.

Beyond target identification, access to large-scale human genetic data linked to phenotypes enables a variety of other key opportunities for drug development efforts. In addition to direct human genetic discovery, these cohorts are incredibly valuable for “instant replication” of insights from other human genetics studies, prediction of potential side-effects, recall sub-studies, rapid validation of targets proposed by other means, and unbiased study of the natural history of rare monogenic diseases of interest to drug developers. Fifteen years ago, the Wellcome Trust Case Control Consortium shifted the field from candidate gene studies to genome-wide association studies (Burton et al., 2007). Now large biobanks are moving us even further in the unbiased spectrum by recruiting population-based cohorts without regard to disease status. The disease-agnostic approach and collection of a wide variety of data and specimens allows an immense number of hypotheses to be tested. The UKB is an exemplar of

such data, a unique resource that provides a rich substrate for a broad spectrum of biomedical research (Bycroft et al., 2018).

The objective of the UKB-ESC is a comprehensive assessment of the protein-coding genetic variation in the half-million UKB participants. Enhancing this already phenotypically rich resource enables researchers to extend their genetic investigations to include rare and private variants not captured in existing UKB chip-based genotypes.

As the exome is enriched for variants that are most readily interpretable and actionable, this consortium has prioritized deep sequencing of the protein-coding portions of the genome. This is the highest value assay to add to the UKB genotype resource, pairing the rarest coding variation with the common and rare variation captured via chip genotyping and imputation. The speed of whole exome sequencing (WES) is enabling a rapid acceleration of actionable scientific discoveries. While WES is the assay of choice for Mendelian disease discovery (Y. Yang et al., 2013) and has provided important novel target discoveries (Abul-Husn et al., 2018), we are also enthusiastic for the arrival of the UKB whole-genome sequencing data (<https://www.ukbiobank.ac.uk/2019/09/uk-biobank-leads-the-way-in-genetics-research-to-tackle-chronic-diseases/>), and recognize the value of continued investment in growing the UKB resource.

Building a successful large-scale drug-discovery collaboration in genomics: lessons learned

Large-scale collaborative projects have driven transformative scientific discoveries, particularly in genetics and genomics (Auton et al., 2015; Bycroft et al., 2018b; Dunham et al., 2012; Lander et al., 2001; S. McCarthy et al., 2016; Methé et al., 2012; “The Cancer Genome Atlas Program - National Cancer Institute,”). Such collaborations provide a unique opportunity to unify scientific communities by enabling a diversity of voices and perspectives to produce insights that would be inaccessible to smaller, more homogeneous groups. Historically, large-scale collaborative life science projects have been typically funded by governments and non-profits to engage academic groups. More recently, government agencies, non-profit organizations, and academic institutions have sought industry partners for big science projects (<https://www.nih.gov/research-training/accelerating-medicines-partnership-amp>, <https://www.alzdiscovery.org/>). The UKB-ESC seeks to further strengthen the ties between academia and industry through a precompetitive collaboration. This model is enabled by a project structure that engages all parties as scientific contributors and incentivizes industry investment in a sharable data resource. We expect the output of this partnership will catalyze novel scientific discoveries, accelerate the development of new therapies, and ultimately improve patient outcomes. As both the co-funders of and scientific teams analyzing the UKB WES data, we have found the following principles essential to the success of this effort.

First and foremost, the UKB-ESC is uniquely enabled by the UKB open data access policy. The biomedical research community has already made great use of UKB data, as evidenced by the rapidly growing body of scientific literature (<https://www.ukbiobank.ac.uk/published-papers/>).

The UKB data access model is valuable in how it enables researchers to openly publish and commercialize results from and derivatives of the UKB data and incentivizes contributions back to the resource. This virtuous cycle enables academia and industry to be engaged and benefit from a body of work larger than any one entity can achieve alone.

Second, the UKB-ESC scale and scope will enable unique, valuable scientific discoveries. The hunt for actionable genetic variants (e.g. rare with profound functional consequences, common with informative phenotypic associations) depends on the scale of available data, in terms of both sample size and phenotypic characterization. Building large, diverse, and deeply characterized cohorts is an enormous multidisciplinary undertaking. This is particularly challenging in high-value phenotyping where different modalities (e.g. imaging, proteomics, EMR data extraction) require diverse expertise, and the data often require significant curation and processing prior to research use. The UKB excels at such tasks within a user-friendly framework attractive to both academic and industrial researchers, creating a roadmap for other projects with similar aspirations.

Third, the UKB's data access and contribution terms invites pre-competitive collaboration by industry partners. The generation of sequencing data at this scale requires a substantial investment of time, personnel, and financial resources (<https://www.ukbiobank.ac.uk/2018/01/regeneron-announces-major-collaboration-to-exome-sequence-uk-biobank-genetic-data-more-quickly/>). The UKB policy of providing a window of data exclusivity for data generators (also a common feature of large-scale academic collaborations) was essential to the business case for an investment of this scale. In addition to the scientific drivers described above, participation in a consortium and the exclusivity period both provide tangible benefits. While the competitive commercial interests of companies may seem to preclude cooperation, an appropriately managed collaboration mitigates individual risk to each partner and provides value larger than the sum of the individual investments. Moreover, the finite window of exclusivity helps focus efforts on scientific projects that are most likely to identify novel targets and accelerate therapeutic development.

Finally, the value of engagement in a large pre-competitive industry collaborative project provides additional value to the participating institutions. Building expertise and functional excellence in important new areas is a key benefit to participating institutions—particularly in the rapidly transforming area of genetics research in support of drug discovery. In our experience, drug discovery teams typically have deep experience in areas such as disease biology, chemistry, and translational research. In contrast, the inclusion of human genetics expertise varies between companies and projects. This discrepancy is often the result of a lack of large-scale data sets that are relevant to specific therapeutic areas or drug development programs. We expect that the increasing availability of large, well-annotated genetics resources will lead to additional industry investment in genetics.

In summary, we have identified the following key features of an effective pre-competitive industry collaboration: Build a large dataset with rich phenotypic characterization, noting that participant recontact is highly valued for answering new questions that emerge from the data.

Provide researchers the opportunity to derive both academic and/or commercial value from the data in an unencumbered way. Provide some exclusivity/first-mover advantage to build a compelling business justification for participation and financing of the project. Enable data access providing insight generation for key internal therapeutic-focused scientists and R&D stakeholders within the collaborating institutions. Provide opportunities for constructive engagement with academic partners, and low-friction data sharing terms and platforms to enable the broadest possible suite of research projects. We are hopeful others will follow the UKB model and build significant data collections designed to engage a wide variety of stakeholders.

Characterization of the first 200,000 exome sequences from the UK Biobank

The release of the first ~200,000 sequenced exomes represent a milestone in the availability of large-scale genomics data. We provide here an overview of the phenotypes and genotypes available as part of this resource. Table 1 includes a summary of the clinical characteristics of the first 50K sequenced individuals, the current 200K sequenced and overall 500K participants. Definitions of UKB phenotypes are provided in Van Hout et al, 2020. Table 2 summarizes the variants observed in the first 200,000 exomes sequenced from the UK Biobank. The targeted region covered 38,997,831 bases and coverage exceeds 20X on 95.6% of sites on average. Approximately 10M variants were observed within the targeted regions. These include: 8,086,176 SNPs, 370,958 indels and 1,596,984 multi-allelic variants. Of the ~8M SNPs observed, 84.5% are coding variants and include 2,139,318 (25.3%) synonymous, 4,549,694 (53.8%) missense, 453,733 (5.4%) predicted loss-of-function (LOF) variants (initiation codon loss, premature stop codons, stop codon loss, splicing and frameshift variants) affecting at least one coding transcript. Analysis of allele frequencies revealed 98% of synonymous, 99% of missense and 99% of LOF (at least one transcript) variants with a minor allele frequency (MAF) less than 1%. On a per individual basis, we observed a median number of 8,586 synonymous, 7,724 missense and 202 LOF variants. Restricting our analysis to variants that affect canonical transcripts, we observe a median of 142 LOFs per individual. The numbers are largely similar to published exome studies (Karczewski et al., 2020; Van Hout et al., 2020). In addition, we also analyzed the increase in number of genes with heterozygous and homozygous LOFs with the increase in number of sequenced samples. Previously, 17,718 genes with heterozygous LOF variants were observed in the UKB 50,000 exomes data (Van Hout et al., 2020), therefore, is it not surprising that we only observe a modest increase in that number (18,045 genes) in 200K exomes. 789 genes with at least one homozygous LOF variant was reported in UKB 50,000 exomes data. The number of genes with homozygous LOFs still appears to be increasing with a projection of 1981 genes with at least one homozygous LOF variant in 500K (1492 genes with at least one homozygous LOF seen in UKB 200K, Table 3).

	UKB 50K	UKB 200K	UKB 500k
Date data downloaded	March 2020	March 2020	March 2020
Number of subjects	49,898	200,633	502,504
Female, n(%)	27,207 (54.5)	110,479 (55.1)	273,378 (54.4)
Age at assessment, years	58 (50-63)	58 (50-63)	58 (50-63)
Body mass index, kg/m ²	26.6 (24-30)	26.6 (24-30)	26.7 (24-30)
Number of imaged participants (%)	14,174 (28.4)	23,818 (11.9)	44,703 (8.9)
Number of current/past smokers, n(%)	22,327 (44.7)	89,391 (44.6)	226,901 (45.2)
Townsend Deprivation Index	-2.0 (-3.6,0.6)	-2.2 (-3.7, 0.5)	-2.1 (-3.6, 0.6)
Inpatient ICD10 codes per patient	8	8	8
Patients with >=1 ICD10 diagnoses, n(%)	43,345 (86.9)	164,452 (82.0)	410,310 (81.7)
Genetic Ancestry Assignment			
African (%)	2.0	1.6	1.6
East Asian (%)	0.35	0.32	0.31
European (%)	93.3	93.7	94.0
Enhanced measures			
Hearing test, n(%)	40,761 (81.7)	84,484 (42.1)	171,826 (34.2)
Pulse Rate, n(%)	40,495 (81.2)	83,974 (41.9)	170,743 (34.0)
Visual Acuity Measured, n(%) eids	39,551 (79.3)	65,248 (32.5)	117,672 (23.4)
IOP measured (left),n(%)	37,889 (75.9)	62,213 (31.0)	111,926 (22.3)
Autorefraction, n(%) eids	39,520 (79.2)	64,807 (32.3)	116,365 (23.2)
Retinal OCT, n(%) eids	38,371 (76.9)	51,286 (25.6)	78,888 (15.7)
ECG at rest, n(%)	12,331 (24.7)	15,953 (8.0)	24,923 (5.0)
Cognitive Function, n(%)	17,340 (34.8)	53,039 (26.4)	123,614 (24.6)
Digestive Health, n(%)	23,922 (47.9)	74,417 (37.1)	174,760 (34.8)
Physical Activity Measurement, n(%)	15,002 (30.1)	44,885 (22.4)	103,684 (20.6)
Cardiometabolic phenotypes			
Coronary Disease, n(%)	4,356 (8.7)	16,224 (8.1)	45,450 (9.0)
Heart Failure, n(%)	707 (1.4)	3,009 (1.5)	8,939 (1.8)
Type 2 Diabetes, n(%)	2,817 (5.6)	11,150 (5.6)	29,266 (5.8)
Respiratory and immunological phenotypes			
Asthma, n(%)	8,471 (17.0)	28,491(14.2)	70,969 (14.1)
COPD, n(%)	1,481 (3.0)	5,135 (2.6)	14,688 (2.9)
Rheumatoid Arthritis, n(%)	951 (1.9)	3,802 (1.9)	9,685 (1.9)
Inflammatory Bowel Disease n(%)	650 (1.3)	2,616 (1.3)	6,851 (1.4)
Neurodegenerative phenotypes			
Alzheimer's Disease, n(%)	47 (0.09)	352 (0.18)	933 (0.19)
Parkinson's Disease, n(%)	131 (0.26)	721 (0.36)	1,938 (0.39)
Multiple Sclerosis, n(%)	163 (0.33)	684 (0.34)	1,766 (0.35)
Myasthenia Gravis, n(%)	23 (0.05)	129 (0.06)	300 (0.06)

Table 1: Demographics and clinical characteristics comparing publicly released whole exome sequencing on 50K, and 200K subjects together with all UKB participants.

Phenotypes were defined using a combination of IC10 codes together with self-reported verbal questionnaire. Values are expressed as median (1st and 3rd quartile). Data corresponds to UKB phenotypic release basket UKBB_41065 (March 2020). Subjects exclude participants withdrawn from UK Biobank as of September 2020.

Variants in WES			Median Per Participant (IQR)	
n=200,643 Participants	# Variants	# Variants MAF < 1%	# Variants	# Variants MAF < 1%
Total	16,846,188	16,709,690	44,719 (567)	3,050 (107)
Targeted Regions	8,457,134	8,395,913	21,380 (240)	1,466 (50)
Variant Type				
SNVs	8,086,176	8,026,702	20,817 (234)	1,403 (49)
Indels	370,958	369,211	563 (23)	64 (8)
Multi Allelic Sites	1,596,984	1,585,412	3768 (68)	223 (19)
Functional Prediction				
Synonymous	2,139,318	2,115,702	8,586 (114)	457 (24)
Missense	4,549,694	4,526,159	7,724 (115)	677 (32)
LOF (any transcript)	453,733	453,007	202 (15)	31 (7)
LOF (strict)	381,717	381,232	142 (12)	22 (6)

Table 2: Summary statistics for variants in publicly released 200K exomes.

Counts of autosomal variants observed across all individuals by type/functional class for all and for MAF<1%. The number of bases targeted for capture by the exome capture reagent was n=38,997,831. Median count and interquartile range (IQR) per individual for all variants, and for MAF<1%. Counts restricted to WES targeted regions. Variant annotation details are included in the Methods section.

LOF (AAF<1%)	189,698 individuals	500,000 individuals
Minimum number of carriers (N)	Observed number of genes with at least N heterozygous carriers	Predicted number of genes with at least N heterozygous carriers
1	18045	18414
5	17411	17968
10	16556	17514
25	14204	16384
50	11380	14853
100	7830	12475
250	3824	8065
500	1905	4711
1000	914	2449
2500	238	924
5000	18	394
10000	1	57

LOF (AAF<1%)	189,698 individuals	500,000 individuals
Minimum number of Carriers	Observed number of genes with at least N homozygotes	Predicted number of genes with at least N homozygotes
1	1492	1981
5	528	954
10	193	627
25	20	199
50	2	42
100	0	5
250	0	0
500	0	0
1000	0	0
2500	0	0
5000	0	0
10000	0	0

Table 3: Observed number of heterozygous and homozygous carriers of LOF variants in ~200K exomes and projections for the number expected in 500K. Predicted number of genes in larger WES sample sizes from existing 200K WES data estimated based on the methodology described in Van Hout et al, 2020. The number of autosomal genes with at least 1, 5, 10, etc. heterozygous and homozygous LOF carriers passing QC filters described in Table 2 for 189,698 UKB participants of European ancestry with WES and predicted number of genes with N heterozygous and homozygous LOF carriers in 500k.

