ESC European Society of Cardiology

Cardiovascular Research (2018) **114**, 1226–1240
doi:10.1093/cvr/cvy068

SPOTLIGHT REVIEW

Challenges and opportunities in stroke genetics

Rainer Malik[1] and Martin Dichgans[1,2]*

[1]Institute for Stroke and Dementia Research (ISD), University Hospital, Ludwig-Maximilians-Universität (LMU) München, Feodor-Lynen-Straße 17, 81377 Munich, Germany; and [2]Munich Cluster for Systems Neurology (SyNergy), Feodor-Lynen-Straße 17, 81377 Munich, Germany

Received 15 September 2017; revised 22 December 2017; editorial decision 4 January 2018; accepted 14 March 2018; online publish-ahead-of-print 15 March 2018

Abstract

Stroke, ischaemic stroke and subtypes of ischaemic stroke display substantial heritability. When compared with related vascular conditions, the number of established risk loci reaching genome-wide significance for association with stroke is still in the lower range, particularly for aetiological stroke subtypes such as large artery atherosclerotic stroke or small vessel stroke. Nevertheless, for individual loci substantial progress has been made in determining the specific mechanisms mediating stroke risk. In this review, we present a roadmap for functional follow-up of common risk variants associated with stroke. First, we discuss *in silico* strategies for characterizing signals in non-coding regions and highlight databases providing information on quantitative trait loci for mRNA and protein expression, as well as methylation, focussing on those with presumed relevance for stroke. Next, we discuss experimental strategies for following up on non-coding risk variants and regions such as massively parallel reporter assays, proteome-wide association studies, and chromatin conformation capture (3C) assays. These and other approaches are relevant for gaining insight into the specific variants and mechanisms mediating genetic stroke risk. Finally, we discuss how genetic findings could influence clinical practice by adding to diagnostic algorithms and eventually improve treatment options for stroke.

Keywords Stroke • Post-GWAS • Ischaemic stroke • GWAS

This article is part of the Spotlight Issue on genome-wide association studies to cardiovascular research.

1. Stroke is a complex phenotype caused by multiple aetiologies

Stroke is the second most common cause of death and disability worldwide.[1] Stroke presents as a heterogeneous disease, with about 20% of stroke cases being of haemorrhagic origin (intracerebral haemorrhage, ICH) and 80% being of ischaemic origin (ischaemic stroke, IS). The ischaemic subtype can further be divided into aetiological subtypes according to the Trial of Org 10172 in Acute Stroke Treatment (TOAST) criteria,[2] which include large artery atherosclerotic stroke (LAS), cardioembolic stroke (CES), stroke due to small vessel disease (SVS), known other aetiologies, and unknown aetiologies.

The incidence and prevalence of aetiological stroke subtypes varies greatly depending on ethnicity. Thus, for example, the incidence of ICH in Asian populations is almost two-times higher than in other ethnicities.[3] The frequency of aetiological IS subtypes likewise varies with SVS being more frequent in Asian ethnicities, while CES is more frequent in patients of European ancestry. Also, there is a significant difference in the proportion of extracranial vs. intracranial atherosclerosis in Asian and European patients, with intracranial atherosclerosis being more prominent in Asian countries.[4] Whether there is a genetic basis to this differential distribution or whether lifestyle factors affecting vascular risk factors play a more prominent role is still under debate.

The complexity and heterogeneous nature of stroke has posed a challenge to the discovery of genetic risk loci because of limited sample sizes for aetiological stroke subtypes and the requirement to harmonize definitions and phenotyping protocols across multiple sites.

2. Stroke phenotyping poses a challenge for stroke genetics

The complex nature of stroke is reflected by substantial efforts needed for precise phenotyping and by the existence of alternative classification systems. Although many phenotypes, especially those on a continuous scale (e.g. height, blood pressure) are relatively easy to assemble, stroke phenotyping is an intricate exercise.

Non-contrast computed tomography remains the standard imaging modality for the initial evaluation of patients with suspected stroke. It may identify early signs of IS, but most importantly is highly sensitive in detecting acute intracerebral or subarachnoid haemorrhage. Detailed phenotyping and sub-classification requires additional investigations including imaging of large arteries for the detection of large artery

* Corresponding author. Tel: +49 894 400 46018; fax: +49-89-4400-46040, Email: martin.dichgans@med.uni-muenchen.de

stenosis, cardiac investigations for the detection of cardiac sources of embolism, and, where available, brain magnetic resonance imaging (MRI) for the detection and detailed assessment of small infarcts and other brain pathologies.

To account for the heterogeneous nature of stroke, other classification schemes for stroke besides the traditional TOAST classification have been developed, with the causative classification for stroke (CCS)[5] being a newly developed, computerized, evidence-based algorithm that provides both causative and phenotypic stroke subtypes in a rule-based manner. CCS can provide 8 or even 16 subtypes that specify the level of confidence of an assignment. Yet, whether CSS or other alternative classification systems such as ASCO[6] are better suited than TOAST for gene identification remains to be determined.

3. Stroke displays substantial heritability

Twin studies suggest a significant heritability for stroke. Monozygotic twins are more likely (odds ratio ~2.0) to be concordant for stroke than dizygotic twins.[7,8] A substantial heritability of stroke is further confirmed by single nucleotide polymorphism (SNP)-based pseudo-heritability measures derived from genome-wide association study (GWAS). However, the two studies[9,10] reporting GWAS-derived measures using the GCTA software[11] show some degree of variation. Both studies agreed on the heritability estimate for stroke as a whole (~40%). However, heritability estimates for stroke subtypes varied markedly. Stills, LAS always showed the highest heritability measures (40.3 and 66%) while SVS showed the lowest measures (16.1 and 10%). With only few genome-wide loci for stroke identified thus far, it seems there is still a substantial proportion of missing heritability in stroke.

4. Common genetic variants associated with stroke

Recent GWAS of stroke and related phenotypes have found several risk loci to be associated with stroke and stroke subtypes (*Figure 1*). Interestingly, most genome-wide associations are largely restricted to aetiological stroke subtypes of IS: *HDAC9*,[12,13] *SERPINA1*,[14] and *TSPAN2*[15] have been shown to associate with LAS, whereas *PITX2*[16] and *ZFHX3*[17] associate with CES. Using traditional GWAS approaches, no genome-wide significant association has been discovered for SVS. Yet, using a covariate-informed approach, a recent study found a locus at chr16p24 to be associated with SVS.[18] This signal is situated in an intergenic region between *JPH3* and *ZCCHC14*, with the lead SNP being associated with decreased expression of *ZCCHC14* (Zinc Finger CCHC-Type Containing 14) in tibial arterial tissue.

Some genome-wide signals for stroke are associated with multiple subtypes of IS. This specifically applies to the chr12q24[19] and the *ABO* locus[13] and might in part be mediated through established vascular risk factors for stroke (e.g. hypertension) acting on multiple IS subtypes simultaneously. Despite an even lower sample size, previous GWAS in ICH have revealed two replicated genome-wide signals on 1q21 (for non-lobar ICH) and 12q21 (for lobar ICH).[20] Two published loci for IS have not been successfully replicated in subsequent studies with larger sample size.[9,21]

4.1 Exonic variants in stroke

As in other complex diseases, most of the GWAS-signals for stroke are non-coding and therefore pose challenges to the biological follow-up of these signals. Yet, some variants are exonic, enabling direct insight into disease pathophysiology through amino-acid exchanges.

The gene encoding SH2B adapter protein 3 (*SH2B3* or *LNK*) is located on chr12q24. Variants at this locus are associated with multiple conditions including cardiovascular risk factors and vascular diseases (reviewed in[22]). Due to the intricate linkage disequilibrium (LD) structure at this locus and the large number of genes contained within the region, it is difficult to pinpoint a specific gene underlying the increased risk for stroke. However, some evidence may come from experimental studies of *SH2B3*.[23] An exonic variant [R262W, minor allele frequency (MAF) 0.43 in Europeans] in this gene has been shown to cause loss of function of *SH2B3*. This amino-acid change is associated with expansion of haematopoietic stem cells and enhanced megakaryopoiesis in human cord blood. In mice, haematopoietic *Sh2b3* deficiency leads to accelerated arterial thrombosis and atherosclerosis, but only in the setting of hypercholesterolaemia. Together, these findings define *SH2B3* as the lead candidate for the causal gene at the chr12q24 locus.

SERPINA1 has recently been added to the list of genes implicated in LAS pathophysiology.[14] *SERPINA1*, encoding alpha-1-antitrypsin (AAT) was identified by an exome-centric approach using the Illumina exome chip. A common variant leading to an amino acid exchange (V213A, MAF 0.17 in Europeans) in AAT was found to be exome-wide significant in a meta-analysis of LAS cases and stroke-free controls from different ethnicities. The availability of microscale thermophoresis assays previously developed in the analysis of AAT provided an excellent opportunity to test the biophysical properties of the mutant vs. wild-type AAT protein. Indeed, mutant AAT M1(V213) showed a higher dissociation constant with its main interaction partner neutrophil elastase (NE) in lipid-rich plasma, but not in lipid-free plasma. Hydrogen-deuterium exchange experiments further showed an enhanced structural flexibility in the mutant protein. The target enzyme of AAT, NE, is expressed by macrophages in advanced atherosclerotic lesions.[24] AAT might facilitate protection against matrix breakdown by NE and clearance of lipoprotein deposits. However, whether AAT is protective against atherosclerosis is still debated. Recent studies have provided evidence for an interaction between AAT and lipoproteins including both low-density lipoprotein (LDL) and high-density lipoprotein (HDL). Binding of AAT to reconstituted HDL reduced the inhibitory capacity of AAT towards NE,[25] similar to removing lipids from plasma. Moreover, the allele-specific differences of the dissociation constant disappeared, supporting the hypothesis of an altered interaction of M1(V213) AAT with lipoproteins in the complex plasma environment.

4.2 Non-coding variants in stroke

As for other conditions, a key task in the post-GWAS era of stroke is to determine how risk variants not interfering with protein structure mediate increased risk for stroke. As indicated, the majority of genome-wide associated signals reside in non-coding regions. Several non-coding variants have been associated with high confidence with stroke or stroke subtypes. Some of them are also associated with related cardiovascular phenotypes, making them high priority for biological follow-up. For these non-coding variants, there are two recent examples where a functional role of the genomic region harbouring the risk variant could be identified through wet-lab experiments.

HDAC9 is not only the strongest genetic risk factor for LAS, but is also among the top associated signals in coronary artery disease (CAD)[26] and

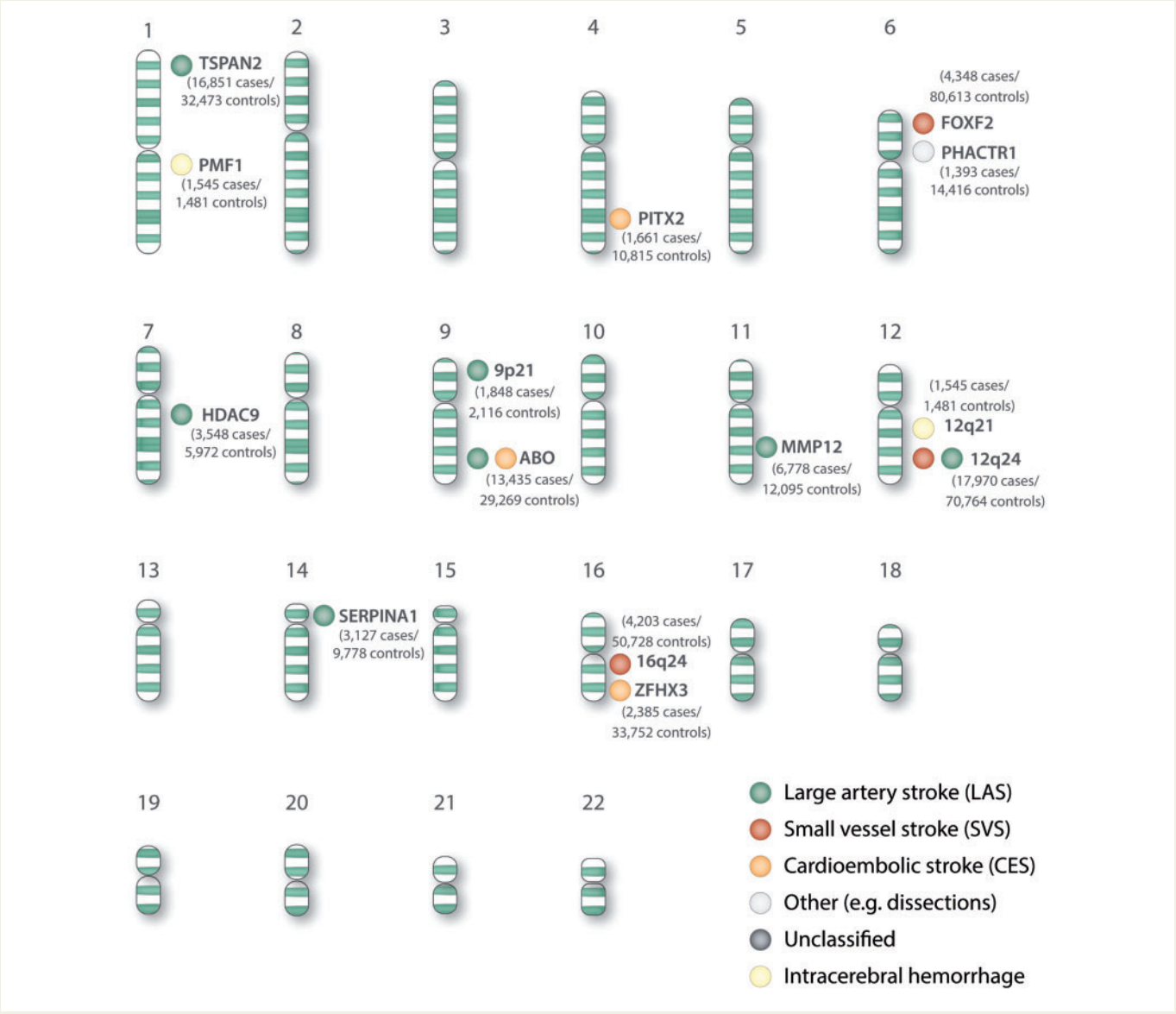

Figure 1 Risk loci for stroke and stroke subtypes. Shown are the genomic locations of published risk loci reaching genome-wide significance in previous GWAS for stroke. In some cases associations are limited to a specific stroke subtype, while others are associated with multiple stroke subtypes. Sample sizes of the original GWAS study are given below each locus.