Summary and closing thoughts

Transforming healthcare with data is at the heart of the UKB-ESC. We are proud of our contribution to the scientific community, enthusiastic about making these data broadly available, and excited to see what the future holds in terms of discoveries from and contributions to the UKB and UKB-ESC resources, particularly as we gain additional insight into functionally consequential variants that can meaningfully inform drug development. Data and methods developed by the UKB-ESC have already contributed to several publications (Liu et al., 2020).

With the UKB-ESC exome sequencing nearly complete, we hope the key features of this collaboration will be adopted as the preferred model for similar projects in the future. As an example, this pre-competitive industry collaboration model could be leveraged to generate complementary deep phenotyping data for UKB subjects and open entirely new avenues of scientific research. We expect the value of the WES data will be enhanced by layering deeper and richer phenotypes, which can provide important insights into disease biology, characterize response to marketed therapies, and identify novel target identification. Thoughtfully designed projects that maximize the engagement of academia, non-profits, and industry will yield valuable data resources and scientific insights that accelerate drug discovery and personalized healthcare.

There is no more important scientific challenge than developing new therapies to treat unmet medical needs. Large-scale collaborations such as the UKB-ESC play an essential role by generating unique, accessible resources that can be used by a diverse community of researchers to address questions critical to advancing human health.

Data accessibility

UK Biobank aims to encourage and provide as wide access as possible to its data and samples for health-related research in the public interest by all bona fide researchers from the academic, charity, public, and commercial sectors, both in the UK and internationally, without preferential access for any user. UK Biobank's publicly available Data Showcase (<http://biobank.ndph.ox.ac.uk/showcase/>) presents the univariate distributions and methods used for collection of all the variables available for health-related research, enabling potential research users to explore what data are available and plan research applications.

All researchers who wish to access the resource must register with UK Biobank via its online Access Management System (AMS) (<https://bbams.ndph.ox.ac.uk/ams/>). Once approved, researchers may apply (via the AMS) to access the resource for specific, well-defined research projects. At the time of publication, over 16,500 researchers were registered with UK Biobank and over 2,000 research applications were approved (see www.ukbiobank.ac.uk/approved-research/ for a summary of research that is currently underway).

Acknowledgements

The authors would like to thank everyone who made this work possible, particularly the UK Biobank team, their funders, the dedicated professionals from the member institutions who contributed to and supported this work, and most especially the UK Biobank participants, without whom this research would not be possible. This research has been conducted using the UK Biobank Resource under application number 26041.

Methods

OQFE Protocol

The OQFE protocol maps raw reads (FASTQ) with BWA-MEM to the GRCh38 reference in a deterministic manner, retaining all supplementary alignments. Mate tags are added with samblaster as specified in the FE protocol (Regier et al., 2018). OQFE CRAMs contain all reads from the input FASTQs and meet all FE tag specifications. Duplicate reads are then marked with Picard 2.21.2, which resolves a known issue with the FE version of Picard (2.4.1), in which the representative read in a set of duplicate reads can depend on the sequence input order, potentially resulting in an order-dependent set of supplementary alignment duplicates. The final OQFE CRAM is compressed with samtools, without any quality score recalibration or binning.

OQFE mapping and variant calling with UK Biobank exome sequencing

The OQFE protocol was applied to more than 200,000 UK Biobank (UKB 200K) exome samples with the OQFE pipeline (<https://hub.docker.com/r/dnanexus/oqfe>). Variants were called on each CRAM with DeepVariant 0.10.0 (Poplin et al., 2018) using a deep learning model retrained on exome data sequenced with the same protocol as was used to sequence the UKB 200K samples. Variant calls were restricted to the exome capture region and the 100 base-pairs flanking each capture target, resulting in a gVCF (genomic VCF) for each sample containing all variant genotypes and compressed representations of reference regions without called variant genotypes. All gVCFs were then merged with GLnexus 1.2.6 and the default 'DeepVariantWES' parameters (M. Lin et al., 2018; Yun et al., 2020).

Variant summary & Annotation

A set of QC metrics were applied to the variants included in Table 1. Variants that passed the following criteria are included: individual and variant missingness <10%, Hardy Weinberg Equilibrium p-value >10⁻¹⁵, minimum read coverage depth of 7 for SNPs and 10 for indels, at least one sample per site passed the allele balance threshold > 0.15 for SNPs and 0.20 for indels. Variants were annotated using SnpEff (Cingolani et al., 2012) based on Ensembl 85 gene model. Transcript aware annotations are derived using protein coding transcripts with defined start and stop coordinates.

Canonical transcript annotation

Transcripts with a “MANE Select v0.9” tag were obtained from Ensembl and were used to identify the canonical transcript of a gene. MANE annotation represents one high-quality transcript per protein-coding gene that is well-supported by experimental data and represents the biology of the gene. In the absence of a “MANE” annotation for a gene, APPRIS (Rodriguez et al., 2018) and Ensembl canonical annotation tags were used to define the canonical transcript set. Ensembl canonical annotation is obtained using the definition provided at <http://uswest.ensembl.org/Help/Glossary>

LOF annotation

While sequence-based annotation of LOF variants is straightforward, identifying true LOF events is not. We subset the predicted LOF variants to “strict LOF” category using the following criteria: 1) Excluded UTR splice variants, start lost and stop lost variants; 2) Excluded variants that are present in the last 5% of the protein and are predicted to escape nonsense-mediated decay; 3) Included only variants that affect the canonical transcript.

References

- Aartsma-Rus, A. (2018). Genetic therapies for spinal muscular atrophy type 1. In *The Lancet Neurology* (Vol. 17, Issue 2, pp. 111–112). Lancet Publishing Group. [https://doi.org/10.1016/S1474-4422\(17\)30436-2](https://doi.org/10.1016/S1474-4422(17)30436-2)
- Abifadel, M., Varret, M., Rabès, J. P., Allard, D., Ouguerram, K., Devillers, M., Cruaud, C., Benjannet, S., Wickham, L., Erlich, D., Derré, A., Villéger, L., Farnier, M., Beucler, I., Bruckert, E., Chambaz, J., Chanu, B., Lecerf, J. M., Luc, G., ... Boileau, C. (2003). Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nature Genetics*, *34*(2), 154–156. <https://doi.org/10.1038/ng1161>
- Abul-Husn, N. S., Cheng, X., Li, A. H., Xin, Y., Schurmann, C., Stevis, P., Liu, Y., Kozlitina, J., Stender, S., Wood, G. C., Stepanchick, A. N., Still, M. D., McCarthy, S., O'Dushlaine, C., Packer, J. S., Balasubramanian, S., Gosalia, N., Esopi, D., Kim, S. Y., ... Dewey, F. E. (2018). A Protein-Truncating *HSD17B13* Variant and Protection from Chronic Liver Disease. *New England Journal of Medicine*, *378*(12), 1096–1106. <https://doi.org/10.1056/NEJMoa1712191>
- Aittomäki, K., Dieguez Lucena, J., Pakarinen, P., Sistonen, P., Tapanainen, J., Gromoll, J., Kaskikari, R., Sankila, E. M., Lehväslaiho, H., Reyes Engel, A., Nieschlag, E., Huhtaniemi, I., & de la Chapelle, A. (1995). Mutation in the follicle-stimulating hormone receptor gene causes hereditary hypergonadotropic ovarian failure. *Cell*, *82*(6), 959–968. [https://doi.org/10.1016/0092-8674\(95\)90275-9](https://doi.org/10.1016/0092-8674(95)90275-9)
- Andersson, S., Berman, D. M., Jenkins, E. P., & Russell, D. W. (1991). Deletion of steroid 5 α -reductase 2 gene in male pseudohermaphroditism. *Nature*, *354*(6349), 159–161. <https://doi.org/10.1038/354159a0>
- Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Bentley, D. R., Chakravarti, A., Clark, A. G., Donnelly, P., Eichler, E. E., Flicek, P., Gabriel, S. B., Gibbs, R. A., Green, E. D., Hurles, M. E., Knoppers, B. M., Korbel, J. O., Lander, E. S., Lee, C., Lehrach, H., ... Schloss, J. A. (2015). A global reference for human genetic variation. In *Nature* (Vol. 526, Issue 7571, pp. 68–74). Nature Publishing Group. <https://doi.org/10.1038/nature15393>
- Ay Ling, R. M., Ross, R., Towner, P., Von Laue, S., Finidori, J., Moutoussamy, S., Buchanan, C. R., Clayton, P. E., & Norman, M. R. (1997). A dominant-negative mutation of the growth hormone receptor causes familial short stature. *Nature Genetics*, *16*(1), 13–14. <https://doi.org/10.1038/ng0597-13>
- Bai, M., Quinn, S., Trivedi, S., Kifor, O., Pearce, S. H. S., Pollak, M. R., Krapcho, K., Hebert, S. C., & Brown, E. M. (1996). Expression and characterization of inactivating and activating mutations in the human Ca²⁺(o)-sensing receptor. *Journal of Biological Chemistry*, *271*(32), 19537–19545. <https://doi.org/10.1074/jbc.271.32.19537>
- Berndt, S. I., Skibola, C. F., Joseph, V., Camp, N. J., Nieters, A., Wang, Z., Cozen, W., Monnereau, A., Wang, S. S., Kelly, R. S., Lan, Q., Teras, L. R., Chatterjee, N., Chung, C. C., Yeager, M., Brooks-Wilson, A. R., Hartge, P., Purdue, M. P., Birmann, B. M., ... Slager, S. L. (2013). Genome-wide association study identifies multiple risk loci for chronic lymphocytic leukemia. *Nature Genetics*, *45*(8), 868–876. <https://doi.org/10.1038/ng.2652>
- Bichet, D. G., Birnbaumer, M., Lonergan, M., Arthus, M. F., Rosenthal, W., Goodyer, P., Nivet, H., Benoit, S., Giampietro, P., Simonetti, S., Fish, A., Whitley, C. B., Jaeger, P., Gertner, J.,

- New, M., DiBona, F. J., Kaplan, B. S., Robertson, G. L., Hendy, G. N., ... Morgan, K. (1994). Nature and recurrence of AVPR2 mutations in X-linked nephrogenic diabetes insipidus. *American Journal of Human Genetics*, 55(2), 278–286.
- Burkhardt, R., Kenny, E. E., Lowe, J. K., Birkeland, A., Josowitz, R., Noel, M., Salit, J., Maller, J. B., Pe'er, I., Daly, M. J., Altshuler, D., Stoffel, M., Friedman, J. M., & Breslow, J. L. (2008). Common SNPs in HMGCR in micronesians and whites associated with LDL-cholesterol levels affect alternative splicing of exon13. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 28(11), 2078–2084. <https://doi.org/10.1161/ATVBAHA.108.172288>
- Burton, P. R., Clayton, D. G., Cardon, L. R., Craddock, N., Deloukas, P., Duncanson, A., Kwiatkowski, D. P., McCarthy, M. I., Ouwehand, W. H., Samani, N. J., Todd, J. A., Donnelly, P., Barrett, J. C., Davison, D., Easton, D., Evans, D., Leung, H. T., Marchini, J. L., Morris, A. P., ... Compston, A. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, 447(7145), 661–678. <https://doi.org/10.1038/nature05911>
- Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., Cortes, A., Welsh, S., Young, A., Effingham, M., McVean, G., Leslie, S., Allen, N., Donnelly, P., & Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature*, 562(7726), 203–209. <https://doi.org/10.1038/s41586-018-0579-z>
- Cargill, M., Schrodi, S. J., Chang, M., Garcia, V. E., Brandon, R., Callis, K. P., Matsunami, N., Ardlie, K. G., Civello, D., Catanese, J. J., Leong, D. U., Panko, J. M., McAllister, L. B., Hansen, C. B., Papenfuss, J., Prescott, S. M., White, T. J., Leppert, M. F., Krueger, G. G., & Begovich, A. B. (2007). A large-scale genetic association study confirms IL12B and leads to the identification of IL23R as psoriasis-risk genes. *American Journal of Human Genetics*, 80(2), 273–290. <https://doi.org/10.1086/511051>
- Carrington, M., Kissner, T., Gerrard, B., Ivanov, S., O'Brien, S. J., & Dean, M. (1997). Novel alleles of the chemokine-receptor gene CCR5. *American Journal of Human Genetics*, 61(6), 1261–1267. <https://doi.org/10.1086/301645>
- Carroll, P., Al-Mojalli, H., Al-Abbad, A., Al-Hassoun, I., Al-Hamed, M., Al-Amr, R., Butt, A. I., & Meyer, B. F. (2006). Novel mutations underlying nephrogenic diabetes insipidus in Arab families. *Genetics in Medicine*, 8(7), 443–447. <https://doi.org/10.1097/01.gim.0000223554.46981.7a>
- Cattaneo, M. (2011). The platelet P2Y12 receptor for adenosine diphosphate: Congenital and drug-induced defects. In *Blood* (Vol. 117, Issue 7, pp. 2102–2112). Blood. <https://doi.org/10.1182/blood-2010-08-263111>
- Chou, Y. H. W., Pollak, M. R., Brandi, M. L., Toss, G., Arnqvist, H., Atkinson, A. B., Papapoulos, S. E., Marx, S., Brown, E. M., Seidman, J. G., & Seidman, C. E. (1995). Mutations in the human Ca²⁺-sensing-receptor gene that cause familial hypocalciuric hypercalcemia. *American Journal of Human Genetics*, 56(5), 1075–1079.
- Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*, 6(2). <https://doi.org/10.4161/fly.19695>
- Cohen, J., Pertsemlidis, A., Kotowski, I. K., Graham, R., Garcia, C. K., & Hobbs, H. H. (2005). Low

- LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nature Genetics*, 37(2), 161–165. <https://doi.org/10.1038/ng1509>
- Cole, D. E. C., Yun, F. H. J., Wong, B. Y. L., Shuen, A. Y., Booth, R. A., Scillitani, A., Pidasheva, S., Zhou, X., Canaff, L., & Hendy, G. N. (2009). Calcium-sensing receptor mutations and denaturing high performance liquid chromatography. *Journal of Molecular Endocrinology*, 42(4), 331–339. <https://doi.org/10.1677/JME-08-0164>
- Colombo, C., Porzio, O., Liu, M., Massa, O., Vasta, M., Salardi, S., Beccaria, L., Monciotti, C., Toni, S., Pedersen, O., Hansen, T., Federici, L., Pesavento, R., Cadario, F., Federici, G., Ghirri, P., Arvan, P., Iafusco, D., & Barbetti, F. (2008). Seven mutations in the human insulin gene linked to permanent neonatal/infancy-onset diabetes mellitus. *Journal of Clinical Investigation*, 118(6), 2148–2156. <https://doi.org/10.1172/JCI33777>
- Cook, D., Brown, D., Alexander, R., March, R., Morgan, P., Satterthwaite, G., & Pangalos, M. N. (2014). Lessons learned from the fate of AstraZeneca's drug pipeline: A five-dimensional framework. In *Nature Reviews Drug Discovery* (Vol. 13, Issue 6, pp. 419–431). Nature Publishing Group. <https://doi.org/10.1038/nrd4309>
- Dolgin, E. (2019). Massive NIH–industry project opens portals to target validation. *Nature Reviews Drug Discovery*. <https://doi.org/10.1038/d41573-019-00033-8>
- Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., Epstein, C. B., Frietze, S., Harrow, J., Kaul, R., Khatun, J., Lajoie, B. R., Landt, S. G., Lee, B. K., Pauli, F., Rosenbloom, K. R., Sabo, P., Safi, A., Sanyal, A., ... Lochovsky, L. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489(7414), 57–74. <https://doi.org/10.1038/nature11247>
- Edghill, E. L., Flanagan, S. E., Patch, A. M., Boustred, C., Parrish, A., Shields, B., Shepherd, M. H., Hussain, K., Kapoor, R. R., Malecki, M., MacDonald, M. J., Støy, J., Steiner, D. F., Philipson, L. H., Bell, G. I., Hattersley, A. T., & Ellard, S. (2008). Insulin mutation screening in 1,044 patients with diabetes mutations in the INS gene are a common cause of neonatal diabetes but a rare cause of diabetes diagnosed in childhood or adulthood. *Diabetes*, 57(4), 1034–1042. <https://doi.org/10.2337/db07-1405>
- Estrada, K., Styrkarsdottir, U., Evangelou, E., Hsu, Y. H., Duncan, E. L., Ntzani, E. E., Oei, L., Albagha, O. M. E., Amin, N., Kemp, J. P., Koller, D. L., Li, G., Liu, C. T., Minster, R. L., Moayyeri, A., Vandenput, L., Willner, D., Xiao, S. M., Yerges-Armstrong, L. M., ... Rivadeneira, F. (2012). Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nature Genetics*, 44(5), 491–501. <https://doi.org/10.1038/ng.2249>
- Eyre, S., Bowes, J., Diogo, D., Lee, A., Barton, A., Martin, P., Zhernakova, A., Stahl, E., Viatte, S., McAllister, K., Amos, C. I., Padyukov, L., Toes, R. E. M., Huizinga, T. W. J., Wijmenga, C., Trynka, G., Franke, L., Westra, H. J., Alfredsson, L., ... Bentley, D. (2012). High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. *Nature Genetics*, 44(12), 1336–1340. <https://doi.org/10.1038/ng.2462>
- Flannick, J., Mercader, J. M., Fuchsberger, C., Udler, M. S., Mahajan, A., Wessel, J., Teslovich, T. M., Caulkins, L., Koesterer, R., Barajas-Olmos, F., Blackwell, T. W., Boerwinkle, E., Brody, J. A., Centeno-Cruz, F., Chen, L., Chen, S., Contreras-Cubas, C., Córdova, E., Correa, A., ... Boehnke, M. (2019). Exome sequencing of 20,791 cases of type 2 diabetes and 24,440 controls. *Nature*, 570(7759), 71–76. <https://doi.org/10.1038/s41586-019-1231-2>

- Furtado, R. H. M., & Giugliano, R. P. (2020). What Lessons Have We Learned and What Remains to be Clarified for PCSK9 Inhibitors? A Review of FOURIER and ODYSSEY Outcomes Trials. *Cardiology and Therapy*, 1–15. <https://doi.org/10.1007/s40119-020-00163-w>
- Giri, A., Hellwege, J. N., Keaton, J. M., Park, J., Qiu, C., Warren, H. R., Torstenson, E. S., Kovesdy, C. P., Sun, Y. V., Wilson, O. D., Robinson-Cohen, C., Roumie, C. L., Chung, C. P., Birdwell, K. A., Damrauer, S. M., DuVall, S. L., Klarin, D., Cho, K., Wang, Y., ... Edwards, T. L. (2019). Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. *Nature Genetics*, 51(1), 51–62. <https://doi.org/10.1038/s41588-018-0303-9>
- Goddard, A. D., Covello, R., Luoh, S. M., Clackson, T., Attie, K. M., Gesundheit, N., Rundle, A. C., Wells, J. A., & Carlsson, L. M. S. (1995). Mutations of the growth hormone receptor in children with idiopathic short stature. *New England Journal of Medicine*, 333(17), 1093–1098. <https://doi.org/10.1056/NEJM199510263331701>
- Hamersma, H., Gardner, J., & Beighton, P. (2003). The natural history of sclerosteosis. *Clinical Genetics*, 63(3), 192–197. <https://doi.org/10.1034/j.1399-0004.2003.00036.x>
- Hinds, D. A., Buil, A., Ziemek, D., Martinez-Perez, A., Malik, R., Folkersen, L., Germain, M., Mälarstig, A., Brown, A., Soria, J. M., Dichgans, M., Bing, N., Franco-Cereceda, A., Souto, J. C., Dermitzakis, E. T., Hamsten, A., Worrall, B. B., Tung, J. Y., & Sabater-Lleal, M. (2016). Genome-wide association analysis of self-reported events in 6135 individuals and 252 827 controls identifies 8 loci associated with thrombosis. *Human Molecular Genetics*, 25(9), 1867–1874. <https://doi.org/10.1093/hmg/ddw037>
- Hoffmann, T. J., Ehret, G. B., Nandakumar, P., Ranatunga, D., Schaefer, C., Kwok, P. Y., Iribarren, C., Chakravarti, A., & Risch, N. (2017). Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. *Nature Genetics*, 49(1), 54–64. <https://doi.org/10.1038/ng.3715>
- Imperato-McGinley, J., Guerrero, L., Gautier, T., & Peterson, R. E. (1974). Steroid 5 α -reductase deficiency in man: An inherited form of male pseudohermaphroditism. *Science*, 186(4170), 1213–1215. <https://doi.org/10.1126/science.186.4170.1213>
- Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., Gauthier, L. D., Brand, H., Solomonson, M., Watts, N. A., Rhodes, D., Singer-Berk, M., England, E. M., Seaby, E. G., Kosmicki, J. A., ... MacArthur, D. G. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, 581(7809), 434–443. <https://doi.org/10.1038/s41586-020-2308-7>
- Kathiresan, S., Melander, O., Guiducci, C., Surti, A., Burt, N. P., Rieder, M. J., Cooper, G. M., Roos, C., Voight, B. F., Havulinna, A. S., Wahlstrand, B., Hedner, T., Corella, D., Tai, E. S., Ordovas, J. M., Berglund, G., Vartiainen, E., Jousilahti, P., Hedblad, B., ... Orho-Melander, M. (2008). Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nature Genetics*, 40(2), 189–197. <https://doi.org/10.1038/ng.75>
- Khatun, I., Walsh, M. T., & Hussain, M. M. (2013). Loss of both phospholipid and triglyceride transfer activities of microsomal triglyceride transfer protein in abetalipoproteinemia. *Journal of Lipid Research*, 54(6), 1541–1549. <https://doi.org/10.1194/jlr.M031658>
- Kichaev, G., Bhatia, G., Loh, P. R., Gazal, S., Burch, K., Freund, M. K., Schoech, A., Pasaniuc, B., & Price, A. L. (2019). Leveraging Polygenic Functional Enrichment to Improve GWAS Power.

- American Journal of Human Genetics*, 104(1), 65–75.
<https://doi.org/10.1016/j.ajhg.2018.11.008>
- King, E. A., Wade Davis, J., & Degner, J. F. (2019). Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval. *PLoS Genetics*, 15(12), e1008489.
<https://doi.org/10.1371/journal.pgen.1008489>
- Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., Fitzhugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczy, J., Levine, R., McEwan, P., ... Morgan, M. J. (2001). Initial sequencing and analysis of the human genome. *Nature*, 409(6822), 860–921.
<https://doi.org/10.1038/35057062>
- Laufer, V. A., Tiwari, H. K., Reynolds, R. J., Danila, M. I., Wang, J., Edberg, J. C., Kimberly, R. P., Kottyan, L. C., Harley, J. B., Mikuls, T. R., Gregersen, P. K., Absher, D. M., Langefeld, C. D., Arnett, D. K., & Bridges, S. L. (2019). Genetic influences on susceptibility to rheumatoid arthritis in African-Americans. *Human Molecular Genetics*, 28(5), 858–874.
<https://doi.org/10.1093/hmg/ddy395>
- Li, Z., Chen, J., Yu, H., He, L., Xu, Y., Zhang, D., Yi, Q., Li, C., Li, X., Shen, J., Song, Z., Ji, W., Wang, M., Zhou, J., Chen, B., Liu, Y., Wang, J., Wang, P., Yang, P., ... Shi, Y. (2017). Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. *Nature Genetics*, 49(11), 1576–1583. <https://doi.org/10.1038/ng.3973>
- Lin, J. M., Wan, L., Tsai, Y. Y., Lin, H. J., Tsai, Y., Lee, C. C., Tsai, C. H., Tseng, S. H., & Tsai, F. J. (2008). Vascular Endothelial Growth Factor Gene Polymorphisms in Age-related Macular Degeneration. *American Journal of Ophthalmology*, 145(6).
<https://doi.org/10.1016/j.ajo.2008.01.027>
- Lin, M., Rodeh, O., Penn, J., Bai, X., Reid, J., Krasheninina, O., & Salerno, W. (2018). GLnexus: joint variant calling for large cohort sequencing. *BioRxiv*, 343970.
<https://doi.org/10.1101/343970>
- Liu, J. Z., Chen, C.-Y., Tsai, E. A., Whelan, C. D., Sexton, D., John, S., & Runz, H. (2020). The burden of rare protein-truncating genetic variants on human lifespan. *BioRxiv*, 2020.06.02.129908. <https://doi.org/10.1101/2020.06.02.129908>
- Mahajan, A., Wessel, J., Willems, S. M., Zhao, W., Robertson, N. R., Chu, A. Y., Gan, W., Kitajima, H., Taliun, D., Rayner, N. W., Guo, X., Lu, Y., Li, M., Jensen, R. A., Hu, Y., Huo, S., Lohman, K. K., Zhang, W., Cook, J. P., ... McCarthy, M. I. (2018). Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes article. *Nature Genetics*, 50(4), 559–571. <https://doi.org/10.1038/s41588-018-0084-1>
- McCarthy, M. I., Abecasis, G. R., Cardon, L. R., Goldstein, D. B., Little, J., Ioannidis, J. P. A., & Hirschhorn, J. N. (2008). Genome-wide association studies for complex traits: Consensus, uncertainty and challenges. *Nature Reviews Genetics*, 9(5), 356–369.
<https://doi.org/10.1038/nrg2344>
- McCarthy, S., Das, S., Kretschmar, W., Delaneau, O., Wood, A. R., Teumer, A., Kang, H. M., Fuchsberger, C., Danecek, P., Sharp, K., Luo, Y., Sidore, C., Kwong, A., Timpson, N., Koskinen, S., Vrieze, S., Scott, L. J., Zhang, H., Mahajan, A., ... Marchini, J. (2016). A reference panel of 64,976 haplotypes for genotype imputation. *Nature Genetics*, 48(10), 1279–1283. <https://doi.org/10.1038/ng.3643>

- Messier, T. L., Wong, C. Y., Bovill, E. G., Long, G. L., & Church, W. R. (1996). Factor X Stockton: a mild bleeding diathesis associated with an active site mutation in factor X. *Blood Coagulation & Fibrinolysis*: An International Journal in Haemostasis and Thrombosis, 7(1), 5–14.
- Methé, B. A., Nelson, K. E., Pop, M., Creasy, H. H., Giglio, M. G., Huttenhower, C., Gevers, D., Petrosino, J. F., Abubucker, S., Badger, J. H., Chinwalla, A. T., Earl, A. M., Fitzgerald, M. G., Fulton, R. S., Hallsworth-Pepin, K., Lobos, E. A., Madupu, R., Magrini, V., Martin, J. C., ... White, O. (2012). A framework for human microbiome research. *Nature*, 486(7402), 215–221. <https://doi.org/10.1038/nature11209>
- Morris, A. P., Voight, B. F., Teslovich, T. M., Ferreira, T., Segrè, A. V., Steinthorsdottir, V., Strawbridge, R. J., Khan, H., Grallert, H., Mahajan, A., Prokopenko, I., Kang, H. M., Dina, C., Esko, T., Fraser, R. M., Kanoni, S., Kumar, A., Lagou, V., Langenberg, C., ... McCarthy, M. I. (2012). Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nature Genetics*, 44(9), 981–990. <https://doi.org/10.1038/ng.2383>
- Morris, J. A., Kemp, J. P., Youlten, S. E., Laurent, L., Logan, J. G., Chai, R. C., Vulpescu, N. A., Forgetta, V., Kleinman, A., Mohanty, S. T., Sergio, C. M., Quinn, J., Nguyen-Yamamoto, L., Luco, A. L., Vijay, J., Simon, M. M., Pramatarova, A., Medina-Gomez, C., Trajanoska, K., ... Richards, J. B. (2019). An atlas of genetic influences on osteoporosis in humans and mice. *Nature Genetics*, 51(2), 258–266. <https://doi.org/10.1038/s41588-018-0302-x>
- Nelson, M. R., Tipney, H., Painter, J. L., Shen, J., Nicoletti, P., Shen, Y., Floratos, A., Sham, P. C., Li, M. J., Wang, J., Cardon, L. R., Whittaker, J. C., & Sanseau, P. (2015). The support of human genetic evidence for approved drug indications. *Nature Genetics*, 47(8), 856–860. <https://doi.org/10.1038/ng.3314>
- Nestorowicz, A., Wilson, B. A., Schoor, K. P., Inoue, H., Glaser, B., Landau, H., Stanley, C. A., Thornton, P. S., Clement IV, J. P., Bryan, J., Aguilar-Bryan, L., & Permutt, M. A. (1996). Mutations in the sulfonylurea receptor gene are associated with familial hyperinsulinism in Ashkenazi Jews. *Human Molecular Genetics*, 5(11), 1813–1822. <https://doi.org/10.1093/hmg/5.11.1813>
- Nguyen, P. A., Born, D. A., Deaton, A. M., Nioi, P., & Ward, L. D. (2019). Phenotypes associated with genes encoding drug targets are predictive of clinical trial side effects. *Nature Communications*, 10(1). <https://doi.org/10.1038/s41467-019-09407-3>
- Nioi, P., Sigurdsson, A., Thorleifsson, G., Helgason, H., Agustsdottir, A. B., Norddahl, G. L., Helgadottir, A., Magnusdottir, A., Jonasdottir, A., Gretarsdottir, S., Jonsdottir, I., Steinthorsdottir, V., Rafnar, T., Swinkels, D. W., Galesloot, T. E., Grarup, N., Jørgensen, T., Vestergaard, H., Hansen, T., ... Stefansson, K. (2016). Variant ASGR1 Associated with a Reduced Risk of Coronary Artery Disease. *New England Journal of Medicine*, 374(22), 2131–2141. <https://doi.org/10.1056/NEJMoa1508419>
- Nissen, P. H., Christensen, S. E., Heickendorff, L., Brixen, K., & Mosekilde, L. (2007). Molecular genetic analysis of the calcium sensing receptor gene in patients clinically suspected to have familial hypocalciuric hypercalcemia: Phenotypic variation and mutation spectrum in a Danish population. *Journal of Clinical Endocrinology and Metabolism*, 92(11), 4373–4379. <https://doi.org/10.1210/jc.2007-0322>
- Novembre, J., Galvani, A. P., & Slatkin, M. (2005). The geographic spread of the CCR5 Δ 32 HIV-