other cardiovascular diseases.[27,28] rs2107595, the lead SNP for LAS at this genomic locus is located in an enhancer element downstream of *HDAC9* and upstream of *TWIST1/FERD3L*. rs2107595 has been shown to act as an expression quantitative trait locus (eQTL) for *HDAC9* in PBMCs.[29] Consistent with directionality, *Hdac9* deficiency was found to attenuate atherosclerosis in two different mouse models for atherosclerosis, *ApoE*$^{-/-}$ and *Ldlr*$^{-/-}$ mice.[29,30] Collectively, these findings imply that the effects of risk alleles at *HDAC9* on stroke risk are at least in part mediated through elevated *HDAC9* expression levels. Thus, inhibition of Class II a histone deacetylases[31] or targeted inhibition of *HDAC9* using *HDAC9*-specific inhibitors have been proposed as a viable strategy for atheroprotection and prevention of related clinical manifestations including stroke.[29]

The discovery of the *FOXF2/FOXQ1* locus exemplifies how a specific gene located in the vicinity of common risk variants can be linked to a stroke phenotype through independent experimental data downstream of the GWAS approach. A common variant at the *FOXF2/FOXQ1* locus was recently found to be associated with any stroke (ischaemic + haemorrhagic stroke) using a traditional GWAS strategy.[32] The lead SNP is located between the two forkhead box genes *FOXQ1* and *FOXF2*, with no eQTL information available. The association signal was particularly strong for SVS. Experimental work further showed that conditional deletion of *Foxf2* in adult mice leads to cerebral infarction, reactive gliosis, and microhaemorrhage. Experimental studies performed in another study provided insights on how *FOXF2* might influence small vessel pathology.[33] *Foxf2* is required for brain pericyte differentiation and for the development and maintenance of the blood-brain-barrier (BBB) in mice. Pericytes from *Foxf2*$^{-/-}$ mice are relatively abundant in brain microvessels but express very low levels of *PDGFRb*, a marker of pericyte differentiation. *Foxf2*$^{-/-}$ mice further develop BBB leakage along with perivascular oedema and a reduction in *transforming growth factor (TGF) beta* signalling. These findings

are well in line with observations from mouse models for monogenic small vessel disease that have suggested a role of pericytes and the BBB[34] as well altered *TGF beta* signalling[35] in cerebral small vessel disease.

5. Approaches for functional follow-up of variants in regulatory regions

Other than for *HDAC9* and *FOXF2*, for most of the known non-coding GWAS signals for stroke, the relevant gene responsible for mediating stroke risk is still elusive. In the following we discuss currently used approaches for discovering mechanisms that mediate stroke risk. We start by providing an overview on *in silico* methods, exploiting information on gene expression, protein abundance and methylation patterns to identify relevant genes and gene products for stroke. These approaches enable identifying or prioritizing relevant tissues and cell types implicated in mediating genetic risk for developing stroke. In the subsequent section we discuss traditional and recently introduced wet-lab techniques providing direct experimental evidence for a specific regulatory mechanism.

5.1 In silico approaches

As information on eQTL, protein QTL (pQTL) and methylation QTL (meQTL) is accumulating,[36–44] exploitation of publicly and non-publicly available data, and integration of the multiple layers of information has become a rich source for functional follow-up of GWAS signals.eQTLs describe variants associated with changes in the transcriptome, either acting in a cell-type specific manner on nearby transcripts (*cis*) or on transcripts being further away in the genome (*trans*). For many complex diseases candidate tissues and cell types are rather obvious [e.g. cardiomyocytes for myocardial infarction (MI); immune cells for autoimmune diseases]. For stroke, however, this is less straight-forward, requiring the inclusion of extensive datasets from multiple sources.

Cellular constituents of the neurovascular unit (NVU, including vascular endothelial cells and pericytes), vascular smooth muscle cells, macrophages and constituents of peripheral blood (e.g. immune cells, platelets, coagulation factors) are all assumed to play a role in stroke pathophysiology. Stroke susceptibility is further influenced by resident brain cells (e.g. neurons, astroglia, oligodendroglia, microglia) which also contribute to stroke recovery, secondary responses, and outcome. Hence, there is a requirement to expand on tailored eQTL datasets for stroke and stroke subtypes, as candidate tissues and cell types also vary between IS subtypes. For instance, studies on LAS require data on carotid and aortic tissue whereas studies on SVS require data on cellular constituents of the NVU. However, the acquisition of such data is intrinsically difficult, as patient samples e.g. brain vascular endothelial cells or brain microvessels are difficult to obtain and, ideally, eQTL information would be extracted from a truly representative sample including cases and controls as this would remove bias in the analysis.

A selection of currently available eQTL datasets with particular relevance for stroke is presented in *Table 1*. These ressources represent the

Table 1 Overview of available eQTL, pQTL, and meQTL datasets relevant for stroke (selection)

Dataset	PMID	Tissues represented in dataset	Number of individuals
eQTL			
GTEx V6	25954001	44 post-mortem tissues[a]	449
GRASP2	25428361	2082 published GWAS results including eQTLs	
HGVD	26911352	Peripheral blood cells	1208
BIOS	27918535	Whole blood	2116
Blueprint epigenome project	22398613	Monocytes, granulocyte neutrophils, eosinophils, macrophages (M0, M1 and M2), naive CD4+ and naive CD8+ cells and cell line samples	197
STARNET	27540175	Atherosclerotic-lesion-free internal mammary artery, atherosclerotic aortic root, subcutaneous fat, visceral abdominal fat, skeletal muscle, and liver	600 patients coronary artery disease
Aortic endothelial cells study	23667179	Aortic endothelial cells	147
Religious order study	22471860	Brain-specific cell types (e.g. neurons, astrocytes)	1240
Rush Memory and Aging Project	22471867	Brain-specific cell types (e.g. neurons, astrocytes)	1752
pQTL			
KORA	28240269	Blood plasma samples	1000
meQTL			
Blueprint epigenome project	22398613	Monocytes, granulocyte neutrophils, eosinophils, macrophages (M0, M1 and M2), naive CD4+ and naive CD8+ cells and cell line samples	197
ARIC	25935004	Leukocytes	2097

Given are the name of the study, the reference to the published data and an overview of the samples included.
[a]In alphabetical order: Adipose—Subcutaneous, Adipose—Visceral (Omentum), Adrenal Gland, Anterior cingulate cortex (BA24), Artery—Aorta, Artery—Tibial, Brain, Breast—Mammary Tissue, Caudate (basal ganglia), Cells—EBV-transformed lymphocytes, Cells—Transformed fibroblasts, Cerebellar Hemisphere, Cerebellum, Colon—Sigmoid, Colon—Transverse, Cortex, Esophagus—Gastroesophageal Junction, Esophagus—Mucosa, Esophagus—Muscularis, Frontal Cortex (BA9), Heart—Atrial Appendage, Heart—Left Ventricle, Hippocampus, Hypothalamus, Liver, Lung, Muscle—Skeletal, Nerve—Tibial, Nucleus accumbens (basal ganglia), Ovary, Pancreas, Pituitary, Prostate, Putamen (basal ganglia), Skin—Not Sun Exposed (Suprapubic), Skin—Sun Exposed (Lower leg), Small Intestine - Terminal Ileum, Spleen, Stomach, Testis, Thyroid, Uterus, Vagina and Whole Blood.

best annotated datasets. While most studies are primary studies with data collected from different tissues of non-diseased individuals, the GRASP2 database collects information on eQTLs by automatically extracting information from publications and their Supplementary material and thus represents the largest curated database on eQTLs with information on many relevant tissues for stroke.

Another dataset with obvious relevance to stroke and cardiovascular disease is STARNET, which provides data on atherosclerotic-lesion-free internal mammary artery, atherosclerotic aortic root, subcutaneous fat, visceral abdominal fat, skeletal muscle, and liver from diseased individuals.[41]

An alternative approach to performing lookups in eQTL databases is performing a transcript-wide association study (TWAS).[45] TWAS integrates GWAS and eQTL data by predicting gene expression levels for individuals with genotypic information from gene expression data of external studies. This way, the expression level of a gene is imputed for each individual with a given genotype. Association between the expression and stroke trait is indirectly estimated by linear combination of SNP-trait standardized effect sizes and weighting the predicted expression while accounting for LD among SNPs. TWAS benefits from a lower multiple-testing burden by probing several thousands of genes. TWAS also has the benefit of prioritizing the likely causal gene within an associated locus that contains multiple genes since the results are gene-based. A limitation of TWAS is the requirement for individual level data whereas large-scale GWAS data for stroke are mostly available as summary association statistics.pQTL analyses (e.g. using the SomaSCAN platform[46] or olink[47]) provide information on multiple proteins that can be assessed for functional follow-up. In pQTL analyses, individual genetic variants are correlated with differential protein expression. pQTL studies so far have mostly been performed on healthy individuals, but recent efforts have started to link genetic variation with changes in protein expression in complex diseases.[48,49] Interestingly, transcriptome and proteome studies have shown a relatively low correlation between mRNA and protein abundance. This likely relates to different protein turnover, RNA stability or other issues.

For meQTL analysis, methylation is treated as a quantitative trait dependent on the genotype. In meQTL studies, hundreds of thousands of CpGs can be studied simultaneously, offering highly complementary data on the regulatory capacity of genomic regions. meQTLs can act in cis or trans, and it has been shown that disease-associated variants can have widespread effects on DNA methylation in *trans*, likely reflecting differential occupancy of *trans*-binding sites by cis-regulated transcription factors.[39] This technique has recently been used to study the association of cerebral white matter hyperintensity (WMH) burden on MRI with accelerated ageing in African-American participants of the Atherosclerosis Risk in Communities (ARIC) study. DNA methylation age acceleration represents the deviation of the DNA methylation-predicted age from the chronologic age.[50] There was a significant association between accelerated epigenetic aging and increased WMH burden, independent of known risk factors, including chronologic age. Another study tested whether biological age estimated by DNA methylation predicts functional outcome 3 months post-IS better than chronological age.[51] Interestingly, biological age outperformed chronological age and remained a significant predictor of clinical even when adjusting for chronological age, stroke severity, and recanalization treatment. This study also highlights the potential influence of (epi)genetic variation on stroke outcome. Taken together, meQTL analysis already in its early stages shows promise in stroke research.

As mentioned earlier, determining the cell types and tissues responsible for mediating stroke risk remains a challenge. An indirect method to determine the specific cell types involved in the pathogenesis of stroke is derived from GWAS summary statistics. The hypothesis behind this approach is that genome-wide significant or sub-threshold signals originating from GWAS will be enriched in functional epigenetic chromatin marks like H3K4me3 (active promoters) or H3K4me1 (active enhancers) in a cell-type specific manner.[52] Tools like epigwas[53] make use of this information compiled by the ENCODE[54] and the Epigenomics Roadmap consortium.[55] Such enrichment has convincingly been shown for rheumatoid arthritis (RA) which is characterized by an enrichment of GWAS signals in regulatory variants from various immune cells.[56]

Expectedly, methods like epigwas will provide insight in the specific cell types and tissues causally involved in stroke pathogenesis. However, this information is reliant on the cell types and tissues provided by ENCODE and Roadmap. Although Roadmap includes blood, brain, and vascular (heart and human umbilical vein endothelial cells) tissue information, many possibly relevant cell types for stroke and stroke subtypes are missing.