- resistance allele. *PLoS Biology*, 3(11), 1954–1962.
<https://doi.org/10.1371/journal.pbio.0030339>
- Okada, Y., Wu, D., Trynka, G., Raj, T., Terao, C., Ikari, K., Kochi, Y., Ohmura, K., Suzuki, A., Yoshida, S., Graham, R. R., Manoharan, A., Ortmann, W., Bhangale, T., Denny, J. C., Carroll, R. J., Eyler, A. E., Greenberg, J. D., Kremer, J. M., ... Plenge, R. M. (2014). Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature*, 506(7488), 376–381. <https://doi.org/10.1038/nature12873>
- Otonkoski, T., Ämmälä, C., Huopio, H., Cote, G. J., Chapman, J., Cosgrove, K., Ashfield, R., Huang, E., Komulainen, J., Ashcroft, F. M., Dunne, M. J., Kere, J., & Thomas, P. M. (1999). A point mutation inactivating the sulfonylurea receptor causes the severe form of persistent hyperinsulinemic hypoglycemia of infancy in Finland. *Diabetes*, 48(2), 408–415.
<https://doi.org/10.2337/diabetes.48.2.408>
- Pardiñas, A. F., Holmans, P., Pocklington, A. J., Escott-Price, V., Ripke, S., Carrera, N., Legge, S. E., Bishop, S., Cameron, D., Hamshere, M. L., Han, J., Hubbard, L., Lynham, A., Mantripragada, K., Rees, E., MacCabe, J. H., McCarroll, S. A., Baune, B. T., Breen, G., ... Walters, J. T. R. (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature Genetics*, 50(3), 381–389.
<https://doi.org/10.1038/s41588-018-0059-2>
- Pasel, K., Schulz, A., Timmermann, K., Linnemann, K., Hoeltzenbein, M., Jääskeläinen, J., Grüters, A., Filler, G., & Schöneberg, T. (2000). Functional Characterization of the Molecular Defects Causing Nephrogenic Diabetes Insipidus in Eight Families 1. *The Journal of Clinical Endocrinology & Metabolism*, 85(4), 1703–1710.
<https://doi.org/10.1210/jcem.85.4.6507>
- Poplin, R., Chang, P. C., Alexander, D., Schwartz, S., Colthurst, T., Ku, A., Newburger, D., Dijamco, J., Nguyen, N., Afshar, P. T., Gross, S. S., Dorfman, L., McLean, C. Y., & Depristo, M. A. (2018). A universal snp and small-indel variant caller using deep neural networks. *Nature Biotechnology*, 36(10), 983. <https://doi.org/10.1038/nbt.4235>
- Pullinger, C. R., Hennessy, L. K., Chatterton, J. E., Liu, W., Love, J. A., Mendel, C. M., Frost, P. H., Malloy, M. J., Schumaker, V. N., & Kane, J. P. (1995). Familial Ligand-Defective Apolipoprotein B: Identification of a New Mutation That Decreases LDL Receptor Binding Affinity. *Journal of Clinical Investigation*, 95(3), 1225–1234.
<https://doi.org/10.1172/JCI117772>
- Quillent, C., Oberlin, E., Braun, J., Rousset, D., Gonzalez-Canali, G., Métais, P., Montagnier, L., Virelizier, J. L., Arenzana-Seisdedos, F., & Beretta, A. (1998). HIV-1-resistance phenotype conferred by combination of two separate inherited mutations of CCR5 gene. *Lancet*, 351(9095), 14–18. [https://doi.org/10.1016/S0140-6736\(97\)09185-X](https://doi.org/10.1016/S0140-6736(97)09185-X)
- Rabès, J. P., Varret, M., Saint-Jore, B., Erlich, D., Jondeau, G., Krempf, M., Giraudet, P., Junien, C., & Boileau, C. (1997). Familial ligand-defective apolipoprotein b-100: Simultaneous detection of the arg3500 → GLN and ARG3531→CYS mutations in a French population. *Human Mutation*, 10(2), 160–163. [https://doi.org/10.1002/\(SICI\)1098-1004\(1997\)10:2<160::AID-HUMU8>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1098-1004(1997)10:2<160::AID-HUMU8>3.0.CO;2-O)
- Rader, D. J., Cohen, J., & Hobbs, H. H. (2003). Monogenic hypercholesterolemia: new insights in pathogenesis and treatment. *Journal of Clinical Investigation*, 111(12), 1795–1803.
<https://doi.org/10.1172/jci18925>

- Regier, A. A., Farjoun, Y., Larson, D. E., Krasheninina, O., Kang, H. M., Howrigan, D. P., Chen, B. J., Kher, M., Banks, E., Ames, D. C., English, A. C., Li, H., Xing, J., Zhang, Y., Matise, T., Abecasis, G. R., Salerno, W., Zody, M. C., Neale, B. M., & Hall, I. M. (2018). Functional equivalence of genome sequencing analysis pipelines enables harmonized variant calling across human genetics projects. *Nature Communications*, *9*(1), 4038. <https://doi.org/10.1038/s41467-018-06159-4>
- Ripke, S., Neale, B. M., Corvin, A., Walters, J. T. R., Farh, K. H., Holmans, P. A., Lee, P., Bulik-Sullivan, B., Collier, D. A., Huang, H., Pers, T. H., Agartz, I., Agerbo, E., Albus, M., Alexander, M., Amin, F., Bacanu, S. A., Begemann, M., Belliveau, R. A., ... O'Donovan, M. C. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, *511*(7510), 421–427. <https://doi.org/10.1038/nature13595>
- Rodriguez, J. M., Rodriguez-Rivas, J., Di Domenico, T., Vázquez, J., Valencia, A., & Tress, M. L. (2018). APPRIS 2017: Principal isoforms for multiple gene sets. *Nucleic Acids Research*, *46*(D1). <https://doi.org/10.1093/nar/gkx997>
- Rosenthal, W., Seibold, A., Antaramian, A., Lonergan, M., Arthus, M. F., Hendy, G. N., Birnbaumer, M., & Bichet, D. G. (1992). Molecular identification of the gene responsible for congenital nephrogenic diabetes insipidus. *Nature*, *359*(6392), 233–235. <https://doi.org/10.1038/359233a0>
- Sabatine, M. S., Giugliano, R. P., Keech, A. C., Honarpour, N., Wiviott, S. D., Murphy, S. A., Kuder, J. F., Wang, H., Liu, T., Wasserman, S. M., Sever, P. S., Pedersen, T. R., Fish, M. P., Abrahamsen, T. E., Im, K., Kanevsky, E., Bonaca, M. P., Lira Pineda, A., Hanlon, K., ... Shifrin, G. (2017). Evolocumab and clinical outcomes in patients with cardiovascular disease. *New England Journal of Medicine*, *376*(18), 1713–1722. <https://doi.org/10.1056/NEJMoa1615664>
- Schwartz, G. G., Steg, P. G., Szarek, M., Bhatt, D. L., Bittner, V. A., Diaz, R., Edelberg, J. M., Goodman, S. G., Hanotin, C., Harrington, R. A., Jukema, J. W., Lecorps, G., Mahaffey, K. W., Moryusef, A., Pordy, R., Quintero, K., Roe, M. T., Sasiela, W. J., Tamby, J. F., ... Ball, E. (2018). Alirocumab and cardiovascular outcomes after acute coronary syndrome. *New England Journal of Medicine*, *379*(22), 2097–2107. <https://doi.org/10.1056/NEJMoa1801174>
- Sinner, M. F., Pfeufer, A., Akyol, M., Beckmann, B. M., Hinterseer, M., Wacker, A., Perz, S., Sauter, W., Illig, T., Näbauer, M., Schmitt, C., Wichmann, H. E., Schömig, A., Steinbeck, G., Meitinger, T., & Kääb, S. (2008). The non-synonymous coding IKr-channel variant KCNH2-K897T is associated with atrial fibrillation: Results from a systematic candidate gene-based analysis of KCNH2 (HERG). *European Heart Journal*, *29*(7), 907–914. <https://doi.org/10.1093/eurheartj/ehm619>
- Soria, L. F., Ludwig, E. H., Clarke, H. R. G., Vega, G. L., Grundy, S. M., & McCarthy, B. J. (1989). Association between a specific apolipoprotein B mutation and familial defective apolipoprotein B-100. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(2), 587–591. <https://doi.org/10.1073/pnas.86.2.587>
- Stitzel, N. O., Won, H. H., Morrison, A. C., Peloso, G. M., Do, R., Lange, L. A., Fontanillas, P., Gupta, N., Duga, S., Goel, A., Farrall, M., Saleheen, D., Ferrario, P., König, I., Asselta, R., Merlini, P. A., Marziliano, N., Notarangelo, M. F., Schick, U., ... Kathiresan, S. (2014). Inactivating mutations in NPC1L1 and protection from coronary heart disease. *New*

- England Journal of Medicine*, 371(22), 2072–2082.
<https://doi.org/10.1056/NEJMoa1405386>
- Støy, J., Edghill, E. L., Flanagan, S. E., Ye, H., Paz, V. P., Pluzhnikov, A., Below, J. E., Hayes, M. G., Cox, N. J., Lipkind, G. M., Lipton, R. B., Greeley, S. A. W., Patch, A. M., Ellard, S., Steiner, D. F., Hattersley, A. T., Philipson, L. H., & Bell, G. I. (2007). Insulin gene mutations as a cause of permanent neonatal diabetes. *Proceedings of the National Academy of Sciences of the United States of America*, 104(38), 15040–15044.
<https://doi.org/10.1073/pnas.0707291104>
- Styrkarsdottir, U., Halldorsson, B. V., Gretarsdottir, S., Gudbjartsson, D. F., Walters, G. B., Ingvarsson, T., Jonsdottir, T., Saemundsdottir, J., Snorradottir, S., Center, J. R., Nguyen, T. V., Alexandersen, P., Gulcher, J. R., Eisman, J. A., Christiansen, C., Sigurdsson, G., Kong, A., Thorsteinsdottir, U., & Stefansson, K. (2009). New sequence variants associated with bone mineral density. *Nature Genetics*, 41(1), 15–17. <https://doi.org/10.1038/ng.284>
- Styrkarsdottir, U., Thorleifsson, G., Gudjonsson, S. A., Sigurdsson, A., Center, J. R., Lee, S. H., Nguyen, T. V., Kwok, T. C. Y., Lee, J. S. W., Ho, S. C., Woo, J., Leung, P. C., Kim, B. J., Rafnar, T., Kiemenev, L. A., Ingvarsson, T., Koh, J. M., Tang, N. L. S., Eisman, J. A., ... Stefansson, K. (2016). Sequence variants in the PTCH1 gene associate with spine bone mineral density and osteoporotic fractures. *Nature Communications*, 7.
<https://doi.org/10.1038/ncomms10129>
- Sullivan, A. D., Wigginton, J., & Kirschner, D. (2001). The coreceptor mutation CCR5Δ32 influences the dynamics of hiv epidemics and is selected for by HIV. *Proceedings of the National Academy of Sciences of the United States of America*, 98(18), 10214–10219.
<https://doi.org/10.1073/pnas.181325198>
- Sung, Y. J., Winkler, T. W., de las Fuentes, L., Bentley, A. R., Brown, M. R., Kraja, A. T., Schwander, K., Ntalla, I., Guo, X., Franceschini, N., Lu, Y., Cheng, C. Y., Sim, X., Vojinovic, D., Marten, J., Musani, S. K., Li, C., Feitosa, M. F., Kilpeläinen, T. O., ... Chasman, D. I. (2018). A Large-Scale Multi-ancestry Genome-wide Study Accounting for Smoking Behavior Identifies Multiple Significant Loci for Blood Pressure. *American Journal of Human Genetics*, 102(3), 375–400. <https://doi.org/10.1016/j.ajhg.2018.01.015>
- Takeuchi, F., Akiyama, M., Matoba, N., Katsuya, T., Nakatochi, M., Tabara, Y., Narita, A., Saw, W. Y., Moon, S., Spracklen, C. N., Chai, J. F., Kim, Y. J., Zhang, L., Wang, C., Li, H., Li, H., Wu, J. Y., Dorajoo, R., Nierenberg, J. L., ... Kato, N. (2018). Interethnic analyses of blood pressure loci in populations of East Asian and European descent. *Nature Communications*, 9(1).
<https://doi.org/10.1038/s41467-018-07345-0>
- The Cancer Genome Atlas Program - National Cancer Institute*. (n.d.).
- Thomas, P. M., Wohlk, N., Huang, E., Kuhnle, U., Rabl, W., Gagel, R. F., & Cote, G. J. (1996). Inactivation of the first nucleotide-binding fold of the sulfonylurea receptor, and familial persistent hyperinsulinemic hypoglycemia of infancy. *American Journal of Human Genetics*, 59(3), 510–518.
- Van Der Graaf, A., Avis, H. J., Kusters, D. M., Vissers, M. N., Hutten, B. A., Defesche, J. C., Huijgen, R., Fouchier, S. W., Wijburg, F. A., Kastelein, J. J. P., & Wiegman, A. (2011). Molecular basis of autosomal dominant hypercholesterolemia: Assessment in a large cohort of hypercholesterolemic children. *Circulation*, 123(11), 1167–1173.
<https://doi.org/10.1161/CIRCULATIONAHA.110.979450>

- Van Hout, C. V., Tachmazidou, I., Backman, J. D., Hoffman, J. D., Liu, D., Pandey, A. K., Gonzaga-Jauregui, C., Khalid, S., Ye, B., Banerjee, N., Li, A. H., O'Dushlaine, C., Marcketta, A., Staples, J., Schurmann, C., Hawes, A., Maxwell, E., Barnard, L., Lopez, A., ... Baras, A. (2020). Exome sequencing and characterization of 49,960 individuals in the UK Biobank. *Nature*. <https://doi.org/10.1038/s41586-020-2853-0>
- Watzke, H. H., Lechner, K., Roberts, H. R., Reddy, S. V., Welsch, D. J., Friedman, P., Mahr, G., Jagadeeswaran, P., Monroe, D. M., & High, K. A. (1990). Molecular defect (Gla+14 → Lys) and its functional consequences in a hereditary factor X deficiency (factor X 'Vorarlberg'). *Journal of Biological Chemistry*, *265*(20), 11982–11989.
- Willer, C. J., Schmidt, E. M., Sengupta, S., Peloso, G. M., Gustafsson, S., Kanoni, S., Ganna, A., Chen, J., Buchkovich, M. L., Mora, S., Beckmann, J. S., Bragg-Gresham, J. L., Chang, H. Y., Demirkan, A., Den Hertog, H. M., Do, R., Donnelly, L. A., Ehret, G. B., Esko, T., ... Abecasis, G. R. (2013). Discovery and refinement of loci associated with lipid levels. *Nature Genetics*, *45*(11), 1274–1285. <https://doi.org/10.1038/ng.2797>
- Winkler, T. W., Justice, A. E., Graff, M., Barata, L., Feitosa, M. F., Chu, S., Czajkowski, J., Esko, T., Fall, T., Kilpeläinen, T. O., Lu, Y., Mägi, R., Mihailov, E., Pers, T. H., Rieger, S., Teumer, A., Ehret, G. B., Ferreira, T., Heard-Costa, N. L., ... Loos, R. J. F. (2015). The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. *PLoS Genetics*, *11*(10), 1–42. <https://doi.org/10.1371/journal.pgen.1005378>
- Wood, A. J. J., & Rittmaster, R. S. (1994). Finasteride. In *New England Journal of Medicine* (Vol. 330, Issue 2, pp. 120–125). <https://doi.org/10.1056/NEJM199401133300208>
- Xiong, H. Y., Alipanahi, B., Lee, L. J., Bretschneider, H., Merico, D., Yuen, R. K. C., Hua, Y., Gueroussov, S., Najafabadi, H. S., Hughes, T. R., Morris, Q., Barash, Y., Krainer, A. R., Jovic, N., Scherer, S. W., Blencowe, B. J., & Frey, B. J. (2015). The human splicing code reveals new insights into the genetic determinants of disease. *Science*, *347*(6218). <https://doi.org/10.1126/science.1254806>
- Xue, A., Wu, Y., Zhu, Z., Zhang, F., Kemper, K. E., Zheng, Z., Yengo, L., Lloyd-Jones, L. R., Sidorenko, J., Wu, Y., Agbessi, M., Ahsan, H., Alves, I., Andiappan, A., Awadalla, P., Battle, A., Beutner, F., Bonder, M. J. J., Boomsma, D., ... Yang, J. (2018). Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. *Nature Communications*, *9*(1). <https://doi.org/10.1038/s41467-018-04951-w>
- Yang, W., Tang, H., Zhang, Y., Tang, X., Zhang, J., Sun, L., Yang, J., Cui, Y., Zhang, L., Hirankarn, N., Cheng, H., Pan, H. F., Gao, J., Lee, T. L., Sheng, Y., Lau, C. S., Li, Y., Chan, T. M., Yin, X., ... Lau, Y. L. (2013). Meta-analysis followed by replication identifies loci in or near CDKN1B, TET3, CD80, DRAM1, and ARID5B as associated with systemic lupus erythematosus in Asians. *American Journal of Human Genetics*, *92*(1), 41–51. <https://doi.org/10.1016/j.ajhg.2012.11.018>
- Yang, Y., Muzny, D. M., Reid, J. G., Bainbridge, M. N., Willis, A., Ward, P. A., Braxton, A., Beuten, J., Xia, F., Niu, Z., Hardison, M., Person, R., Bekheirnia, M. R., Leduc, M. S., Kirby, A., Pham, P., Scull, J., Wang, M., Ding, Y., ... Eng, C. M. (2013). Clinical Whole-Exome Sequencing for the Diagnosis of Mendelian Disorders. *New England Journal of Medicine*, *369*(16), 1502–1511. <https://doi.org/10.1056/NEJMoa1306555>
- Yu, Y., Bhangale, T. R., Fagerness, J., Ripke, S., Thorleifsson, G., Tan, P. L., Souied, E. H.,

- Richardson, A. J., Merriam, J. E., Buitendijk, G. H. S., Reynolds, R., Raychaudhuri, S., Chin, K. A., Sobrin, L., Evangelou, E., Lee, P. H., Lee, A. Y., Leveziel, N., Zack, D. J., ... Seddon, J. M. (2011). Common variants near FRK/COL10A1 and VEGFA are associated with advanced age-related macular degeneration. *Human Molecular Genetics*, *20*(18), 3699–3709. <https://doi.org/10.1093/hmg/ddr270>
- Yun, T., Li, H., Chang, P.-C., Lin, M., Carroll, A., & McLean, C. (2020). Accurate, scalable cohort variant calls using DeepVariant and GLnexus. *BioRxiv*, 2020.02.10.942086. <https://doi.org/10.1101/2020.02.10.942086>
- Zajickova, K., Vrbikova, J., Canaff, L., Pawelek, P. D., Goltzman, D., & Hendy, G. N. (2007). Identification and functional characterization of a novel mutation in the calcium-sensing receptor gene in familial hypocalciuric hypercalcemia: Modulation of clinical severity by vitamin D status. *Journal of Clinical Endocrinology and Metabolism*, *92*(7), 2616–2623. <https://doi.org/10.1210/jc.2007-0123>
- Zhao, W., Rasheed, A., Tikkanen, E., Lee, J. J., Butterworth, A. S., Howson, J. M. M., Assimes, T. L., Chowdhury, R., Orho-Melander, M., Damrauer, S., Small, A., Asma, S., Imamura, M., Yamauch, T., Chambers, J. C., Chen, P., Sapkota, B. R., Shah, N., Jabeen, S., ... Saleheen, D. (2017). Identification of new susceptibility loci for type 2 diabetes and shared etiological pathways with coronary heart disease. *Nature Genetics*, *49*(10), 1450–1457. <https://doi.org/10.1038/ng.3943>
- Zimmerman, P. A., Buckler-White, A., Alkhatib, G., Spalding, T., Kubofcik, J., Combadiere, C., Weissman, D., Cohen, O., Rubbert, A., Lam, G., Vaccarezza, M., Kennedy, P. E., Kumaraswami, V., Giorgi, J. V., Detels, R., Hunter, J., Chopek, M., Berger, E. A., Fauci, A. S., ... Murphy, P. M. (1997). Inherited resistance to HIV-1 conferred by an inactivating mutation in CC chemokine receptor 5: Studies in populations with contrasting clinical phenotypes, defined racial background, and quantified risk. *Molecular Medicine*, *3*(1), 23–36. <https://doi.org/10.1007/bf03401665>

UKB-ESC Research Team

¹**Bristol Myers Squibb**

Oleg Moiseyenko, Carlos Rios, Saurabh Saha

²**Regeneron Pharmaceuticals Inc.**

Goncalo Abecasis, Nilanjana Banerjee, Christina Beechert, Boris Boutkov, Michael Cantor, Giovanni Coppola, Aris Economides, Gisu Eom, Caitlin Forsythe, Erin D. Fuller, Zhenhua Gu, Lukas Habegger, Marcus B. Jones, Rouel Lanche, Michael Lattari, Michelle LeBlanc, Dadong Li, Luca A. Lotta, Kia Manoochchri, Adam J. Mansfield, Evan K. Maxwell, Jason Mighty, Mrunali Nafde, Sean O'Keeffe, Max Orelus, Maria Sotiropoulos Padilla, Razvan Panea, Tommy Polanco, Manasi Pradhan, Ayesha Rasool, Thomas D. Schleicher, Deepika Sharma, Alan Shuldiner, Jeffrey C. Staples, Cristopher V. Van Hout, Louis Widom, Sarah E. Wolf

³**Biogen Inc.**

Sally John, Chia-Yen Chen, David Sexton, Varant Kupelian, Eric Marshall, Timothy Swan, Susan Eaton, Jimmy Z. Liu, Stephanie Loomis, Megan Jensen, Saranya Duraisamy

⁴**Takeda Pharmaceutical Company Ltd**

Jason Tetrault, David Merberg, Sunita Badola

⁵**Abbvie Inc.**

Mark Reppell, Jason Grundstad, Xiuwen Zheng

⁶**Anylam Pharmaceuticals**

Aimee M. Deaton, Margaret M. Parker, Lucas D. Ward, Alexander O. Flynn-Carroll

⁷**AstraZeneca**

Caroline Austin (Business Development); Ruth March (Precision Medicine & Biosamples); Menelas N. Pangalos (BioPharmaceuticals R&D); Adam Platt (Translational Science & Experimental Medicine, Research and Early Development, Respiratory and Immunology); Mike Snowden (Discovery Sciences); Athena Matakidou, Sebastian Wasilewski, Quanli Wang, Sri Deevi, Keren Carss, Katherine Smith (Centre for Genomics Research, Discovery Sciences, BioPharmaceuticals R&D)

⁸**Pfizer, Inc.**

Morten Sogaard, Xinli Hu, Xing Chen, Zhan Ye

Gene Symbol	Gene Description	Generic drug name	drug type	mode of action	indication for drug	References
ABCC8	ATP binding cassette subfamily C member 8	sulphonylureas	small molecule	inhibitor	diabetes	(Mahajan et al., 2018; Nestorowicz et al., 1996; Otonkoski et al., 1999; Thomas et al., 1996)
KCNJ11	potassium voltage-gated channel subfamily J member 11	sulphonylureas	small molecule	inhibitor	diabetes	(Flannick et al., 2019; Mahajan et al., 2018; A. P. Morris et al., 2012; Xue et al., 2018)
ADRB1	beta1-adrenergic receptor	metoprolol	small molecule	inhibitor	hypertension	(Giri et al., 2019; Hoffmann et al., 2017; Sung et al., 2018; Takeuchi et al., 2018)
APOB	apolipoprotein B	mipomersen	antisense therapeutic	inhibitor	familial hypercholesterolemia	(Pullinger et al., 1995; Rabès et al., 1997; Soria et al., 1989; Van Der Graaf et al., 2011)
AVPR2	arginine vasopressin receptor 2	vasopressin/desmopressin	hormone	agonist	diabetes insipidus	(Bichet et al., 1994; Carroll et al., 2006; Pasel et al., 2000; Rosenthal et al., 1992)
CASR	calcium-sensing receptor	cinacalcet	small molecule	agonist	hypercalcemia	(Bai et al., 1996; Chou et al., 1995; Cole et al., 2009; Nissen et al., 2007; Zajickova et al., 2007)
CCR5	C-C chemokine receptor type 5	maraviroc	small molecule	inhibitor	HIV infection	(Carrington et al., 1997; Novembre et al., 2005; Quillent et al., 1998; Sullivan et al., 2001; Zimmerman et al., 1997)
CTLA4	cytotoxic T-lymphocyte associated protein 4	abatacept	protein replacement therapy	agonist	rheumatoid arthritis	(Carrington et al., 1997; Novembre et al., 2005; Quillent et al., 1998; Sullivan et al., 2001; Zimmerman et al., 1997)
DRD2	dopamine receptor D2	risperidone	small molecule	inhibitor	schizophrenia	(Li et al., 2017; Pardiñas et al., 2018; Ripke et al., 2014)
EPOR	erythropoietin receptor	Recombinant erythropoietin	hormone	agonist	anemia	(Kichayev et al., 2019; Xiong et al., 2015)
ESR1	estrogen receptor alpha	estrogens	hormone	agonist	osteoporosis	(J. A. Morris et al., 2019)
F10	factor X	rivaroxaban	small molecule	inhibitor	thrombosis	(Messier et al., 1996; Watzke et al., 1990)
F2	thrombin	dabigatran	peptide mimetic	inhibitor	thrombosis	(Hinds et al., 2016)
FSHR	follicle-stimulating hormone receptor	recombinant FSH	hormone	agonist	infertility associated with PCOS	(Aittomäki et al., 1995)
GHR	growth hormone receptor	recombinant GH	hormone	agonist	growth stature	(Ay Ling et al., 1997; Goddard et al., 1995)
HMGCR	HMG-CoA reductase	atorvastatin	small molecule	inhibitor	hypercholesterolemia	(Burkhardt et al., 2008; Kathiresan et al.,

						2008)
IL12B	interleukin-12	ustekinumab	monoclonal antibody	inhibitor	psoriasis	(Cargill et al., 2007)
IL6R	interleukin-6 receptor	tocilizumab	monoclonal antibody	inhibitor	rheumatoid arthritis	(Eyre et al., 2012; Lauffer et al., 2019; Okada et al., 2014)
INS	insulin	recombinant insulin	hormone	agonist	diabetes	(Colombo et al., 2008; Edgill et al., 2008; Støy et al., 2007)
KCNH2	hERG	amiodarone	small molecule	inhibitor	atrial fibrillation	(Sinner et al., 2008)
MTP	microsomal triglyceride transfer protein, MTP	lomitapide	small molecule	inhibitor	familial hypercholesterolemia	(Khatun et al., 2013)
NPC1L1	NPC1 like intracellular cholesterol transporter 1	ezetimibe	small molecule	inhibitor	Hypercholesterolemia	(Stitzel et al., 2014)
P2RY12	P2Y receptor 12	clopidogrel	small molecule	inhibitor	"atherosclerosis, stroke"	(Cattaneo, 2011)
PCSK9	PCSK9	alirocumab, evolocumab	monoclonal antibody	inhibitor	hypercholesterolemia	(Abifadel et al., 2003; Cohen et al., 2005)
PNLIP	pancreatic lipase	orlistat	small molecule	inhibitor	obesity	(Winkler et al., 2015)
PPARG	peroxisome proliferator activated receptor gamma	rosiglitazone	small molecule	inhibitor	diabetes	(Zhao et al., 2017)
TNFSF11	TNF superfamily member 11	denosumab	monoclonal antibody	inhibitor	osteoporosis	(Estrada et al., 2012)
SOST	sclerostin	romosozumab	monoclonal antibody	inhibitor	osteoporosis	(Estrada et al., 2012; Hamersma et al., 2003; Styrcars dottir et al., 2009, 2016)
VEGFA	vascular endothelial growth factor A	afilbercet	recombinant protein	decoy receptor	AMD; Macular Edema Following Retinal Vein Occlusion; Diabetic Macular Edema; Diabetic Retinopathy	(J. M. Lin et al., 2008; Yu et al., 2011)
VEGFA	vascular endothelial growth factor A	ranimzumab; bevacizumab; brolucizumab	monoclonal antibody	inhibitor	AMD	(J. M. Lin et al., 2008; Yu et al., 2011)
VEGFA	vascular endothelial growth factor A	pegaptanib	small molecule	inhibitor	AMD	(J. M. Lin et al., 2008; Yu et al., 2011)
PPARA	peroxisome proliferator activated receptor alp	clofibrate; gemfibrozil; ciprofibrate; bezafibrate; and fenofibrate	small molecule	agonist	hypercholesterolemia	(Willer et al., 2013)
BCL2	BCL2, apoptosis regulator	venetoclax	small molecule	inhibitor	chronic lymphocytic leukemia	(Berndt et al., 2013)
CD80	CD80 molecule	belimumab	monoclonal antibody	inhibitor	systemic lupus erythematosus	(W. Yang et al., 2013)

SMN2	survival of motor neuron 2, centromeric	nusinersen	antisense therapeutic	inhibitor	spinal muscular atrophy	(Aartsma-Rus, 2018)
LDLR	low density lipoprotein receptor	statins	small molecule		hyperlipidaemia	(Rader et al., 2003)
SRD5A2	steroid 5 alpha-reductase 2	finasteride	small molecule	inhibitor	benign prostrate hyperplasia	(Andersson et al., 1991; Imperato-McGinley et al., 1974; Wood & Rittmaster, 1994)

Table S1: Approved drugs with targets supported by human genetics evidence.