Given these considerations, information on cell-type-specific enrichment of epigenetic marks might guide the prioritization of cell types for lookups in eQTL, meQTL or pQTL databases. The information obtained through these sources can then be integrated into a 'biological score' or 'functional score'.[57] This approach assigns an arbitrary score to each gene in risk loci for stroke for each category (e.g. eQTL, pQTL, exonic variant) that is deemed to be relevant for functional follow-up, resulting in an ordered list of genes that might be causally involved in stroke pathogenesis.

5.2 Mutation databases

Identification of disease-related genes by whatever method helps prioritizing genes for targeted analysis of rare exonic mutations potentially linked more directly to the disease. The Genome Aggregation Database (gnomAD, formerly EXAC) database[58] offers free access to whole exome and whole genome data, making it an invaluable resource for further bioinformatics analyses. The latest gnomAD release spans 123 136 exomes and 15 496 genomes from unrelated individuals sequenced as part of various disease-specific and population genetic studies.

In a similar approach, deCODE sequenced 2636 Icelanders to a median depth of 20× and imputed variants of 104 220 individuals down to a MAF of 0.1%.[59] Additional targeted re-sequencing of risk loci will provide further evidence of potential rare mutations contributing to the missing heritability in stroke.

5.3 Experimental approaches

A complementary approach to *in silico*-based methods of gathering information through published databases is the *de novo* creation of functional data through wet-lab experiments. In this section, we briefly highlight three techniques that show promise for the experimental follow-up of non-coding variants. We further briefly review gene editing methods that can be used to study the effect of a specific genetic variant in induced pluripotent stem cells that can subsequently be differentiated to specific (vascular) cell types.

Massively parallel reporter assays (MPRAs) can be utilized to prioritize GWAS variants according to their transcriptional activity.[60,61] A hallmark of this method is the identification of proximal promoter or enhancer elements in a genomic region associated with stroke and testing the effect of natural genetic variation in an identified set of putative promoters, enhancers or silencers. The overall workflow of an MPRA

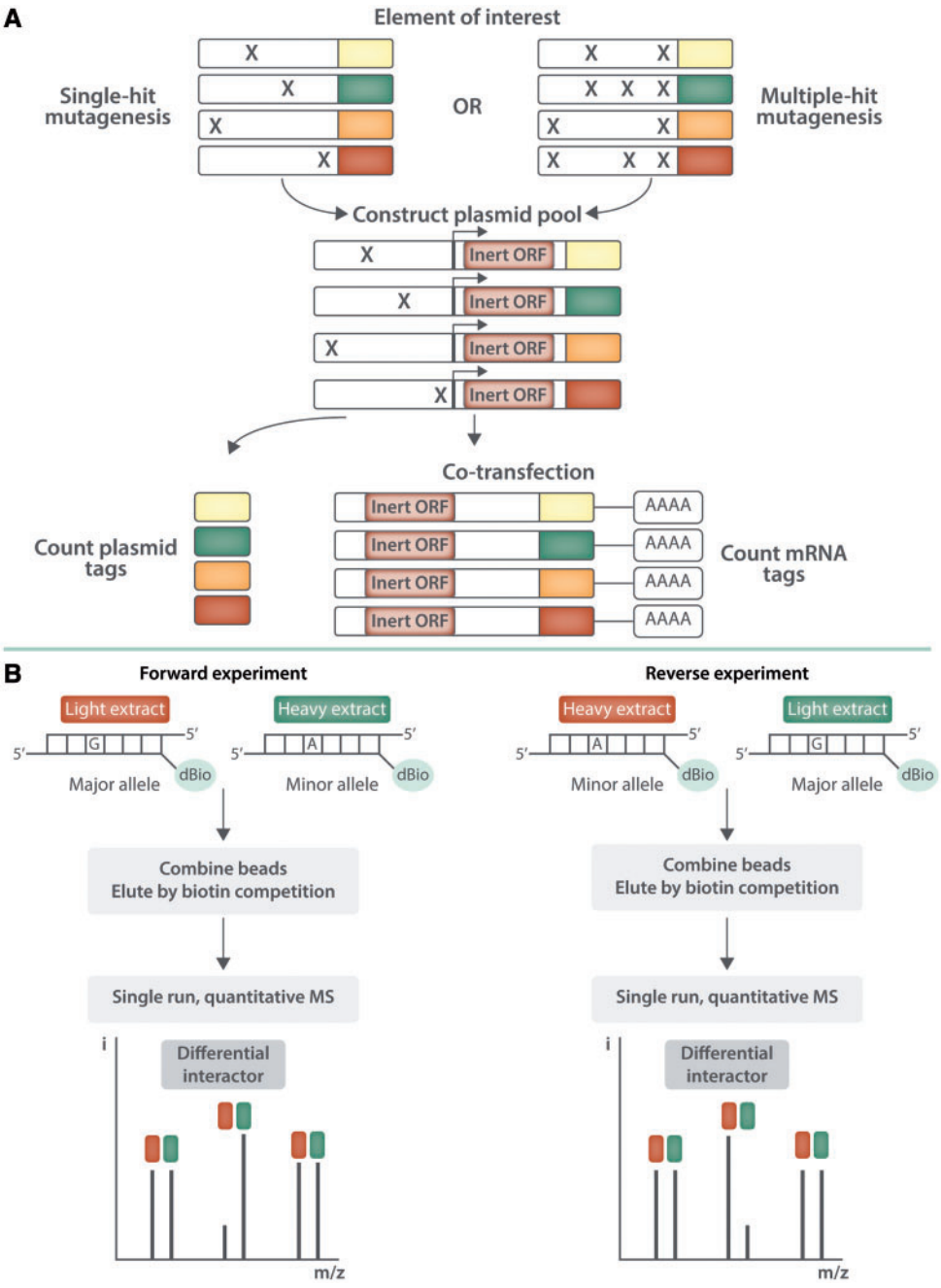

Figure 2 (*A*) Schematic illustration of a MPRA experiment. Oligonucleotides containing enhancer variants coupled to distinguishing tags are first generated using microarray-based DNA synthesis. An invariant promoter-open reading frame (ORF) segment is then inserted between the variants and tags by double digestion and directional ligation. The resulting reporter plasmid pool is co-transfected into a population of cells. The relative regulatory activities of the transfected variants are inferred by sequencing and counting their corresponding tags from the cellular mRNA and the transfected plasmid pool. (*B*) Schematic illustration of a PWAS experiment. Synthetic oligonucleotides containing the SNP are phosphorylated, polymerized and subsequently labelled by desthiobiotin in a strand-specific manner. The immobilized DNA fragments are incubated with either light or heavy nuclear extract from the cell type of interest. After removal of unbound proteins, bead fractions are combined and DNA-protein complexes are eluted with biotin. The eluate is precipitated, digested and analysed by single-run, high resolution, quantitative MS. Specific interaction partners result in a ratio different from 1:1, demonstrating specific enrichment at one variant of the single nucleotide polymorphism.

experiment is illustrated in *Figure 2A*. In short, the method involves synthesizing, in parallel, the promoters or enhancers along with a transcriptional start site that is followed by a specific 20-bp 'barcode' sequence. After co-transfection of multiple of these constructs, the mRNA and plasmid tags are counted. Regulatory sequences result in transcription of the barcode sequence thus identifying functional elements. By introducing genetic variation, the consequences of the variant of interest on transcriptional activity can be quantified.

Mass spectrometry (MS/MS)-based *proteome-wide association studies* (PWAS, *Figure 2B*)[62] allow identifying allele-specific binding of transcription factors to regulatory sequences harbouring the variant of interest. Variants disrupting or generating a transcription factor-binding site result in altered affinity to the respective transcription factor, subsequently resulting in a change of transcript levels. Differential binding can be measured using quantitative MS, thus revealing which transcription factors and transcriptional cascades could be contributing to stroke.

Chromatin conformation capture (3C) assays allow identifying physical interactions between chromosomal elements including interactions between elements that are physically far apart from one other.[63] These assays can be used to quantify the number of interactions between genomic loci that are located in close spatial proximity regardless of their distance in the nucleotide sequence. A typical application is studying the interaction between a promoter and an enhancer element, either between two single loci (3C) or one locus vs. the entire genome (4C). A more detailed description of these methods is provided in recent topical reviews.[64]

In order to be informative with respect to stroke the three methods described earlier should ideally be conducted in appropriate cell systems and expression frameworks.[65,66] Ideally, this would involve experiments in r multiple cell types relevant for stroke to gather a complete picture of the complex regulatory actions involved in disease pathogenesis.

5.4 Gene editing

An increasingly used technology to functionally characterize individual variants is genome editing using CRISPR/CAS9.[67–69] Induced pluripotent stem cells (iPSCs) or embryonic stem cells (ESCs) with knock-out or knock-in-mutations can be compared within the same genomic context, making this strategy particularly informative. Although primarily used in the context of exonic mutations, genome editing can also be utilized to study more subtle effects of regulatory variants. This also includes the generation of cells heterozygous for the variant of interest.[70] In a landmark study using CRISPR/Cas9 models, Gupta *et al.*[71] recently examined the role of a single SNP (rs9349379) in *PHACTR1* on five vascular diseases (*Figure 1*) previously shown to associate with this variant on a genome-wide level of significance. The minor allele was shown to be associated with an increased risk of both CAD and coronary calcification and with a reduced risk of migraine, cervical artery dissection, fibromuscular dysplasia, and hypertension. Cervical artery dissection is a frequent cause of juvenile stroke. However, variants at *PHACTR1* have so far not been associated with stroke. rs9349379 was further shown to be associated with reduced arterial stiffness and a reduction in flow-mediated vasodilation. iPSC-derived vascular endothelial and smooth muscle cells carrying a 88-bp deletion around rs934939 were shown to exhibit increased endothelin-1 expression. This effect was further confirmed in ESCs and in plasma samples of human subjects, suggesting that activation of the endothelin-A receptor results in vasoconstriction, vascular smooth muscle cell (VSMC) proliferation, extracellular matrix (ECM) production, and fibrosis. Whether this specific mechanism also plays a role in stroke, needs to be determined.

Again, choosing the right cell types for CRISPR/CAS9 models derived from stroke GWAS loci remains a challenge. Still, the use of iPSC- and ESC-derived genome-edited cellular models has huge potential. Although vascular cell types will be more informative for LAS, brain cell types could be more of interest for SVS.

6. Translating genetic findings to clinical practice

Translating recent discoveries in stroke genetics into clinical practice remains a challenge. Potential applications include the identification of novel disease markers, risk prediction using genetic information, drug re-purposing and -positioning, and the determination of causal relationships using mendelian randomization (MR).

6.1 Biomarker discovery

Although for other cardiovascular diseases blood-based diagnostic markers with high clinical relevance have been available for years (e.g. troponin for MI,[72] D-dimer for venous thrombosis[73,74]), there are no such markers for stroke. Among the clinically most relevant questions in patients with acute stroke is the discrimination between ischaemic and haemorrhagic stroke. Efforts have been made to identify blood-based biomarkers, with mixed results. Glial fibrillary acidic protein (GFAP) shows promise for differentiating ischaemic from haemorrhagic stroke.[75] GFAP is anastroglial protein that is detectable in very low concentrations in the blood of healthy individuals. Owing to the instantaneous destruction of astrocytes and disruption of the BBB following ICH GFAP is instantly released into the plasma. In contrast, in IS, the structural integrity of brain cells and the BBB remain preserved for some time after symptom onset.[76,77] However, due to a limited sensitivity and specificity for ICH GFAP has so far not entered the clinic.

Later, it has been shown that not only proteins can act as biomarkers in stroke. A combination of three microRNAs has recently been shown to yield superior discrimination between IS, stroke mimics, transient ischemic attacks (TIAs) and healthy controls compared with established protein-based biomarkers.[78] Whether proteins derived from GWAS also display discriminative power to act as biomarkers needs to be determined.