Table S1 References

- Aartsma-Rus, A. (2018). Genetic therapies for spinal muscular atrophy type 1. In *The Lancet Neurology* (Vol. 17, Issue 2, pp. 111–112). Lancet Publishing Group. [https://doi.org/10.1016/S1474-4422\(17\)30436-2](https://doi.org/10.1016/S1474-4422(17)30436-2)
- Abifadel, M., Varret, M., Rabès, J. P., Allard, D., Ouguerram, K., Devillers, M., Cruaud, C., Benjannet, S., Wickham, L., Erlich, D., Derré, A., Villéger, L., Farnier, M., Beucler, I., Bruckert, E., Chambaz, J., Chanu, B., Lecerf, J. M., Luc, G., ... Boileau, C. (2003). Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nature Genetics*, 34(2), 154–156. <https://doi.org/10.1038/ng1161>
- Abul-Husn, N. S., Cheng, X., Li, A. H., Xin, Y., Schurmann, C., Stevis, P., Liu, Y., Kozlitina, J., Stender, S., Wood, G. C., Stepanchick, A. N., Still, M. D., McCarthy, S., O’Dushlaine, C., Packer, J. S., Balasubramanian, S., Gosalia, N., Esopi, D., Kim, S. Y., ... Dewey, F. E. (2018). A Protein-Truncating *HSD17B13* Variant and Protection from Chronic Liver Disease. *New England Journal of Medicine*, 378(12), 1096–1106. <https://doi.org/10.1056/NEJMoa1712191>
- Aittomäki, K., Dieguez Lucena, J., Pakarinen, P., Sistonen, P., Tapanainen, J., Gromoll, J., Kaskikari, R., Sankila, E. M., Lehväslaiho, H., Reyes Engel, A., Nieschlag, E., Huhtaniemi, I., & de la Chapelle, A. (1995). Mutation in the follicle-stimulating hormone receptor gene causes hereditary hypergonadotropic ovarian failure. *Cell*, 82(6), 959–968. [https://doi.org/10.1016/0092-8674\(95\)90275-9](https://doi.org/10.1016/0092-8674(95)90275-9)
- Andersson, S., Berman, D. M., Jenkins, E. P., & Russell, D. W. (1991). Deletion of steroid 5 α -reductase 2 gene in male pseudohermaphroditism. *Nature*, 354(6349), 159–161. <https://doi.org/10.1038/354159a0>
- Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Bentley, D. R., Chakravarti, A., Clark, A. G., Donnelly, P., Eichler, E. E., Flicek, P., Gabriel, S. B., Gibbs, R. A., Green, E. D., Hurles, M. E., Knoppers, B. M., Korbel, J. O., Lander, E. S., Lee, C., Lehrach, H., ... Schloss, J. A. (2015). A global reference for human genetic variation. In *Nature* (Vol. 526, Issue 7571, pp. 68–74). Nature Publishing Group. <https://doi.org/10.1038/nature15393>
- Ay Ling, R. M., Ross, R., Towner, P., Von Laue, S., Finidori, J., Moutoussamy, S., Buchanan, C. R., Clayton, P. E., & Norman, M. R. (1997). A dominant-negative mutation of the growth hormone receptor causes familial short stature. *Nature Genetics*, 16(1), 13–14. <https://doi.org/10.1038/ng0597-13>
- Bai, M., Quinn, S., Trivedi, S., Kifor, O., Pearce, S. H. S., Pollak, M. R., Krapcho, K., Hebert, S. C., &

- Brown, E. M. (1996). Expression and characterization of inactivating and activating mutations in the human Ca²⁺(o)-sensing receptor. *Journal of Biological Chemistry*, 271(32), 19537–19545. <https://doi.org/10.1074/jbc.271.32.19537>
- Berndt, S. I., Skibola, C. F., Joseph, V., Camp, N. J., Nieters, A., Wang, Z., Cozen, W., Monnereau, A., Wang, S. S., Kelly, R. S., Lan, Q., Teras, L. R., Chatterjee, N., Chung, C. C., Yeager, M., Brooks-Wilson, A. R., Hartge, P., Purdue, M. P., Birmann, B. M., ... Slager, S. L. (2013). Genome-wide association study identifies multiple risk loci for chronic lymphocytic leukemia. *Nature Genetics*, 45(8), 868–876. <https://doi.org/10.1038/ng.2652>
- Bichet, D. G., Birnbaumer, M., Lonergan, M., Arthus, M. F., Rosenthal, W., Goodyer, P., Nivet, H., Benoit, S., Giampietro, P., Simonetti, S., Fish, A., Whitley, C. B., Jaeger, P., Gertner, J., New, M., DiBona, F. J., Kaplan, B. S., Robertson, G. L., Hendy, G. N., ... Morgan, K. (1994). Nature and recurrence of AVPR2 mutations in X-linked nephrogenic diabetes insipidus. *American Journal of Human Genetics*, 55(2), 278–286.
- Burkhardt, R., Kenny, E. E., Lowe, J. K., Birkeland, A., Josowitz, R., Noel, M., Salit, J., Maller, J. B., Pe'er, I., Daly, M. J., Altshuler, D., Stoffel, M., Friedman, J. M., & Breslow, J. L. (2008). Common SNPs in HMGCR in micronesians and whites associated with LDL-cholesterol levels affect alternative splicing of exon13. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 28(11), 2078–2084. <https://doi.org/10.1161/ATVBAHA.108.172288>
- Burton, P. R., Clayton, D. G., Cardon, L. R., Craddock, N., Deloukas, P., Duncanson, A., Kwiatkowski, D. P., McCarthy, M. I., Ouwehand, W. H., Samani, N. J., Todd, J. A., Donnelly, P., Barrett, J. C., Davison, D., Easton, D., Evans, D., Leung, H. T., Marchini, J. L., Morris, A. P., ... Compston, A. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, 447(7145), 661–678. <https://doi.org/10.1038/nature05911>
- Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., Cortes, A., Welsh, S., Young, A., Effingham, M., McVean, G., Leslie, S., Allen, N., Donnelly, P., & Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature*, 562(7726), 203–209. <https://doi.org/10.1038/s41586-018-0579-z>
- Cargill, M., Schrodi, S. J., Chang, M., Garcia, V. E., Brandon, R., Callis, K. P., Matsunami, N., Ardlie, K. G., Civello, D., Catanese, J. J., Leong, D. U., Panko, J. M., McAllister, L. B., Hansen, C. B., Papenfuss, J., Prescott, S. M., White, T. J., Leppert, M. F., Krueger, G. G., & Begovich, A. B. (2007). A large-scale genetic association study confirms IL12B and leads to the identification of IL23R as psoriasis-risk genes. *American Journal of Human Genetics*, 80(2), 273–290. <https://doi.org/10.1086/511051>
- Carrington, M., Kissner, T., Gerrard, B., Ivanov, S., O'Brien, S. J., & Dean, M. (1997). Novel alleles of the chemokine-receptor gene CCR5. *American Journal of Human Genetics*, 61(6), 1261–1267. <https://doi.org/10.1086/301645>
- Carroll, P., Al-Mojalli, H., Al-Abbad, A., Al-Hassoun, I., Al-Hamed, M., Al-Amr, R., Butt, A. I., & Meyer, B. F. (2006). Novel mutations underlying nephrogenic diabetes insipidus in Arab families. *Genetics in Medicine*, 8(7), 443–447. <https://doi.org/10.1097/01.gim.0000223554.46981.7a>
- Cattaneo, M. (2011). The platelet P2Y₁₂ receptor for adenosine diphosphate: Congenital and drug-induced defects. In *Blood* (Vol. 117, Issue 7, pp. 2102–2112). Blood.

<https://doi.org/10.1182/blood-2010-08-263111>

- Chou, Y. H. W., Pollak, M. R., Brandi, M. L., Toss, G., Arnqvist, H., Atkinson, A. B., Papapoulos, S. E., Marx, S., Brown, E. M., Seidman, J. G., & Seidman, C. E. (1995). Mutations in the human Ca²⁺-sensing-receptor gene that cause familial hypocalciuric hypercalcemia. *American Journal of Human Genetics*, *56*(5), 1075–1079.
- Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*, *6*(2). <https://doi.org/10.4161/fly.19695>
- Cohen, J., Pertsemlidis, A., Kotowski, I. K., Graham, R., Garcia, C. K., & Hobbs, H. H. (2005). Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nature Genetics*, *37*(2), 161–165. <https://doi.org/10.1038/ng1509>
- Cole, D. E. C., Yun, F. H. J., Wong, B. Y. L., Shuen, A. Y., Booth, R. A., Scillitani, A., Pidasheva, S., Zhou, X., Canaff, L., & Hendy, G. N. (2009). Calcium-sensing receptor mutations and denaturing high performance liquid chromatography. *Journal of Molecular Endocrinology*, *42*(4), 331–339. <https://doi.org/10.1677/JME-08-0164>
- Colombo, C., Porzio, O., Liu, M., Massa, O., Vasta, M., Salardi, S., Beccaria, L., Monciotti, C., Toni, S., Pedersen, O., Hansen, T., Federici, L., Pesavento, R., Cadario, F., Federici, G., Ghirri, P., Arvan, P., Iafusco, D., & Barbetti, F. (2008). Seven mutations in the human insulin gene linked to permanent neonatal/infancy-onset diabetes mellitus. *Journal of Clinical Investigation*, *118*(6), 2148–2156. <https://doi.org/10.1172/JCI33777>
- Cook, D., Brown, D., Alexander, R., March, R., Morgan, P., Satterthwaite, G., & Pangalos, M. N. (2014). Lessons learned from the fate of AstraZeneca's drug pipeline: A five-dimensional framework. In *Nature Reviews Drug Discovery* (Vol. 13, Issue 6, pp. 419–431). Nature Publishing Group. <https://doi.org/10.1038/nrd4309>
- Dolgin, E. (2019). Massive NIH–industry project opens portals to target validation. *Nature Reviews Drug Discovery*. <https://doi.org/10.1038/d41573-019-00033-8>
- Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., Epstein, C. B., Fretz, S., Harrow, J., Kaul, R., Khatun, J., Lajoie, B. R., Landt, S. G., Lee, B. K., Pauli, F., Rosenbloom, K. R., Sabo, P., Safi, A., Sanyal, A., ... Lochovsky, L. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, *489*(7414), 57–74. <https://doi.org/10.1038/nature11247>
- Edghill, E. L., Flanagan, S. E., Patch, A. M., Boustred, C., Parrish, A., Shields, B., Shepherd, M. H., Hussain, K., Kapoor, R. R., Malecki, M., MacDonald, M. J., Støy, J., Steiner, D. F., Philipson, L. H., Bell, G. I., Hattersley, A. T., & Ellard, S. (2008). Insulin mutation screening in 1,044 patients with diabetes mutations in the INS gene are a common cause of neonatal diabetes but a rare cause of diabetes diagnosed in childhood or adulthood. *Diabetes*, *57*(4), 1034–1042. <https://doi.org/10.2337/db07-1405>
- Estrada, K., Styrkarsdottir, U., Evangelou, E., Hsu, Y. H., Duncan, E. L., Ntzani, E. E., Oei, L., Albagha, O. M. E., Amin, N., Kemp, J. P., Koller, D. L., Li, G., Liu, C. T., Minster, R. L., Moayyeri, A., Vandenput, L., Willner, D., Xiao, S. M., Yerges-Armstrong, L. M., ... Rivadeneira, F. (2012). Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nature Genetics*, *44*(5), 491–501. <https://doi.org/10.1038/ng.2249>

- Eyre, S., Bowes, J., Diogo, D., Lee, A., Barton, A., Martin, P., Zhernakova, A., Stahl, E., Viatte, S., McAllister, K., Amos, C. I., Padyukov, L., Toes, R. E. M., Huizinga, T. W. J., Wijmenga, C., Trynka, G., Franke, L., Westra, H. J., Alfredsson, L., ... Bentley, D. (2012). High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. *Nature Genetics*, *44*(12), 1336–1340. <https://doi.org/10.1038/ng.2462>
- Flannick, J., Mercader, J. M., Fuchsberger, C., Udler, M. S., Mahajan, A., Wessel, J., Teslovich, T. M., Caulkins, L., Koesterer, R., Barajas-Olmos, F., Blackwell, T. W., Boerwinkle, E., Brody, J. A., Centeno-Cruz, F., Chen, L., Chen, S., Contreras-Cubas, C., Córdova, E., Correa, A., ... Boehnke, M. (2019). Exome sequencing of 20,791 cases of type 2 diabetes and 24,440 controls. *Nature*, *570*(7759), 71–76. <https://doi.org/10.1038/s41586-019-1231-2>
- Furtado, R. H. M., & Giugliano, R. P. (2020). What Lessons Have We Learned and What Remains to be Clarified for PCSK9 Inhibitors? A Review of FOURIER and ODYSSEY Outcomes Trials. *Cardiology and Therapy*, 1–15. <https://doi.org/10.1007/s40119-020-00163-w>
- Giri, A., Hellwege, J. N., Keaton, J. M., Park, J., Qiu, C., Warren, H. R., Torstenson, E. S., Kovesdy, C. P., Sun, Y. V., Wilson, O. D., Robinson-Cohen, C., Roumie, C. L., Chung, C. P., Birdwell, K. A., Damrauer, S. M., DuVall, S. L., Klarin, D., Cho, K., Wang, Y., ... Edwards, T. L. (2019). Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. *Nature Genetics*, *51*(1), 51–62. <https://doi.org/10.1038/s41588-018-0303-9>
- Goddard, A. D., Covello, R., Luoh, S. M., Clackson, T., Attie, K. M., Gesundheit, N., Rundle, A. C., Wells, J. A., & Carlsson, L. M. S. (1995). Mutations of the growth hormone receptor in children with idiopathic short stature. *New England Journal of Medicine*, *333*(17), 1093–1098. <https://doi.org/10.1056/NEJM199510263331701>
- Hamersma, H., Gardner, J., & Beighton, P. (2003). The natural history of sclerosteosis. *Clinical Genetics*, *63*(3), 192–197. <https://doi.org/10.1034/j.1399-0004.2003.00036.x>
- Hinds, D. A., Buil, A., Ziemek, D., Martinez-Perez, A., Malik, R., Folkersen, L., Germain, M., Mälarstig, A., Brown, A., Soria, J. M., Dichgans, M., Bing, N., Franco-Cereceda, A., Souto, J. C., Dermitzakis, E. T., Hamsten, A., Worrall, B. B., Tung, J. Y., & Sabater-Lleal, M. (2016). Genome-wide association analysis of self-reported events in 6135 individuals and 252 827 controls identifies 8 loci associated with thrombosis. *Human Molecular Genetics*, *25*(9), 1867–1874. <https://doi.org/10.1093/hmg/ddw037>
- Hoffmann, T. J., Ehret, G. B., Nandakumar, P., Ranatunga, D., Schaefer, C., Kwok, P. Y., Iribarren, C., Chakravarti, A., & Risch, N. (2017). Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. *Nature Genetics*, *49*(1), 54–64. <https://doi.org/10.1038/ng.3715>
- Imperato-McGinley, J., Guerrero, L., Gautier, T., & Peterson, R. E. (1974). Steroid 5 α -reductase deficiency in man: An inherited form of male pseudohermaphroditism. *Science*, *186*(4170), 1213–1215. <https://doi.org/10.1126/science.186.4170.1213>
- Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., Gauthier, L. D., Brand, H., Solomonson, M., Watts, N. A., Rhodes, D., Singer-Berk, M., England, E. M., Seaby, E. G., Kosmicki, J. A., ... MacArthur, D. G. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, *581*(7809), 434–443. <https://doi.org/10.1038/s41586-020-2308-7>
- Kathiresan, S., Melander, O., Guiducci, C., Surti, A., Burt, N. P., Rieder, M. J., Cooper, G. M.,