6.2 Drug re-purposing

An increasingly recognized application of GWAS data is to link locus and gene discovery to information on drug targets and determine whether already existing drugs approved for other indications can be re-purposed to target gene products and mechanisms relevant in the respective phenotypes under investigation. For several drugs, their immediate biological target (i.e. protein or peptide) is known. This information is held in the Therapeutic Targets Database (TTD)[79] and DrugBank.[80] Okada *et al.*[56] successfully exploited this information by linking data from the TTD and Drugbank to GWAS data on RA. Analysing 98 biological candidate genes they found a significant enrichment in approved drugs for RA suggesting that drugs previously not approved for RA but targeting RA candidate genes could potentially be repurposed to treat RA. This approach has successfully been used in various other conditions through their respective GWAS hits and may provide the stroke field with new candidate drugs derived from novel candidate genes.[81]

6.3 Mendelian randomization

MR is the method of choice to prove a causal role of a variant or a set of variants (genetic instrument) acting through an intermediate phenotype or risk factor on the final outcome of stroke (*Figure 3*).[82,83] MR analyses for IS subtypes have documented a causal relationship between Type 2 diabetes and LAS/SVS, but not fasting glucose levels, fasting insulin levels and body mass index.[84] Other MR studies on stroke have yielded negative results.[85–87] Whether these negative findings reflect a a truly negative

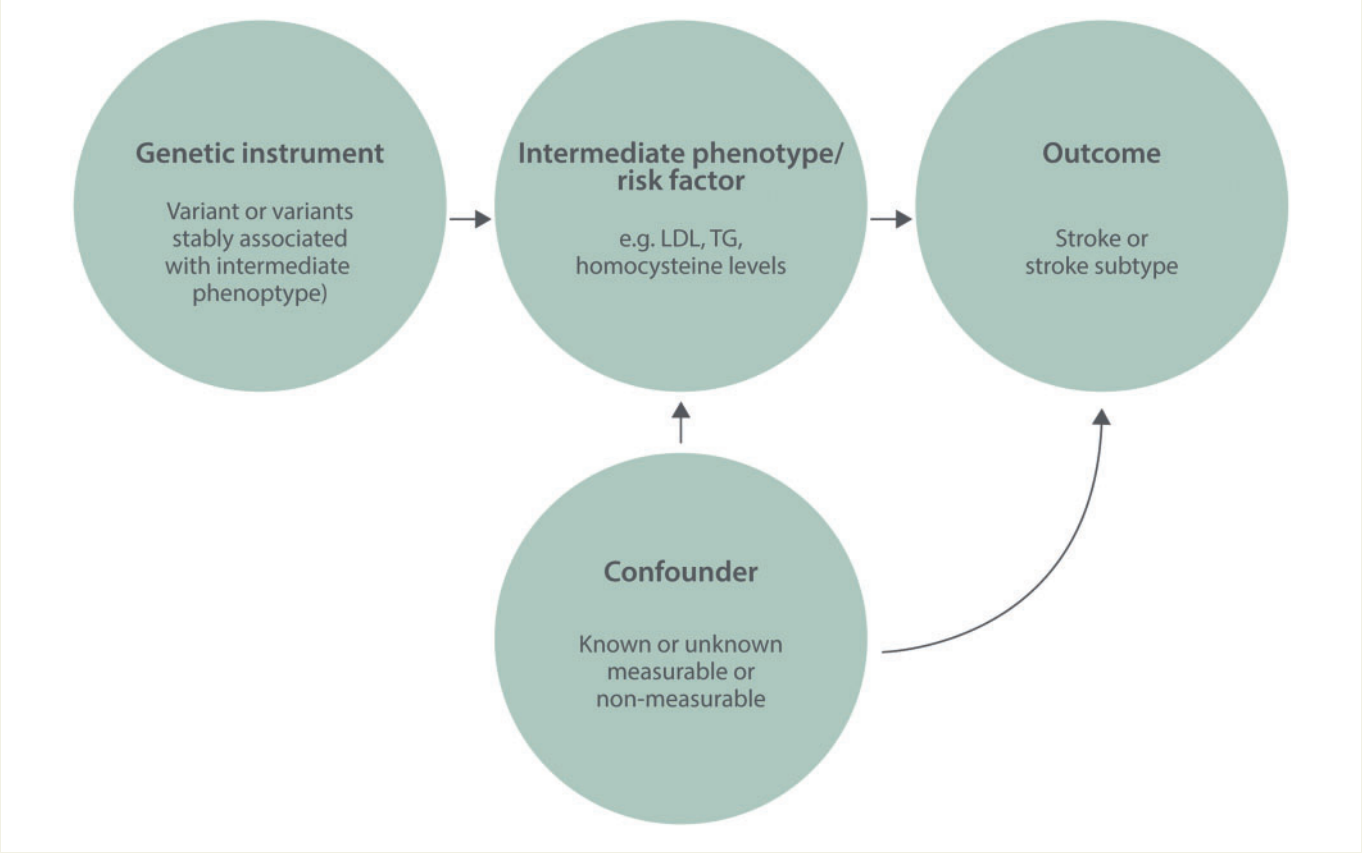

Figure 3 Schematic presentation of a MR experiment. MR works under following assumptions. Assumption 1: The genetic instrument (a SNP or a combination of multiple SNPs) is robustly associated with the intermediate phenotype or risk factor (e.g. blood lipids). This is typically accounted for by focussing on SNPs reaching genome-wide significance. Assumption 2: The genetic instrument is unrelated to any confounder biasing the relationship between the intermediate phenotype and the outcome. Assumption 3: The genetic instrument is related to the outcome only through its association with the intermediate phenotype. In other words, the genetic instrument must not be related to a second risk factor or intermediate phenotype.

association or relate to a lack of statistical power remains to be determined.MR is extensively reviewed in another part of this special issue. Still, we would like emphasize that sufficiently large sample sizes (outcome variable) are particularly important in MR analysis to prove causality. This remains a challenge in studies on stroke.

7. Future directions and summary

Stroke is a complex phenotype with a large proportion of heritability still unexplained. Larger GWAS meta-analyses with substantially increased sample sizes and integrating genetic information from multiple ethnicities to improve fine-mapping are currently underway. These studies are expected to reveal novel genome-wide significant loci for stroke and stroke subtypes. *Figure 4* exemplifies how these novel findings might be integrated into a work flow for functional follow-up of GWAS signals using SVS as an example and summarizing the *in silico* and wet-lab strategies mentioned earlier. GWAS remain key to the identification of risk regions. These risk regions can be further tested, using fine-mapping strategies, MPRA, PWAS or sequencing approaches for the identification of rare variants to compile a comprehensive list of genetic variants directly influencing stroke risk. Such rare variants can be exonic, leading to testable hypotheses regarding the mechanisms linking the respective gene and gene products to SVS. In the event of non-coding variants, eQTL/pQTL/meQTL information in relevant cell types combined with wet-lab experiments help prioritizing the candidate genes. These genes can further be studied in knock-out or knock-in models of relevant cell types or animal models. For SVS, the complete NVU, comprised of neurons, astrocytes, vascular endothelial cells pericytes and ECM components is expected to be a major factor in mediating disease risk and should therefore serve, in parts or as a whole, as the model of choice.

In addition to traditional GWAS analysis, the analysis of endophenotypes can also inform about variants and genes implicated in stroke risk. Endophenotypes are intermediate disease states that are associated with IS subtypes in the population, are independently and jointly heritable with IS, and are present in individuals with and without IS. For many IS subtypes including LAS, SVS, and CES, there are endophenotypes that have been investigated in depth. For instance, carotid plaque and carotid intima media thickness serve as endophenotypes for LAS and CAD/MI.[88] Genome-wide signals were reported in or around *ZHX2*, *APOC1*, and *PINX1* for cIMT and in *PIK3CG* and *EDNRA* for plaque presence. Silent brain infarcts and microbleeds are considered endophenotypes for SVS. This also applies to WMH, a quantitative phenotype that is currently moving into focus. A recent meta-analysis found *TRIM65*, *SH3PXD2A*, *EFEMP1*, *PMF1*, and *HAAO* to be associated with WMH burden on a genome-wide level.[89] Atrial fibrillation, a major cause of CES, is among

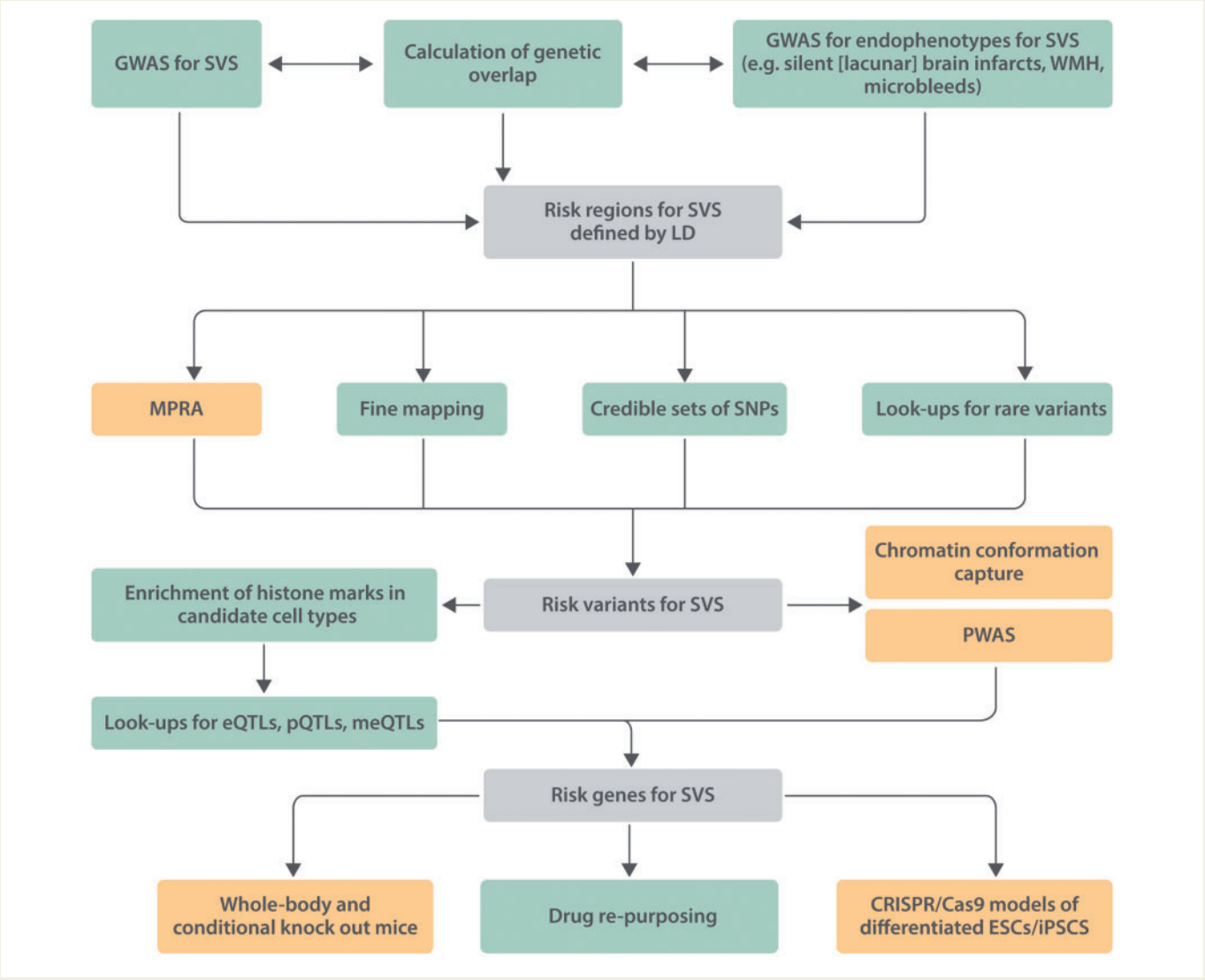

Figure 4 Potential work flow for functional follow-up of GWAS signals as exemplified for small vessel stroke: Information from *in silico* methods (green) and *in vitro* experiments (orange) can be integrated to form comprehensive lists of risk regions, risk variants, and risk genes for SVS. Technical advances in the generation of transgenic mice and CRISPR/Cas9 mediated genome editing facilitate functional exploration of risk variants and risk genes in *in vivo* and cellular models including differentiated ESCs/iPSCs. Candidate cell types for SVS include pericytes, vascular endothelial and smooth muscle cells, astrocytes, microglia, neurons and oligodendrocytes.

the most extensively studied endophenotypes with up to 25 genome-wide significant loci being reported thus far.[90] In general, quantitative traits such as WMH offer higher statistical power for locus discovery than dichotomous traits. As such, these endophenotypes remain of particular interest for stroke genetics.

Most, if not all GWAS studies have been performed on stroke risk. However, another important measure would be functional and cognitive outcome after stroke as measured by the modified Rankin scale[91] and targeted cognitive tests,[92] respectively. The discovery of genetic factors impacting on stroke outcome might help stratifying patients and offering personalized care to individuals at risk of worse outcome, including functional dependency or death

In contrast to other vascular conditions, there have been no sex-specific GWAS on stroke conducted so far. Epidemiological data show that men suffer from stroke more often than women, but even this unequal distribution is subtype-specific. An obstacle to sex-specific studies is the reduction in statistical power due to even lower sample sizes. Sex-specific signals are likely for stroke,[93] probably on a lower scale than for anthropomorphic traits.[94]

On a similar note, analysing non-autosomal data have been shown to yield significant results in several diseases.[95] The epidemiological difference between men and women in terms of stroke could potentially be attributed to genetic variation on the X-chromosome. With specialized imputation protocols, X-chromosomal markers can be analysed with the same accuracy as autosomal markers. Of note however, previous studies on CAD identified no genome-wide signals on the X-chromosome,[96] despite epidemiological evidence for sex-stratified risk.

Cross-phenotype analysis of diseases using LD score regression[97,98] or gwas-pw[99] can add to understanding disease mechanisms that are specific to stroke or a specific stroke subtype as compared with more

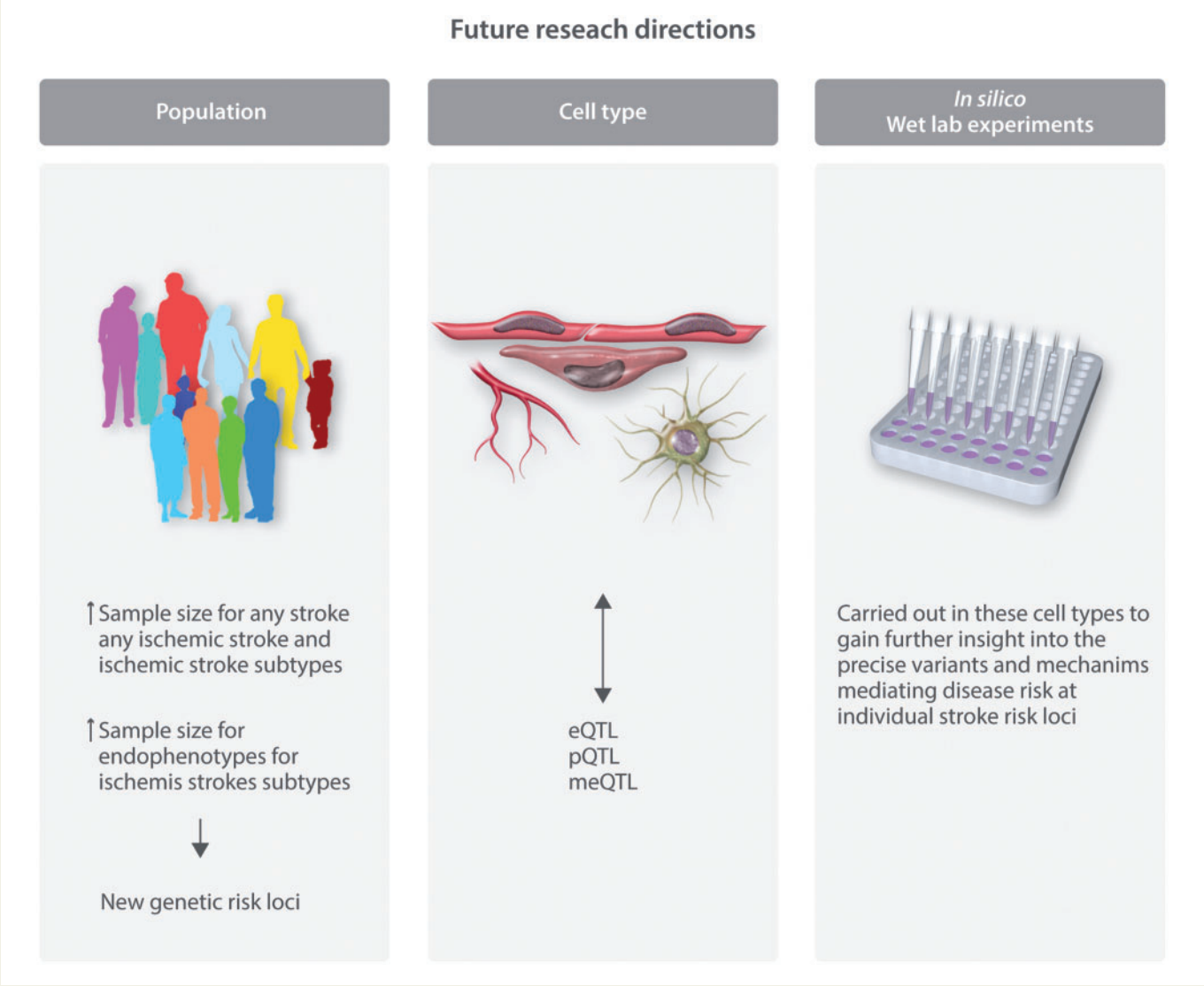

Figure 5 Key messages for readers.

pleiotropic signals or mechanisms associated with multiple related conditions and multiple stroke subtypes. Results from cross-phenotype analyses are available in the LDHub portal.[100] Along the same lines, it is now possible to conduct a complete phenome-wide association study where a single lead SNP or a panel of lead variants istested against multiple phenotypes, delivering a true account of genotype-to-phenotype relationships.[101,102] Recent data releases from the UK Biobank[103] offer a unique opportunity to study these relationships with regard to stroke GWAS signals.

7.1 Summary

Stroke is a complex phenotype. For some time insights into stroke pathophysiology from genetics have been lagging behind discoveries in other vascular phenotypes. Yet, with large multi-ancestry GWAS currently underway the number of reported risk loci for stroke and aetiological stroke subtypes is likely to increase in the near future. Combining *in silico* and experimental approaches to their functional exploration will improve our understanding of disease mechanisms in stroke and likely provide the field with novel targets and pathways for intervention (*Figure 5*).

Conflict of interest: none declared.

Funding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No 666881, SVDs@target (to M Dichgans) and No 667375, CoSTREAM (to M Dichgans); the DFG as part of the Munich Cluster for Systems Neurology (EXC 1010 SyNergy) and the CRC 1123 (B3) (to M Dichgans); the Corona Foundation (to M Dichgans); the Fondation Leducq (Transatlantic Network of Excellence on the Pathogenesis of Small Vessel Disease of the Brain) (to M

Dichgans); the e:Med program (e:AtheroSysMed) (to M Dichgans) and the FP7/2007-2103 European Union project CVgenes@target (grant agreement number Health-F2-2013-601456) (to M Dichgans).