- Roos, C., Voight, B. F., Havulinna, A. S., Wahlstrand, B., Hedner, T., Corella, D., Tai, E. S., Ordovas, J. M., Berglund, G., Vartiainen, E., Jousilahti, P., Hedblad, B., ... Orho-Melander, M. (2008). Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nature Genetics*, *40*(2), 189–197. <https://doi.org/10.1038/ng.75>
- Khatun, I., Walsh, M. T., & Hussain, M. M. (2013). Loss of both phospholipid and triglyceride transfer activities of microsomal triglyceride transfer protein in abetalipoproteinemia. *Journal of Lipid Research*, *54*(6), 1541–1549. <https://doi.org/10.1194/jlr.M031658>
- Kichaev, G., Bhatia, G., Loh, P. R., Gazal, S., Burch, K., Freund, M. K., Schoech, A., Pasaniuc, B., & Price, A. L. (2019). Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *American Journal of Human Genetics*, *104*(1), 65–75. <https://doi.org/10.1016/j.ajhg.2018.11.008>
- King, E. A., Wade Davis, J., & Degner, J. F. (2019). Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval. *PLoS Genetics*, *15*(12), e1008489. <https://doi.org/10.1371/journal.pgen.1008489>
- Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., Fitzhugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., Levine, R., McEwan, P., ... Morgan, M. J. (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*(6822), 860–921. <https://doi.org/10.1038/35057062>
- Laufer, V. A., Tiwari, H. K., Reynolds, R. J., Danila, M. I., Wang, J., Edberg, J. C., Kimberly, R. P., Kottyan, L. C., Harley, J. B., Mikuls, T. R., Gregersen, P. K., Absher, D. M., Langefeld, C. D., Arnett, D. K., & Bridges, S. L. (2019). Genetic influences on susceptibility to rheumatoid arthritis in African-Americans. *Human Molecular Genetics*, *28*(5), 858–874. <https://doi.org/10.1093/hmg/ddy395>
- Li, Z., Chen, J., Yu, H., He, L., Xu, Y., Zhang, D., Yi, Q., Li, C., Li, X., Shen, J., Song, Z., Ji, W., Wang, M., Zhou, J., Chen, B., Liu, Y., Wang, J., Wang, P., Yang, P., ... Shi, Y. (2017). Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. *Nature Genetics*, *49*(11), 1576–1583. <https://doi.org/10.1038/ng.3973>
- Lin, J. M., Wan, L., Tsai, Y. Y., Lin, H. J., Tsai, Y., Lee, C. C., Tsai, C. H., Tseng, S. H., & Tsai, F. J. (2008). Vascular Endothelial Growth Factor Gene Polymorphisms in Age-related Macular Degeneration. *American Journal of Ophthalmology*, *145*(6). <https://doi.org/10.1016/j.ajo.2008.01.027>
- Lin, M., Rodeh, O., Penn, J., Bai, X., Reid, J., Krashenina, O., & Salerno, W. (2018). GLnexus: joint variant calling for large cohort sequencing. *BioRxiv*, 343970. <https://doi.org/10.1101/343970>
- Liu, J. Z., Chen, C.-Y., Tsai, E. A., Whelan, C. D., Sexton, D., John, S., & Runz, H. (2020). The burden of rare protein-truncating genetic variants on human lifespan. *BioRxiv*, 2020.06.02.129908. <https://doi.org/10.1101/2020.06.02.129908>
- Mahajan, A., Wessel, J., Willems, S. M., Zhao, W., Robertson, N. R., Chu, A. Y., Gan, W., Kitajima, H., Taliun, D., Rayner, N. W., Guo, X., Lu, Y., Li, M., Jensen, R. A., Hu, Y., Huo, S., Lohman, K. K., Zhang, W., Cook, J. P., ... McCarthy, M. I. (2018). Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes article. *Nature*

- Genetics*, 50(4), 559–571. <https://doi.org/10.1038/s41588-018-0084-1>
- McCarthy, M. I., Abecasis, G. R., Cardon, L. R., Goldstein, D. B., Little, J., Ioannidis, J. P. A., & Hirschhorn, J. N. (2008). Genome-wide association studies for complex traits: Consensus, uncertainty and challenges. *Nature Reviews Genetics*, 9(5), 356–369. <https://doi.org/10.1038/nrg2344>
- McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A. R., Teumer, A., Kang, H. M., Fuchsberger, C., Danecek, P., Sharp, K., Luo, Y., Sidore, C., Kwong, A., Timpson, N., Koskinen, S., Vrieze, S., Scott, L. J., Zhang, H., Mahajan, A., ... Marchini, J. (2016). A reference panel of 64,976 haplotypes for genotype imputation. *Nature Genetics*, 48(10), 1279–1283. <https://doi.org/10.1038/ng.3643>
- Messier, T. L., Wong, C. Y., Bovill, E. G., Long, G. L., & Church, W. R. (1996). Factor X Stockton: a mild bleeding diathesis associated with an active site mutation in factor X. *Blood Coagulation & Fibrinolysis: An International Journal in Haemostasis and Thrombosis*, 7(1), 5–14.
- Methé, B. A., Nelson, K. E., Pop, M., Creasy, H. H., Giglio, M. G., Huttenhower, C., Gevers, D., Petrosino, J. F., Abubucker, S., Badger, J. H., Chinwalla, A. T., Earl, A. M., Fitzgerald, M. G., Fulton, R. S., Hallsworth-Pepin, K., Lobos, E. A., Madupu, R., Magrini, V., Martin, J. C., ... White, O. (2012). A framework for human microbiome research. *Nature*, 486(7402), 215–221. <https://doi.org/10.1038/nature11209>
- Morris, A. P., Voight, B. F., Teslovich, T. M., Ferreira, T., Segre, A. V., Steinthorsdottir, V., Strawbridge, R. J., Khan, H., Grallert, H., Mahajan, A., Prokopenko, I., Kang, H. M., Dina, C., Esko, T., Fraser, R. M., Kanoni, S., Kumar, A., Lagou, V., Langenberg, C., ... McCarthy, M. I. (2012). Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nature Genetics*, 44(9), 981–990. <https://doi.org/10.1038/ng.2383>
- Morris, J. A., Kemp, J. P., Youlten, S. E., Laurent, L., Logan, J. G., Chai, R. C., Vulpescu, N. A., Forgetta, V., Kleinman, A., Mohanty, S. T., Sergio, C. M., Quinn, J., Nguyen-Yamamoto, L., Luco, A. L., Vijay, J., Simon, M. M., Pramatarova, A., Medina-Gomez, C., Trajanoska, K., ... Richards, J. B. (2019). An atlas of genetic influences on osteoporosis in humans and mice. *Nature Genetics*, 51(2), 258–266. <https://doi.org/10.1038/s41588-018-0302-x>
- Nelson, M. R., Tipney, H., Painter, J. L., Shen, J., Nicoletti, P., Shen, Y., Floratos, A., Sham, P. C., Li, M. J., Wang, J., Cardon, L. R., Whittaker, J. C., & Sanseau, P. (2015). The support of human genetic evidence for approved drug indications. *Nature Genetics*, 47(8), 856–860. <https://doi.org/10.1038/ng.3314>
- Nestorowicz, A., Wilson, B. A., Schoor, K. P., Inoue, H., Glaser, B., Landau, H., Stanley, C. A., Thornton, P. S., Clement IV, J. P., Bryan, J., Aguilar-Bryan, L., & Permutt, M. A. (1996). Mutations in the sulfonyleurea receptor gene are associated with familial hyperinsulinism in Ashkenazi Jews. *Human Molecular Genetics*, 5(11), 1813–1822. <https://doi.org/10.1093/hmg/5.11.1813>
- Nguyen, P. A., Born, D. A., Deaton, A. M., Nioi, P., & Ward, L. D. (2019). Phenotypes associated with genes encoding drug targets are predictive of clinical trial side effects. *Nature Communications*, 10(1). <https://doi.org/10.1038/s41467-019-09407-3>
- Nioi, P., Sigurdsson, A., Thorleifsson, G., Helgason, H., Agustsdottir, A. B., Norddahl, G. L., Helgadottir, A., Magnusdottir, A., Jonasdottir, A., Gretarsdottir, S., Jonsdottir, I.,

- Steinthorsdottir, V., Rafnar, T., Swinkels, D. W., Galesloot, T. E., Grarup, N., Jørgensen, T., Vestergaard, H., Hansen, T., ... Stefansson, K. (2016). Variant *ASGR1* Associated with a Reduced Risk of Coronary Artery Disease. *New England Journal of Medicine*, *374*(22), 2131–2141. <https://doi.org/10.1056/NEJMoa1508419>
- Nissen, P. H., Christensen, S. E., Heickendorff, L., Brixen, K., & Mosekilde, L. (2007). Molecular genetic analysis of the calcium sensing receptor gene in patients clinically suspected to have familial hypocalciuric hypercalcemia: Phenotypic variation and mutation spectrum in a Danish population. *Journal of Clinical Endocrinology and Metabolism*, *92*(11), 4373–4379. <https://doi.org/10.1210/jc.2007-0322>
- Novembre, J., Galvani, A. P., & Slatkin, M. (2005). The geographic spread of the CCR5 Δ 32 HIV-resistance allele. *PLoS Biology*, *3*(11), 1954–1962. <https://doi.org/10.1371/journal.pbio.0030339>
- Okada, Y., Wu, D., Trynka, G., Raj, T., Terao, C., Ikari, K., Kochi, Y., Ohmura, K., Suzuki, A., Yoshida, S., Graham, R. R., Manoharan, A., Ortmann, W., Bhangale, T., Denny, J. C., Carroll, R. J., Eyler, A. E., Greenberg, J. D., Kremer, J. M., ... Plenge, R. M. (2014). Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature*, *506*(7488), 376–381. <https://doi.org/10.1038/nature12873>
- Otonkoski, T., Ämmälä, C., Huopio, H., Cote, G. J., Chapman, J., Cosgrove, K., Ashfield, R., Huang, E., Komulainen, J., Ashcroft, F. M., Dunne, M. J., Kere, J., & Thomas, P. M. (1999). A point mutation inactivating the sulfonylurea receptor causes the severe form of persistent hyperinsulinemic hypoglycemia of infancy in Finland. *Diabetes*, *48*(2), 408–415. <https://doi.org/10.2337/diabetes.48.2.408>
- Pardiñas, A. F., Holmans, P., Pocklington, A. J., Escott-Price, V., Ripke, S., Carrera, N., Legge, S. E., Bishop, S., Cameron, D., Hamshere, M. L., Han, J., Hubbard, L., Lynham, A., Mantripragada, K., Rees, E., MacCabe, J. H., McCarroll, S. A., Baune, B. T., Breen, G., ... Walters, J. T. R. (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature Genetics*, *50*(3), 381–389. <https://doi.org/10.1038/s41588-018-0059-2>
- Pasel, K., Schulz, A., Timmermann, K., Linnemann, K., Hoeltzenbein, M., Jääskeläinen, J., Grüters, A., Filler, G., & Schöneberg, T. (2000). Functional Characterization of the Molecular Defects Causing Nephrogenic Diabetes Insipidus in Eight Families 1. *The Journal of Clinical Endocrinology & Metabolism*, *85*(4), 1703–1710. <https://doi.org/10.1210/jcem.85.4.6507>
- Poplin, R., Chang, P. C., Alexander, D., Schwartz, S., Colthurst, T., Ku, A., Newburger, D., Dijamco, J., Nguyen, N., Afshar, P. T., Gross, S. S., Dorfman, L., McLean, C. Y., & DePristo, M. A. (2018). A universal snp and small-indel variant caller using deep neural networks. *Nature Biotechnology*, *36*(10), 983. <https://doi.org/10.1038/nbt.4235>
- Pullinger, C. R., Hennessy, L. K., Chatterton, J. E., Liu, W., Love, J. A., Mendel, C. M., Frost, P. H., Malloy, M. J., Schumaker, V. N., & Kane, J. P. (1995). Familial Ligand-Defective Apolipoprotein B: Identification of a New Mutation That Decreases LDL Receptor Binding Affinity. *Journal of Clinical Investigation*, *95*(3), 1225–1234. <https://doi.org/10.1172/JCI117772>
- Quillent, C., Oberlin, E., Braun, J., Rousset, D., Gonzalez-Canali, G., Métais, P., Montagnier, L., Virelizier, J. L., Arenzana-Seisdedos, F., & Beretta, A. (1998). HIV-1-resistance phenotype