References

1. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, de Ferranti S, Despres JP, Fullerton HJ, Howard VJ, Huffman MD, Judd SE, Kissela BM, Lackland DT, Lichtman JH, Lisabeth LD, Liu S, Mackey RH, Matchar DB, McGuire DK, Mohler ER III, Moy CS, Muntner P, Mussolino ME, Nasir K, Neumar RW, Nichol G, Palaniappan L, Pandey DK, Reeves MJ, Rodriguez CJ, Sorlie PD, Stein J, Towfighi A, Turan TN, Virani SS, Willey JZ, Woo D, Yeh RW, Turner MB, American Heart Association Statistics Committee, Stroke Statistics Subcommittee. Heart disease and stroke statistics–2015 update: a report from the American Heart Association. *Circulation* 2015;**131**:e29–322.
2. Adams HP Jr, Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, Marsh EE III. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke* 1993;**24**:35–41.
3. Ikram MA, Wieberdink RG, Koudstaal PJ. International epidemiology of intracerebral hemorrhage. *Curr Atheroscler Rep* 2012;**14**:300–306.
4. Kim BJ, Kim JS. Ischemic stroke subtype classification: an asian viewpoint. *J Stroke* 2014;**16**:8–17.
5. Ay H, Benner T, Arsava EM, Furie KL, Singhal AB, Jensen MB, Ayata C, Towfighi A, Smith EE, Chong JY, Koroshetz WJ, Sorensen AG. A computerized algorithm for etiologic classification of ischemic stroke: the Causative Classification of Stroke System. *Stroke* 2007;**38**:2979–2984.
6. Desai JA, Abuzinadah AR, Imoukhuede O, Bernbaum ML, Modi J, Demchuk AM, Coutts SB. Etiologic classification of TIA and minor stroke by A-S-C-O and causative classification system as compared to TOAST reduces the proportion of patients categorized as cause undetermined. *Cerebrovasc Dis* 2014;**38**:121–126.
7. Flossmann E, Schulz UG, Rothwell PM. Systematic review of methods and results of studies of the genetic epidemiology of ischemic stroke. *Stroke* 2004;**35**:212–227.
8. de Faire U, Friberg L, Lundman T. Concordance for mortality with special reference to ischaemic heart disease and cerebrovascular disease. A study on the Swedish Twin Registry. *Prev Med* 1975;**4**:509–517.
9. Holliday EG, Maguire JM, Evans TJ, Koblar SA, Jannes J, Sturm JW, Hankey GJ, Baker R, Golledge J, Parsons MW, Malik R, McEvoy M, Biros E, Lewis MD, Lincz LF, Peel R, Oldmeadow C, Smith W, Moscato P, Barlera S, Bevan S, Bis JC, Boerwinkle E, Boncoraglio GB, Brott TG, Brown RD Jr, Cheng YC, Cole JW, Cotlarciuc I, Devan WJ, Fornage M, Furie KL, Gretarsdottir S, Gschwendtner A, Ikram MA, Longstreth WT Jr, Meschia JF, Mitchell BD, Mosley TH, Nalls MA, Parati EA, Psaty BM, Sharma P, Stefansson K, Thorleifsson G, Thorsteinsdottir U, Traylor M, Verhaaren BF, Wiggins KL, Worrall BB, Sudlow C, Rothwell PM, Farrall M, Dichgans M, Rosand J, Markus HS, Scott RJ, Levi C, Attia J, Australian Stroke Genetics C, International Stroke Genetics C, Wellcome Trust Case Control C. Common variants at 6p21.1 are associated with large artery atherosclerotic stroke. *Nat Genet* 2012;**44**:1147–1151.
10. Bevan S, Traylor M, Adib-Samii P, Malik R, Paul NL, Jackson C, Farrall M, Rothwell PM, Sudlow C, Dichgans M, Markus HS. Genetic heritability of ischemic stroke and the contribution of previously reported candidate gene and genomewide associations. *Stroke* 2012;**43**:3161–3167.
11. Lee SH, Wray NR, Goddard ME, Visscher PM. Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet* 2011;**88**:294–305.
12. International Stroke Genetics Consortium (ISGC), Wellcome Trust Case Control Consortium, Bellenguez C, Bevan S, Gschwendtner A, Spencer CC, Burgess AI, Pirinen M, Jackson CA, Traylor M, Strange A, Su Z, Band G, Syme PD, Malik R, Pera J, Norrving B, Lemmens R, Freeman C, Schanz R, James T, Poole D, Murphy L, Segal H, Cortellini L, Cheng YC, Woo D, Nalls MA, Muller-Myhsok B, Meisinger C, Seedorf U, Ross-Adams H, Boonen S, Wloch-Kopec D, Valant V, Slark J, Furie K, Delavaran H, Langford C, Deloukas P, Edkins S, Hunt S, Gray E, Dronov S, Peltonen L, Gretarsdottir S, Thorleifsson G, Thorsteinsdottir U, Stefansson K, Boncoraglio GB, Parati EA, Attia J, Holliday E, Levi C, Franzosi MG, Goel A, Helgadottir A, Blackwell JM, Bramon E, Brown MA, Casas JP, Corvin A, Duncanson A, Jankowski J, Mathew CG, Palmer CN, Plomin R, Rautanen A, Sawcer SJ, Trembath RC, Viswanathan AC, Wood NW, Worrall BB, Kittner SJ, Mitchell BD, Kissela B, Meschia JF, Thijs V, Lindgren A, Macleod MJ, Slowik A, Walters M, Rosand J, Sharma P, Farrall M, Sudlow CL, Rothwell PM, Dichgans M, Donnelly P, Markus HS. Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. *Nat Genet* 2012;**44**:328–333.
13. Malik R, Traylor M, Pulit SL, Bevan S, Hopewell JC, Holliday EG, Zhao W, Abrantes P, Amouyel P, Attia JR, Battey TW, Berger K, Boncoraglio GB, Chauhan G, Cheng YC, Chen WM, Clarke R, Cotlarciuc I, Debette S, Falcone GJ, Ferro JM, Gamble DM, Ilinca A, Kittner SJ, Kourkoulis CE, Lemmens R, Levi CR, Lichtner P, Lindgren A, Liu J, Meschia JF, Mitchell BD, Oliveira SA, Pera J, Reiner AP, Rothwell PM, Sharma P, Slowik A, Sudlow CL, Tatlisumak T, Thijs V, Vicente AM, Woo D, Seshadri S, Saleheen D, Rosand J, Markus HS, Worrall BB, Dichgans M, Group IA, Collaboration M, Wellcome Trust Case Control C, Network NSG. Low-frequency and common genetic variation in ischemic stroke: the METASTROKE collaboration. *Neurology* 2016;**86**:1217–1226.
14. Malik R, Dau T, Gonik M, Sivakumar A, Deredge DJ, Edeleva EV, Gotzfried J, van der Laan SW, Pasterkamp G, Beaufort N, Seixas S, Bevan S, Lincz LF, Holliday EG, Burgess AI, Rannikmae K, Minnerup J, Kriebel J, Waldenberger M, Muller-Nurasyid M, Lichtner P, Saleheen D, International Stroke Genetics C, Rothwell PM, Levi C, Attia J, Sudlow CL, Braun D, Markus HS, Wintrode PL, Berger K, Jenne DE, Dichgans M. Common coding variant in SERPINA1 increases the risk for large artery stroke. *Proc Natl Acad Sci U S A* 2017;**114**:3613–3618.
15. NINDS Stroke Genetics Network (SiGN), International Stroke Genetics Consortium (ISGC). Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study. *Lancet Neurol* 2015;**15**:174–184.
16. Gudbjartsson DF, Arnar DO, Helgadottir A, Gretarsdottir S, Holm H, Sigurdsson A, Jonasdottir A, Baker A, Thorleifsson G, Kristjansson K, Palsson A, Blondal T, Sulem P, Backman VM, Hardarson GA, Palsdottir E, Helgason A, Sigurjonsdottir R, Sverrisson JT, Kostulas K, Ng MC, Baum L, So WY, Wong KS, Chan JC, Furie KL, Greenberg SM, Sale M, Kelly P, MacRae CA, Smith EE, Rosand J, Hillert J, Ma RC, Ellinor PT, Thorgeirsson G, Gulcher JR, Kong A, Thorsteinsdottir U, Stefansson K. Variants conferring risk of atrial fibrillation on chromosome 4q25. *Nature* 2007;**448**:353–357.
17. Gudbjartsson DF, Holm H, Gretarsdottir S, Thorleifsson G, Walters GB, Thorgeirsson G, Gulcher J, Mathiesen EB, Njolstad I, Nyrnes A, Wilsgaard T, Hald EM, Hveem K, Stoltenberg C, Kucera G, Stubblefield T, Carter S, Roden D, Ng MC, Baum L, So WY, Wong KS, Chan JC, Gieger C, Wichmann HE, Gschwendtner A, Dichgans M, Kuhlenbaumer G, Berger K, Ringelstein EB, Bevan S, Markus HS, Kostulas K, Hillert J, Sveinbjornsdottir S, Valdimarsson EM, Lochen ML, Ma RC, Darbar D, Kong A, Arnar DO, Thorsteinsdottir U, Stefansson K. A sequence variant in ZFHX3 on 16q22 associates with atrial fibrillation and ischemic stroke. *Nat Genet* 2009;**41**:876–878.
18. Traylor M, Malik R, Nalls MA, Cotlarciuc I, Radmanesh F, Thorleifsson G, Hanscombe KB, Langefeld C, Saleheen D, Rost NS, Yet I, Spector TD, Bell JT, Hannon E, Mill J, Chauhan G, Debette S, Bis JC, Longstreth WT, Ikram MA, Launer LJ, Seshadri S, Hamilton-Bruce MA, Jimenez-Conde J, Cole JW, Schmidt R, Słowik A, Lemmens R, Lindgren A, Melander O, Grewal RP, Sacco RL, Rundek T, Rexrode K, Arnett DK, Johnson JA, Benavente OR, Wasssertheil-Smoller S, Lee J-M, Pulit SL, Wong Q, Rich SS, de Bakker PIW, McArdle PF, Woo D, Anderson CD, Xu H, Heitsch L, Fornage M, Jern C, Stefansson K, Thorsteinsdottir U, Gretarsdottir S, Lewis CM, Sharma P, Sudlow CLM, Rothwell PM, Boncoraglio GB, Thijs V, Levi C, Meschia JF, Rosand J, Kittner SJ, Mitchell BD, Dichgans M, Worrall BB, Markus HS, International Stroke Genetics C. Genetic Variation at 16q24.2 is associated with small vessel stroke. *Ann Neurol* 2017;**81**:383–394.
19. Kilarski LL, Achterberg S, Devan WJ, Traylor M, Malik R, Lindgren A, Pare G, Sharma P, Slowik A, Thijs V, Walters M, Worrall BB, Sale MM, Algra A, Kappelle LJ, Wijmenga C, Norrving B, Sandling JK, Ronnblom L, Goris A, Franke A, Sudlow C, Rothwell PM, Levi C, Holliday EG, Fornage M, Psaty B, Gretarsdottir S, Thorsteinsdottir U, Seshadri S, Mitchell BD, Kittner S, Clarke R, Hopewell JC, Bis JC, Boncoraglio GB, Meschia J, Ikram MA, Hansen BM, Montaner J, Thorleifsson G, Stefanson K, Rosand J, de Bakker PI, Farrall M, Dichgans M, Markus HS, Bevan S, Garnet Collaborative Research Group WTCCCASGCtMC, the International Stroke Genetics C. Meta-analysis in more than 17, 900 cases of ischemic stroke reveals a novel association at 12q24.12. *Neurology* 2014;**83**:678–685.
20. Woo D, Falcone GJ, Devan WJ, Brown WM, Biffi A, Howard TD, Anderson CD, Brouwers HB, Valant V, Battey TW, Radmanesh F, Raffeld MR, Baedorf-Kassis S, Deka R, Woo JG, Martin LJ, Haverbusch M, Moomaw CJ, Sun G, Broderick JP, Flaherty ML, Martini SR, Kleindorfer DO, Kissela B, Comeau ME, Jagiella JM, Schmidt H, Freudenberger P, Pichler A, Enzinger C, Hansen BM, Norrving B, Jimenez-Conde J, Giralt-Steinhauer E, Elosua R, Cuadrado-Godia E, Soriano C, Roquer J, Kraft P, Ayres AM, Schwab K, McCauley JL, Pera J, Urbanik A, Rost NS, Goldstein JN, Viswanathan A, Stogerer EM, Tirschwell DL, Selim M, Brown DL, Silliman SL, Worrall BB, Meschia JF, Kidwell CS, Montaner J, Fernandez-Cadenas I, Delgado P, Malik R, Dichgans M, Greenberg SM, Rothwell PM, Lindgren A, Slowik A, Schmidt R, Langefeld CD, Rosand J, International Stroke Genetics C. Meta-analysis of genome-wide association studies identifies 1q22 as a susceptibility locus for intracerebral hemorrhage. *Am J Hum Genet* 2014;**94**:511–521.
21. Ikram MA, Seshadri S, Bis JC, Fornage M, DeStefano AL, Aulchenko YS, Debette S, Lumley T, Folsom AR, van den Herik EG, Bos MJ, Beiser A, Cushman M, Launer LJ, Shahar E, Struchalin M, Du Y, Glazer NL, Rosamond WD, Rivadeneira F, Kelly-Hayes M, Lopez OL, Coresh J, Hofman A, DeCarli C, Heckbert SR, Koudstaal PJ, Yang Q, Smith NL, Kase CS, Rice K, Haritunians T, Roks G, de Kort PL, Taylor KD, de Lau LM, Oostra BA, Uitterlinden AG, Rotter JI, Boerwinkle E, Psaty BM, Mosley TH, van Duijn CM, Breteler MM, Longstreth WT, Jr, Wolf, PA. Genomewide association studies of stroke. *N Engl J Med* 2009;**360**:1718–1728.
22. Webb TR, Erdmann J, Stirrups KE, Stitziel NO, Masca NG, Jansen H, Kanoni S, Nelson CP, Ferrario PG, Konig IR, Eicher JD, Johnson AD, Hamby SE, Betsholtz C, Ruusalepp A, Franzen O, Schadt EE, Bjorkegren JL, Weeke PE, Auer PL, Schick UM, Lu Y, Zhang H, Dube MP, Goel A, Farrall M, Peloso GM, Won HH, Do R, van Iperen E, Kruppa J, Mahajan A, Scott RA, Willenborg C, Braund PS, van Capelleveen JC, Doney AS, Donnelly LA, Asselta R, Merlini PA, Duga S, Marziliano N, Denny JC,

Shaffer C, El-Mokhtari NE, Franke A, Heilmann S, Hengstenberg C, Hoffmann P, Holmen OL, Hveem K, Jansson JH, Jockel KH, Kessler T, Kriebel J, Laugwitz KL, Marouli E, Martinelli N, McCarthy MI, Van Zuydam NR, Meisinger C, Esko T, Mihailov E, Escher SA, Alver M, Moebus S, Morris AD, Virtamo J, Nikpay M, Olivieri O, Provost S, AlQarawi A, Robertson NR, Akinsansya KO, Reilly DF, Vogt TF, Yin W, Asselbergs FW, Kooperberg C, Jackson RD, Stahl E, Muller-Nurasyid M, Strauch K, Varga TV, Waldenberger M, Wellcome Trust Case Control C, Zeng L, Chowdhury R, Salomaa V, Ford I, Jukema JW, Amouyel P, Kontto J, Investigators M, Nordestgaard BG, Ferrieres J, Saleheen D, Sattar N, Surendran P, Wagner A, Young R, Howson JM, Butterworth AS, Danesh J, Ardissino D, Bottinger EP, Erbel R, Franks PW, Girelli D, Hall AS, Hovingh GK, Kastrati A, Lieb W, Meitinger T, Kraus WE, Shah SH, McPherson R, Orho-Melander M, Melander O, Metspalu A, Palmer CN, Peters A, Rader DJ, Reilly MP, Loos RJ, Reiner AP, Roden DM, Tardif JC, Thompson JR, Wareham NJ, Watkins H, Willer CJ, Samani NJ, Schunkert H, Deloukas P, Kathiresan S, Myocardial Infarction G, Investigators CAEC. Systematic evaluation of pleiotropy identifies 6 further loci associated with coronary artery disease. *J Am Coll Cardiol* 2017;**69**:823–836.

23. Wang W, Tang Y, Wang Y, Tascau L, Balcerek J, Tong W, Levine RL, Welch C, Tall AR, Wang N. LNK/SH2B3 loss of function promotes atherosclerosis and thrombosis. *Circ Res* 2016;**119**:e91–e103.
24. Larionov S, Dedeck O, Birkenmeier G, Thal DR. Expression of alpha2-macroglobulin, neutrophil elastase, and interleukin-1alpha differs in early-stage and late-stage atherosclerotic lesions in the arteries of the circle of Willis. *Acta Neuropathol* 2007;**113**:33–43.
25. Gordon SM, McKenzie B, Kemeh G, Sampson M, Perl S, Young NS, Fessler MB, Remaley AT. Rosuvastatin alters the proteome of high density lipoproteins: generation of alpha-1-antitrypsin enriched particles with anti-inflammatory properties. *Mol Cell Proteomics* 2015;**14**:3247–3257.
26. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP, Hopewell JC, Webb TR, Zeng L, Dehghan A, Alver M, Armasu SM, Auro K, Bjonnes A, Chasman DI, Chen S, Ford I, Franceschini N, Gieger C, Grace C, Gustafsson S, Huang J, Hwang SJ, Kim YK, Kleber ME, Lau KW, Lu X, Lu Y, Lyytikainen LP, Mihailov E, Morrison AC, Pervjakova N, Qu L, Rose LM, Salfati E, Saxena R, Scholz M, Smith AV, Tikkanen E, Uitterlinden A, Yang X, Zhang W, Zhao W, de Andrade M, de Vries PS, van Zuydam NR, Anand SS, Bertram L, Beutner F, Dedoussis G, Frossard P, Gauguier D, Goodall AH, Gottesman O, Haber M, Han BG, Huang J, Jalilzadeh S, Kessler T, Konig IR, Lannfelt L, Lieb W, Lind L, Lindgren CM, Lokki ML, Magnusson PK, Mallick NH, Mehra N, Meitinger T, Memon FU, Morris AP, Nieminen MS, Pedersen NL, Peters A, Rallidis LS, Rasheed A, Samuel M, Shah SH, Sinisalo J, Stirrups KE, Trompet S, Wang L, Zaman KS, Ardissino D, Boerwinkle E, Borecki IB, Bottinger EP, Buring JE, Chambers JC, Collins R, Cupples LA, Danesh J, Demuth I, Elosua R, Epstein SE, Esko T, Feitosa MF, Franco OH, Franzosi MG, Granger CB, Gu D, Gudnason V, Hall AS, Hamsten A, Harris TB, Hazen SL, Hengstenberg C, Hofman A, Ingelsson E, Iribarren C, Jukema JW, Karhunen PJ, Kim BJ, Kooner JS, Kullo IJ, Lehtimaki T, Loos RJ, Melander O, Metspalu A, Marz W, Palmer CN, Perola M, Quertermous T, Rader DJ, Ridker PM, Ripatti S, Roberts R, Salomaa V, Sanghera DK, Schwartz SM, Seedorf U, Stewart AF, Stott DJ, Thiery J, Zalloua PA, O'Donnell CJ, Reilly MP, Assimes TL, Thompson JR, Erdmann J, Clarke R, Watkins H, Kathiresan S, McPherson R, Deloukas P, Schunkert H, Samani NJ, Farrall M, Consortium CAD. A comprehensive 1, 000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* 2015;**47**:1121–1130.
27. Foroud T, Lai D, Koller D, Van't Hof F, Kurki MI, Anderson CS, Brown RD, Connolly ES, Eriksson JG, Flaherty M, Fornage M, von und zu Fraunberg M, Gaál EI, Laakso A, Hernesniemi J, Huston J, Jääskeläinen JE, Kiemeney LA, Kivisaari R, Kleindorfer D, Ko N, Lehto H, Mackey J, Meissner I, Moomaw CJ, Mosley TH, Moskala M, Niemelä M, Palotie A, Pera J, Rinkel G, Ripke S, Rouleau G, Ruigrok Y, Sauerbeck L, Słowik A, Vermeulen SH, Woo D, Worrall BB, Broderick J, Familial Intracranial Aneurysm Study I. Genome-wide association study of intracranial aneurysm identifies a new association on chromosome 7. *Stroke* 2014;**45**:3194–3199.
28. Matsukura M, Ozaki K, Takahashi A, Onouchi Y, Morizono T, Komai H, Shigematsu H, Kudo T, Inoue Y, Kimura H, Hosaka A, Shigematsu K, Miyata T, Watanabe T, Tsunoda T, Kubo M, Tanaka T. Genome-wide association study of peripheral arterial disease in a Japanese population. *PLoS One* 2015;**10**:e0139262.
29. Azghandi S, Prell C, van der Laan SW, Schneider M, Malik R, Berer K, Gerdes N, Pasterkamp G, Weber C, Haffner C, Dichgans M. Deficiency of the stroke relevant HDAC9 gene attenuates atherosclerosis in accord with allele-specific effects at 7p21.1. *Stroke* 2015;**46**:197–202.
30. Cao Q, Rong S, Repa JJ, Clair RS, Parks JS, Mishra N. Histone deacetylase 9 represses cholesterol efflux and alternatively activated macrophages in atherosclerosis development. *Arterioscler Thromb Vasc Biol* 2014;**34**:1871–1879.
31. Lobera M, Madauss KP, Pohlhaus DT, Wright QG, Trocha M, Schmidt DR, Baloglu E, Trump RP, Head MS, Hofmann GA, Murray-Thompson M, Schwartz B, Chakravorty S, Wu Z, Mander PK, Kruidenier L, Reid RA, Burkhart W, Turunen BJ, Rong JX, Wagner C, Moyer MB, Wells C, Hong X, Moore JT, Williams JD, Soler D, Ghosh S, Nolan MA. Selective class IIa histone deacetylase inhibition via a nonchelating zinc-binding group. *Nat Chem Biol* 2013;**9**:319–325.
32. Neurology Working Group of the Cohorts for Heart Aging Research in Genomic Epidemiology (CHARGE) Consortium, Stroke Genetics Network (SiGN), International Stroke Genetics Consortium (ISGC). Identification of additional risk loci for stroke and small vessel disease: a meta-analysis of genome-wide association studies. *Lancet Neurol* 2016;**15**:695–707.
33. Reyahi A, Nik AM, Ghiami M, Gritli-Linde A, Ponten F, Johansson BR, Carlsson P. Foxf2 Is required for brain pericyte differentiation and development and maintenance of the blood-brain barrier. *Dev Cell* 2015;**34**:19–32.
34. Ghosh M, Balbi M, Hellal F, Dichgans M, Lindauer U, Plesnila N. Pericytes are involved in the pathogenesis of cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy. *Ann Neurol* 2015;**78**:887–900.
35. Beaufort N, Scharrer E, Kremmer E, Lux V, Ehrmann M, Huber R, Houlden H, Werring D, Haffner C, Dichgans M. Cerebral small vessel disease-related protease HtrA1 processes latent TGF-beta binding protein 1 and facilitates TGF-beta signaling. *Proc Natl Acad Sci U S A* 2014;**111**:16496–16501.
36. GTEx Consortium. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 2015;**348**:648–660.
37. Eicher JD, Landowski C, Stackhouse B, Sloan A, Chen W, Jensen N, Lien JP, Leslie R, Johnson AD. GRASP v2.0: an update on the genome-wide repository of associations between SNPs and phenotypes. *Nucleic Acids Res* 2015;**43**:D799–D804.
38. Higasa K, Miyake N, Yoshimura J, Okamura K, Niihori T, Saitsu H, Doi K, Shimizu M, Nakabayashi K, Aoki Y, Tsurusaki Y, Morishita S, Kawaguchi T, Migita O, Nakayama K, Nakashima M, Mitsui J, Narahara M, Hayashi K, Funayama R, Yamaguchi D, Ishiura H, Ko W-Y, Hata K, Nagashima T, Yamada R, Matsubara Y, Umezawa A, Tsuji S, Matsumoto N, Matsuda F. Human genetic variation database, a reference database of genetic variations in the Japanese population. *J Hum Genet* 2016;**61**:547–553.
39. Bonder MJ, Luijk R, Zhernakova DV, Moed M, Deelen P, Vermaat M, van Iterson M, van Dijk F, van Galen M, Bot J, Slieker RC, Jhamai PM, Verbiest M, Suchiman HE, Verkerk M, van der Breggen R, van Rooij J, Lakenberg N, Arindrarto W, Kielbasa SM, Jonkers I, van 't Hof P, Nooren I, Beekman M, Deelen J, van Heemst D, Zhernakova A, Tigchelaar EF, Swertz MA, Hofman A, Uitterlinden AG, Pool R, van Dongen J, Hottenga JJ, Stehouwer CD, van der Kallen CJ, Schalkwijk CG, van den Berg LH, van Zwet EW, Mei H, Li Y, Lemire M, Hudson TJ, Consortium B, Slagboom PE, Wijmenga C, Veldink JH, van Greevenbroek MM, van Duijn CM, Boomsma DI, Isaacs A, Jansen R, van Meurs JB, T Hoen PA, Franke L, Heijmans BT. Disease variants alter transcription factor levels and methylation of their binding sites. *Nat Genet* 2017;**49**:131–138.
40. Adams D, Altucci L, Antonarakis SE, Ballesteros J, Beck S, Bird A, Bock C, Boehm B, Campo E, Caricasole A, Dahl F, Dermitzakis ET, Enver T, Esteller M, Estivill X, Ferguson-Smith A, Fitzgibbon J, Flicek P, Giehl C, Graf T, Grosveld F, Guigo R, Gut I, Helin K, Jarvius J, Kuppers R, Lehrach H, Lengauer T, Lernmark A, Leslie D, Loeffler M, Macintyre E, Mai A, Martens JH, Minucci S, Ouwehand WH, Pelicci PG, Pendeville H, Porse B, Rakyan V, Reik W, Schrappe M, Schubeler D, Seifert M, Siebert R, Simmons D, Soranzo N, Spicuglia S, Stratton M, Stunnenberg HG, Tanay A, Torrents D, Valencia A, Vellenga E, Vingron M, Walter J, Willcocks S. BLUEPRINT to decode the epigenetic signature written in blood. *Nat Biotechnol* 2012;**30**:224–226.
41. Franzén O, Ermel R, Cohain A, Akers NK, Di Narzo A, Talukdar HA, Foroughi-Asl H, Giambartolomei C, Fullard JF, Sukhavasi K, Köks S, Gan L-M, Giannarelli C, Kovacic JC, Betsholtz C, Losic B, Michoel T, Hao K, Roussos P, Skogsberg J, Ruusalepp A, Schadt EE, Björkegren JLM. Cardiometabolic risk loci share downstream cis- and trans-gene regulation across tissues and diseases. *Science* 2016;**353**:827–830.
42. Erbilgin A, Civelek M, Romanoski CE, Pan C, Hagopian R, Berliner JA, Lusis AJ. Identification of CAD candidate genes in GWAS loci and their expression in vascular cells. *J Lipid Res* 2013;**54**:1894–1905.
43. Bennett DA, Schneider JA, Arvanitakis Z, Wilson RS. Overview and findings from the religious orders study. *Curr Alzheimer Res* 2012;**9**:628–645.
44. Bennett DA, Schneider JA, Buchman AS, Barnes LL, Boyle PA, Wilson RS. Overview and findings from the rush Memory and Aging Project. *Car* 2012;**9**:646–663.
45. Mancuso N, Shi H, Goddard P, Kichaev G, Gusev A, Pasaniuc B. Integrating gene expression with summary association statistics to identify genes associated with 30 complex traits. *Am J Hum Genet* 2017;**100**:473–487.
46. Suhre K, Arnold M, Bhagwat AM, Cotton RJ, Engelke R, Raffler J, Sarwath H, Thareja G, Wahl A, DeLisle RK, Gold L, Pezer M, Lauc G, El-Din Selim MA, Mook-Kanamori DO, Al-Dous EK, Mohamoud YA, Malek J, Strauch K, Grallert H, Peters A, Kastenmüller G, Gieger C, Graumann J. Connecting genetic risk to disease end points through the human blood plasma proteome. *Nat Commun* 2017;**8**:14357.
47. Assarsson E, Lundberg M, Holmquist G, Bjorkesten J, Thorsen SB, Ekman D, Eriksson A, Rennel Dickens E, Ohlsson S, Edfeldt G, Andersson AC, Lindstedt P, Stenvang J, Gullberg M, Fredriksson S. Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. *PLoS One* 2014;**9**:e95192.
48. Di Narzo AF, Telesco SE, Brodmerkel C, Argmann C, Peters LA, Li K, Kidd B, Dudley J, Cho J, Schadt EE, Kasarskis A, Dobrin R, Hao K. High-throughput characterization of blood serum proteomics of IBD patients with respect to aging and genetic factors. *PLoS Genet* 2017;**13**:e1006565.
49. Sattlecker M, Kiddle SJ, Newhouse S, Proitsi P, Nelson S, Williams S, Johnston C, Killick R, Simmons A, Westman E, Hodges A, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Lovestone S, AddNeuroMed C, Dobson RJ. Alzheimer's disease