- conferred by combination of two separate inherited mutations of CCR5 gene. *Lancet*, 351(9095), 14–18. [https://doi.org/10.1016/S0140-6736\(97\)09185-X](https://doi.org/10.1016/S0140-6736(97)09185-X)
- Rabès, J. P., Varret, M., Saint-Jore, B., Erlich, D., Jondeau, G., Krempf, M., Giraudet, P., Junien, C., & Boileau, C. (1997). Familial ligand-defective apolipoprotein b-100: Simultaneous detection of the arg3500 → GLN and ARG3531→CYS mutations in a French population. *Human Mutation*, 10(2), 160–163. [https://doi.org/10.1002/\(SICI\)1098-1004\(1997\)10:2<160::AID-HUMU8>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1098-1004(1997)10:2<160::AID-HUMU8>3.0.CO;2-O)
- Rader, D. J., Cohen, J., & Hobbs, H. H. (2003). Monogenic hypercholesterolemia: new insights in pathogenesis and treatment. *Journal of Clinical Investigation*, 111(12), 1795–1803. <https://doi.org/10.1172/jci18925>
- Regier, A. A., Farjoun, Y., Larson, D. E., Krasheninina, O., Kang, H. M., Howrigan, D. P., Chen, B. J., Kher, M., Banks, E., Ames, D. C., English, A. C., Li, H., Xing, J., Zhang, Y., Matise, T., Abecasis, G. R., Salerno, W., Zody, M. C., Neale, B. M., & Hall, I. M. (2018). Functional equivalence of genome sequencing analysis pipelines enables harmonized variant calling across human genetics projects. *Nature Communications*, 9(1), 4038. <https://doi.org/10.1038/s41467-018-06159-4>
- Ripke, S., Neale, B. M., Corvin, A., Walters, J. T. R., Farh, K. H., Holmans, P. A., Lee, P., Bulik-Sullivan, B., Collier, D. A., Huang, H., Pers, T. H., Agartz, I., Agerbo, E., Albus, M., Alexander, M., Amin, F., Bacanu, S. A., Begemann, M., Belliveau, R. A., ... O'Donovan, M. C. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, 511(7510), 421–427. <https://doi.org/10.1038/nature13595>
- Rodriguez, J. M., Rodriguez-Rivas, J., Di Domenico, T., Vázquez, J., Valencia, A., & Tress, M. L. (2018). APPRIS 2017: Principal isoforms for multiple gene sets. *Nucleic Acids Research*, 46(D1). <https://doi.org/10.1093/nar/gkx997>
- Rosenthal, W., Seibold, A., Antaramian, A., Lonergan, M., Arthus, M. F., Hendy, G. N., Birnbaumer, M., & Bichet, D. G. (1992). Molecular identification of the gene responsible for congenital nephrogenic diabetes insipidus. *Nature*, 359(6392), 233–235. <https://doi.org/10.1038/359233a0>
- Sabatine, M. S., Giugliano, R. P., Keech, A. C., Honarpour, N., Wiviott, S. D., Murphy, S. A., Kuder, J. F., Wang, H., Liu, T., Wasserman, S. M., Sever, P. S., Pedersen, T. R., Fish, M. P., Abrahamsen, T. E., Im, K., Kanevsky, E., Bonaca, M. P., Lira Pineda, A., Hanlon, K., ... Shifrin, G. (2017). Evolocumab and clinical outcomes in patients with cardiovascular disease. *New England Journal of Medicine*, 376(18), 1713–1722. <https://doi.org/10.1056/NEJMoa1615664>
- Schwartz, G. G., Steg, P. G., Szarek, M., Bhatt, D. L., Bittner, V. A., Diaz, R., Edelberg, J. M., Goodman, S. G., Hanotin, C., Harrington, R. A., Jukema, J. W., Lecorps, G., Mahaffey, K. W., Moryusef, A., Pordy, R., Quintero, K., Roe, M. T., Sasiela, W. J., Tamby, J. F., ... Ball, E. (2018). Alirocumab and cardiovascular outcomes after acute coronary syndrome. *New England Journal of Medicine*, 379(22), 2097–2107. <https://doi.org/10.1056/NEJMoa1801174>
- Sinner, M. F., Pfeufer, A., Akyol, M., Beckmann, B. M., Hinterseer, M., Wacker, A., Perz, S., Sauter, W., Illig, T., Näbauer, M., Schmitt, C., Wichmann, H. E., Schömig, A., Steinbeck, G., Meitinger, T., & Käb, S. (2008). The non-synonymous coding IKr-channel variant KCNH2-K897T is associated with atrial fibrillation: Results from a systematic candidate gene-based

- analysis of KCNH2 (HERG). *European Heart Journal*, 29(7), 907–914.
<https://doi.org/10.1093/eurheartj/ehm619>
- Soria, L. F., Ludwig, E. H., Clarke, H. R. G., Vega, G. L., Grundy, S. M., & McCarthy, B. J. (1989). Association between a specific apolipoprotein B mutation and familial defective apolipoprotein B-100. *Proceedings of the National Academy of Sciences of the United States of America*, 86(2), 587–591. <https://doi.org/10.1073/pnas.86.2.587>
- Stitzel, N. O., Won, H. H., Morrison, A. C., Peloso, G. M., Do, R., Lange, L. A., Fontanillas, P., Gupta, N., Duga, S., Goel, A., Farrall, M., Saleheen, D., Ferrario, P., König, I., Asselta, R., Merlini, P. A., Marziliano, N., Notarangelo, M. F., Schick, U., ... Kathiresan, S. (2014). Inactivating mutations in NPC1L1 and protection from coronary heart disease. *New England Journal of Medicine*, 371(22), 2072–2082.
<https://doi.org/10.1056/NEJMoa1405386>
- Støy, J., Edghill, E. L., Flanagan, S. E., Ye, H., Paz, V. P., Pluzhnikov, A., Below, J. E., Hayes, M. G., Cox, N. J., Lipkind, G. M., Lipton, R. B., Greeley, S. A. W., Patch, A. M., Ellard, S., Steiner, D. F., Hattersley, A. T., Philipson, L. H., & Bell, G. I. (2007). Insulin gene mutations as a cause of permanent neonatal diabetes. *Proceedings of the National Academy of Sciences of the United States of America*, 104(38), 15040–15044.
<https://doi.org/10.1073/pnas.0707291104>
- Styrkarsdottir, U., Halldorsson, B. V., Gretarsdottir, S., Gudbjartsson, D. F., Walters, G. B., Ingvarsson, T., Jonsdottir, T., Saemundsdottir, J., Snorraddottir, S., Center, J. R., Nguyen, T. V., Alexandersen, P., Gulcher, J. R., Eisman, J. A., Christiansen, C., Sigurdsson, G., Kong, A., Thorsteinsdottir, U., & Stefansson, K. (2009). New sequence variants associated with bone mineral density. *Nature Genetics*, 41(1), 15–17. <https://doi.org/10.1038/ng.284>
- Styrkarsdottir, U., Thorleifsson, G., Gudjonsson, S. A., Sigurdsson, A., Center, J. R., Lee, S. H., Nguyen, T. V., Kwok, T. C. Y., Lee, J. S. W., Ho, S. C., Woo, J., Leung, P. C., Kim, B. J., Rafnar, T., Kiemenev, L. A., Ingvarsson, T., Koh, J. M., Tang, N. L. S., Eisman, J. A., ... Stefansson, K. (2016). Sequence variants in the PTCH1 gene associate with spine bone mineral density and osteoporotic fractures. *Nature Communications*, 7.
<https://doi.org/10.1038/ncomms10129>
- Sullivan, A. D., Wigginton, J., & Kirschner, D. (2001). The coreceptor mutation CCR5Δ32 influences the dynamics of hiv epidemics and is selected for by HIV. *Proceedings of the National Academy of Sciences of the United States of America*, 98(18), 10214–10219.
<https://doi.org/10.1073/pnas.181325198>
- Sung, Y. J., Winkler, T. W., de las Fuentes, L., Bentley, A. R., Brown, M. R., Kraja, A. T., Schwander, K., Ntalla, I., Guo, X., Franceschini, N., Lu, Y., Cheng, C. Y., Sim, X., Vojinovic, D., Marten, J., Musani, S. K., Li, C., Feitosa, M. F., Kilpeläinen, T. O., ... Chasman, D. I. (2018). A Large-Scale Multi-ancestry Genome-wide Study Accounting for Smoking Behavior Identifies Multiple Significant Loci for Blood Pressure. *American Journal of Human Genetics*, 102(3), 375–400. <https://doi.org/10.1016/j.ajhg.2018.01.015>
- Takeuchi, F., Akiyama, M., Matoba, N., Katsuya, T., Nakatochi, M., Tabara, Y., Narita, A., Saw, W. Y., Moon, S., Spracklen, C. N., Chai, J. F., Kim, Y. J., Zhang, L., Wang, C., Li, H., Li, H., Wu, J. Y., Dorajoo, R., Nierenberg, J. L., ... Kato, N. (2018). Interethnic analyses of blood pressure loci in populations of East Asian and European descent. *Nature Communications*, 9(1).
<https://doi.org/10.1038/s41467-018-07345-0>

The Cancer Genome Atlas Program - National Cancer Institute. (n.d.).

- Thomas, P. M., Wohllk, N., Huang, E., Kuhnle, U., Rabl, W., Gagel, R. F., & Cote, G. J. (1996). Inactivation of the first nucleotide-binding fold of the sulfonyleurea receptor, and familial persistent hyperinsulinemic hypoglycemia of infancy. *American Journal of Human Genetics*, *59*(3), 510–518.
- Van Der Graaf, A., Avis, H. J., Kusters, D. M., Vissers, M. N., Hutten, B. A., Defesche, J. C., Huijgen, R., Fouchier, S. W., Wijburg, F. A., Kastelein, J. J. P., & Wiegman, A. (2011). Molecular basis of autosomal dominant hypercholesterolemia: Assessment in a large cohort of hypercholesterolemic children. *Circulation*, *123*(11), 1167–1173. <https://doi.org/10.1161/CIRCULATIONAHA.110.979450>
- Van Hout, C. V., Tachmazidou, I., Backman, J. D., Hoffman, J. D., Liu, D., Pandey, A. K., Gonzaga-Jauregui, C., Khalid, S., Ye, B., Banerjee, N., Li, A. H., O’Dushlaine, C., Marcketta, A., Staples, J., Schurmann, C., Hawes, A., Maxwell, E., Barnard, L., Lopez, A., ... Baras, A. (2020). Exome sequencing and characterization of 49,960 individuals in the UK Biobank. *Nature*. <https://doi.org/10.1038/s41586-020-2853-0>
- Watzke, H. H., Lechner, K., Roberts, H. R., Reddy, S. V., Welsch, D. J., Friedman, P., Mahr, G., Jagadeeswaran, P., Monroe, D. M., & High, K. A. (1990). Molecular defect (Gla+14 → Lys) and its functional consequences in a hereditary factor X deficiency (factor X ‘Vorarlberg’). *Journal of Biological Chemistry*, *265*(20), 11982–11989.
- Willer, C. J., Schmidt, E. M., Sengupta, S., Peloso, G. M., Gustafsson, S., Kanoni, S., Ganna, A., Chen, J., Buchkovich, M. L., Mora, S., Beckmann, J. S., Bragg-Gresham, J. L., Chang, H. Y., Demirkan, A., Den Hertog, H. M., Do, R., Donnelly, L. A., Ehret, G. B., Esko, T., ... Abecasis, G. R. (2013). Discovery and refinement of loci associated with lipid levels. *Nature Genetics*, *45*(11), 1274–1285. <https://doi.org/10.1038/ng.2797>
- Winkler, T. W., Justice, A. E., Graff, M., Barata, L., Feitosa, M. F., Chu, S., Czajkowski, J., Esko, T., Fall, T., Kilpeläinen, T. O., Lu, Y., Mägi, R., Mihailov, E., Pers, T. H., Rieger, S., Teumer, A., Ehret, G. B., Ferreira, T., Heard-Costa, N. L., ... Loos, R. J. F. (2015). The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. *PLoS Genetics*, *11*(10), 1–42. <https://doi.org/10.1371/journal.pgen.1005378>
- Wood, A. J. J., & Rittmaster, R. S. (1994). Finasteride. In *New England Journal of Medicine* (Vol. 330, Issue 2, pp. 120–125). <https://doi.org/10.1056/NEJM199401133300208>
- Xiong, H. Y., Alipanahi, B., Lee, L. J., Bretschneider, H., Merico, D., Yuen, R. K. C., Hua, Y., Gueroussov, S., Najafabadi, H. S., Hughes, T. R., Morris, Q., Barash, Y., Krainer, A. R., Jovic, N., Scherer, S. W., Blencowe, B. J., & Frey, B. J. (2015). The human splicing code reveals new insights into the genetic determinants of disease. *Science*, *347*(6218). <https://doi.org/10.1126/science.1254806>
- Xue, A., Wu, Y., Zhu, Z., Zhang, F., Kemper, K. E., Zheng, Z., Yengo, L., Lloyd-Jones, L. R., Sidorenko, J., Wu, Y., Agbessi, M., Ahsan, H., Alves, I., Andiappan, A., Awadalla, P., Battle, A., Beutner, F., Bonder, M. J. J., Boomsma, D., ... Yang, J. (2018). Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. *Nature Communications*, *9*(1). <https://doi.org/10.1038/s41467-018-04951-w>
- Yang, W., Tang, H., Zhang, Y., Tang, X., Zhang, J., Sun, L., Yang, J., Cui, Y., Zhang, L., Hirankarn, N., Cheng, H., Pan, H. F., Gao, J., Lee, T. L., Sheng, Y., Lau, C. S., Li, Y., Chan, T. M., Yin, X., ...

- Lau, Y. L. (2013). Meta-analysis followed by replication identifies loci in or near CDKN1B, TET3, CD80, DRAM1, and ARID5B as associated with systemic lupus erythematosus in Asians. *American Journal of Human Genetics*, *92*(1), 41–51. <https://doi.org/10.1016/j.ajhg.2012.11.018>
- Yang, Y., Muzny, D. M., Reid, J. G., Bainbridge, M. N., Willis, A., Ward, P. A., Braxton, A., Beuten, J., Xia, F., Niu, Z., Hardison, M., Person, R., Bekheirnia, M. R., Leduc, M. S., Kirby, A., Pham, P., Scull, J., Wang, M., Ding, Y., ... Eng, C. M. (2013). Clinical Whole-Exome Sequencing for the Diagnosis of Mendelian Disorders. *New England Journal of Medicine*, *369*(16), 1502–1511. <https://doi.org/10.1056/NEJMoa1306555>
- Yu, Y., Bhangale, T. R., Fagerness, J., Ripke, S., Thorleifsson, G., Tan, P. L., Souied, E. H., Richardson, A. J., Merriam, J. E., Buitendijk, G. H. S., Reynolds, R., Raychaudhuri, S., Chin, K. A., Sobrin, L., Evangelou, E., Lee, P. H., Lee, A. Y., Leveziel, N., Zack, D. J., ... Seddon, J. M. (2011). Common variants near FRK/COL10A1 and VEGFA are associated with advanced age-related macular degeneration. *Human Molecular Genetics*, *20*(18), 3699–3709. <https://doi.org/10.1093/hmg/ddr270>
- Yun, T., Li, H., Chang, P.-C., Lin, M., Carroll, A., & McLean, C. (2020). Accurate, scalable cohort variant calls using DeepVariant and GLnexus. *BioRxiv*, 2020.02.10.942086. <https://doi.org/10.1101/2020.02.10.942086>
- Zajickova, K., Vrbikova, J., Canaff, L., Pawelek, P. D., Goltzman, D., & Hendy, G. N. (2007). Identification and functional characterization of a novel mutation in the calcium-sensing receptor gene in familial hypocalciuric hypercalcemia: Modulation of clinical severity by vitamin D status. *Journal of Clinical Endocrinology and Metabolism*, *92*(7), 2616–2623. <https://doi.org/10.1210/jc.2007-0123>
- Zhao, W., Rasheed, A., Tikkanen, E., Lee, J. J., Butterworth, A. S., Howson, J. M. M., Assimes, T. L., Chowdhury, R., Orho-Melander, M., Damrauer, S., Small, A., Asma, S., Imamura, M., Yamauch, T., Chambers, J. C., Chen, P., Sapkota, B. R., Shah, N., Jabeen, S., ... Saleheen, D. (2017). Identification of new susceptibility loci for type 2 diabetes and shared etiological pathways with coronary heart disease. *Nature Genetics*, *49*(10), 1450–1457. <https://doi.org/10.1038/ng.3943>
- Zimmerman, P. A., Buckler-White, A., Alkhatib, G., Spalding, T., Kubofcik, J., Combadiere, C., Weissman, D., Cohen, O., Rubbert, A., Lam, G., Vaccarezza, M., Kennedy, P. E., Kumaraswami, V., Giorgi, J. V., Detels, R., Hunter, J., Chopek, M., Berger, E. A., Fauci, A. S., ... Murphy, P. M. (1997). Inherited resistance to HIV-1 conferred by an inactivating mutation in CC chemokine receptor 5: Studies in populations with contrasting clinical phenotypes, defined racial background, and quantified risk. *Molecular Medicine*, *3*(1), 23–36. <https://doi.org/10.1007/bf03401665>