biomarker discovery using SOMAscan multiplexed protein technology. *Alzheimers Dement* 2014;**10**:724–734.
50. Raina A, Zhao X, Grove ML, Bressler J, Gottesman RF, Guan W, Pankow JS, Boerwinkle E, Mosley TH, Fornage M. Cerebral white matter hyperintensities on MRI and acceleration of epigenetic aging: the atherosclerosis risk in communities study. *Clin Epigenet* 2017;**9**:21.
51. Soriano-Tárraga C, Mola-Caminal M, Giralt-Steinhauer E, Ois A, Rodríguez-Campello A, Cuadrado-Godia E, Gómez-González A, Vivanco-Hidalgo RM, Fernández-Cadenas I, Cullell N, Roquer J, Jiménez-Conde J. Biological age is better than chronological as predictor of 3-month outcome in ischemic stroke. *Neurology* 2017;**89**:830–836.
52. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, Reynolds AP, Sandstrom R, Qu H, Brody J, Shafer A, Neri F, Lee K, Kutyavin T, Stehling-Sun S, Johnson AK, Canfield TK, Giste E, Diegel M, Bates D, Hansen RS, Neph S, Sabo PJ, Heimfeld S, Raubitschek A, Ziegler S, Cotsapas C, Sotoodehnia N, Glass I, Sunyaev SR, Kaul R, Stamatoyannopoulos JA. Systematic localization of common disease-associated variation in regulatory DNA. *Science* 2012;**337**:1190–1195.
53. Trynka G, Sandor C, Han B, Xu H, Stranger BE, Liu XS, Raychaudhuri S. Chromatin marks identify critical cell types for fine mapping complex trait variants. *Nat Genet* 2013;**45**:124–130.
54. Encode Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012;**489**:57–74.
55. Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, Ziller MJ, Amin V, Whitaker JW, Schultz MD, Ward LD, Sarkar A, Quon G, Sandstrom RS, Eaton ML, Wu Y-C, Pfenning AR, Wang X, Claussnitzer M, Liu Y, Coarfa C, Harris RA, Shoresh N, Epstein CB, Gjoneska E, Leung D, Xie W, Hawkins RD, Lister R, Hong C, Gascard P, Mungall AJ, Moore R, Chuah E, Tam A, Canfield TK, Hansen RS, Kaul R, Sabo PJ, Bansal MS, Carles A, Dixon JR, Farh K-H, Feizi S, Karlic R, Kim A-R, Kulkarni A, Li D, Lowdon R, Elliott GNell, Mercer TR, Neph SJ, Onuchic V, Polak P, Rajagopal N, Ray P, Sallari RC, Siebenthall KT, Sinnott-Armstrong NA, Stevens M, Thurman RE, Wu J, Zhang B, Zhou X, Beaudet AE, Boyer LA, De Jager PL, Farnham PJ, Fisher SJ, Haussler D, Jones SJM, Li W, Marra MA, McManus MT, Sunyaev S, Thomson JA, Tlsty TD, Tsai L-H, Wang W, Waterland RA, Zhang MQ, Chadwick LH, Bernstein BE, Costello JF, Ecker JR, Hirst M, Meissner A, Milosavljevic A, Ren B, Stamatoyannopoulos JA, Wang T, Kellis M. Integrative analysis of 111 reference human epigenomes. *Nature* 2015;**518**:317–330.
56. Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, Kochi Y, Ohmura K, Suzuki A, Yoshida S, Graham RR, Manoharan A, Ortmann W, Bhangale T, Denny JC, Carroll RJ, Eyler AE, Greenberg JD, Kremer JM, Pappas DA, Jiang L, Yin J, Ye L, Su DF, Yang J, Xie G, Keystone E, Westra HJ, Esko T, Metspalu A, Zhou X, Gupta N, Mirel D, Stahl EA, Diogo D, Cui J, Liao K, Guo MH, Myouzen K, Kawaguchi T, Coenen MJ, van Riel PL, van de Laar MA, Guchelaar HJ, Huizinga TW, Dieude P, Mariette X, Bridges SL Jr, Zhernakova A, Toes RE, Tak PP, Miceli-Richard C, Bang SY, Lee HS, Martin J, Gonzalez-Gay MA, Rodriguez-Rodriguez L, Rantapaa-Dahlqvist S, Arlestig L, Choi HK, Kamatani Y, Galan P, Lathrop M, Consortium R, Consortium G, Eyre S, Bowes J, Barton A, de Vries N, Moreland LW, Criswell LA, Karlson EW, Taniguchi A, Yamada R, Kubo M, Liu JS, Bae SC, Worthington J, Padyukov L, Klareskog L, Gregersen PK, Raychaudhuri S, Stranger BE, De Jager PL, Franke L, Visscher PM, Brown MA, Yamanaka H, Mimori T, Takahashi A, Xu H, Behrens TW, Siminovitch KA, Momohara S, Matsuda F, Yamamoto K, Plenge RM. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* 2014;**506**:376–381.
57. Brænne I, Civelek M, Vilne B, Di Narzo A, Johnson AD, Zhao Y, Reiz B, Codoni V, Webb TR, Foroughi Asl H, Hamby SE, Zeng L, Trégouët D-A, Hao K, Topol EJ, Schadt EE, Yang X, Samani NJ, Björkegren JLM, Erdmann J, Schunkert H, Lusis AJ, Leducq Consortium CADGd. Prediction of causal candidate genes in coronary artery disease loci. *Arterioscler Thromb Vasc Biol* 2015;**35**:2207–2217.
58. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, Tukiainen T, Birnbaum DP, Kosmicki JA, Duncan LE, Estrada K, Zhao F, Zou J, Pierce-Hoffman E, Berghout J, Cooper DN, Deflaux N, DePristo M, Do R, Flannick J, Fromer M, Gauthier L, Goldstein J, Gupta N, Howrigan D, Kiezun A, Kurki MI, Moonshine AL, Natarajan P, Orozco L, Peloso GM, Poplin R, Rivas MA, Ruano-Rubio V, Rose SA, Ruderfer DM, Shakir K, Stenson PD, Stevens C, Thomas BP, Tiao G, Tusie-Luna MT, Weisburd B, Won HH, Yu D, Altshuler DM, Ardissino D, Boehnke M, Danesh J, Donnelly S, Elosua R, Florez JC, Gabriel SB, Getz G, Glatt SJ, Hultman CM, Kathiresan S, Laakso M, McCarroll S, McCarthy MI, McGovern D, McPherson R, Neale BM, Palotie A, Purcell SM, Saleheen D, Scharf JM, Sklar P, Sullivan PF, Tuomilehto J, Tsuang MT, Watkins HC, Wilson JG, Daly MJ, MacArthur DG, Exome Aggregation C. Analysis of protein-coding genetic variation in 60, 706 humans. *Nature* 2016;**536**:285–291.
59. Gudbjartsson DF, Helgason H, Gudjonsson SA, Zink F, Oddson A, Gylfason A, Besenbacher S, Magnusson G, Halldorsson BV, Hjartarson E, Sigurdsson GT, Stacey SN, Frigge ML, Holm H, Saemundsdottir J, Helgadottir HT, Johannsdottir H, Sigfusson G, Thorgeirsson G, Sverrisson JT, Gretarsdottir S, Walters GB, Rafnar T, Thjodleifsson B, Bjornsson ES, Olafsson S, Thorarinsdottir H, Steingrimsdottir T, Gudmundsdottir TS, Theodors A, Jonasson JG, Sigurdsson A, Bjornsdottir G, Jonsson JJ, Thorarensen O, Ludvigsson P, Gudbjartsson H, Eyjolfsson GI, Sigurdardottir O, Olafsson I, Arnar DO, Magnusson OT, Kong A, Masson G, Thorsteinsdottir U, Helgason A, Sulem P, Stefansson K. Large-scale whole-genome sequencing of the Icelandic population. *Nat Genet* 2015;**47**:435–444.
60. Melnikov A, Murugan A, Zhang X, Tesileanu T, Wang L, Rogov P, Feizi S, Gnirke A, Callan CG Jr, Kinney JB, Kellis M, Lander ES, Mikkelsen TS. Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nat Biotechnol* 2012;**30**:271–277.
61. Patwardhan RP, Hiatt JB, Witten DM, Kim MJ, Smith RP, May D, Lee C, Andrie JM, Lee SI, Cooper GM, Ahituv N, Pennacchio LA, Shendure J. Massively parallel functional dissection of mammalian enhancers in vivo. *Nat Biotechnol* 2012;**30**:265–270.
62. Butter F, Davison L, Viturawong T, Scheibe M, Vermeulen M, Todd JA, Mann M. Proteome-wide analysis of disease-associated SNPs that show allele-specific transcription factor binding. *PLoS Genet* 2012;**8**:e1002982.
63. Haitjema S, Meddens CA, van der Laan SW, Kofink D, Harakalova M, Tragante V, Foroughi Asl H, van Setten J, Brandt MM, Bis JC, O'Donnell C, Cheng C, Hoefer IE, Waltenberger J, Biessen E, Jukema JW, Doevendans PAFM, Nieuwenhuis EES, Erdmann J, Björkegren JLM, Pasterkamp G, Asselbergs FW, den Ruijter HM, Mokry M. Additional candidate genes for human atherosclerotic disease identified through annotation based on chromatin organization. *Circ Cardiovasc Genet* 2017;**10**:e001664.
64. Dekker J, Marti-Renom MA, Mirny LA. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat Rev Genet* 2013;**14**:390–403.
65. Heinz S, Romanoski CE, Benner C, Glass CK. The selection and function of cell type-specific enhancers. *Nat Rev Mol Cell Biol* 2015;**16**:144–154.
66. Mercer TR, Edwards SL, Clark MB, Neph SJ, Wang H, Stergachis AB, John S, Sandstrom R, Li G, Sandhu KS, Ruan Y, Nielsen LK, Mattick JS, Stamatoyannopoulos JA. DNase I-hypersensitive exons colocalize with promoters and distal regulatory elements. *Nat Genet* 2013;**45**:852–859.
67. Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol* 2014;**32**:347–355.
68. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. RNA-guided human genome engineering via Cas9. *Science* 2013;**339**:823–826.
69. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 2014;**157**:1262–1278.
70. Paquet D, Kwart D, Chen A, Sproul A, Jacob S, Teo S, Olsen KM, Gregg A, Noggle S, Tessier-Lavigne M. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* 2016;**533**:125–129.
71. Gupta RM, Hadaya J, Trehan A, Zekavat SM, Roselli C, Klarin D, Emdin CA, Hilvering CRE, Bianchi V, Mueller C, Khera AV, Ryan RJH, Engreitz JM, Issner R, Shoresh N, Epstein CB, de Laat W, Brown JD, Schnabel RB, Bernstein BE, Kathiresan S. A genetic variant associated with five vascular diseases is a distal regulator of endothelin-1 gene expression. *Cell* 2017;**170**:522–533.
72. Jaffe AS, Babuin L, Apple FS. Biomarkers in acute cardiac disease: the present and the future. *J Am Coll Cardiol* 2006;**48**:1–11.
73. Agnelli G, Becattini C. Acute pulmonary embolism. *N Engl J Med* 2010;**363**:266–274.
74. Wells PS, Anderson DR, Rodger M, Forgie M, Kearon C, Dreyer J, Kovacs G, Mitchell M, Lewandowski B, Kovacs MJ. Evaluation of D-dimer in the diagnosis of suspected deep-vein thrombosis. *N Engl J Med* 2003;**349**:1227–1235.
75. Foerch C, Niessner M, Back T, Bauerle M, De Marchis GM, Ferbert A, Grehl H, Hamann GF, Jacobs A, Kastrup A, Klimpe S, Palm F, Thomalla G, Worthmann H, Sitzer M, Group BFS. Diagnostic accuracy of plasma glial fibrillary acidic protein for differentiating intracerebral hemorrhage and cerebral ischemia in patients with symptoms of acute stroke. *Clin Chem* 2012;**58**:237–245.
76. Dvorak F, Haberer I, Sitzer M, Foerch C. Characterisation of the diagnostic window of serum glial fibrillary acidic protein for the differentiation of intracerebral haemorrhage and ischaemic stroke. *Cerebrovasc Dis* 2009;**27**:37–41.
77. Foerch C, Curdt I, Yan B, Dvorak F, Hermans M, Berkefeld J, Raabe A, Neumann-Haefelin T, Steinmetz H, Sitzer M. Serum glial fibrillary acidic protein as a biomarker for intracerebral haemorrhage in patients with acute stroke. *J Neurol Neurosurg Psychiatry* 2006;**77**:181–184.
78. Tiedt S, Prestel M, Malik R, Schieferdecker N, During M, Kautzky V, Stoycheva I, Bock J, Northoff BH, Klein M, Dorn F, Krohn K, Teupser D, Liesz A, Plesnila N, Holdt LM, Dichgans M. RNA seq identifies circulating miR-125a-5p, miR-125b-5p and miR-143-3p as potential biomarkers for acute ischemic stroke. *Circ Res* 2017;**121**:970–980.
79. Yang H, Qin C, Li YH, Tao L, Zhou J, Yu CY, Xu F, Chen Z, Zhu F, Chen YZ. Therapeutic target database update 2016: enriched resource for bench to clinical drug target and targeted pathway information. *Nucleic Acids Res* 2016;**44**:D1069–D1074.
80. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, Chang Z, Woolsey J. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res* 2006;**34**:D668–D672.
81. Sanseau P, Agarwal P, Barnes MR, Pastinen T, Richards JB, Cardon LR, Mooser V. Use of genome-wide association studies for drug repositioning. *Nat Biotechnol* 2012;**30**:317–320.
82. Davey Smith G, Paternoster L, Relton C. When will mendelian randomization become relevant for clinical practice and public health? *Jama* 2017;**317**:589–591.

83. Hartwig FP, Davies NM, Hemani G, Davey Smith G. Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. *Int J Epidemiol* 2016;**45**:1717–1726.
84. Larsson SC, Scott RA, Traylor M, Langenberg CC, Hindy G, Melander O, Orho-Melander M, Seshadri S, Wareham NJ, Markus HS, Collaboration M, Network NSG. Type 2 diabetes, glucose, insulin, BMI, and ischemic stroke subtypes: mendelian randomization study. *Neurology* 2017;**89**:454–460.
85. Keenan T, Zhao W, Rasheed A, Ho WK, Malik R, Felix JF, Young R, Shah N, Samuel M, Sheikh N, Mucksavage ML, Shah O, Li J, Morley M, Laser A, Mallick NH, Zaman KS, Ishaq M, Rasheed SZ, Memon FU, Ahmed F, Hanif B, Lakhani MS, Fahim M, Ishaq M, Shardha NK, Ahmed N, Mahmood K, Iqbal W, Akhtar S, Raheel R, O'Donnell CJ, Hengstenberg C, Marz W, Kathiresan S, Samani N, Goel A, Hopewell JC, Chambers J, Cheng YC, Sharma P, Yang Q, Rosand J, Boncoraglio GB, Kazmi SU, Hakonarson H, Kottgen A, Kalogeropoulos A, Frossard P, Kamal A, Dichgans M, Cappola T, Reilly MP, Danesh J, Rader DJ, Voight BF, Saleheen D. Causal assessment of serum urate levels in cardiometabolic diseases through a mendelian randomization study. *J Am Coll Cardiol* 2016;**67**:407–416.
86. van der Laan SW, Fall T, Soumaré A, Teumer A, Sedaghat S, Baumert J, Zabaneh D, van Setten J, Isgum I, Galesloot TE, Arpegård J, Amouyel P, Trompet S, Waldenberger M, Dörr M, Magnusson PK, Giedraitis V, Larsson A, Morris AP, Felix JF, Morrison AC, Franceschini N, Bis JC, Kavousi M, O'Donnell C, Drenos F, Tragante V, Munroe PB, Malik R, Dichgans M, Worrall BB, Erdmann J, Nelson CP, Samani NJ, Schunkert H, Marchini J, Patel RS, Hingorani AD, Lind L, Pedersen NL, de Graaf J, Kiemeney LALM, Baumeister SE, Franco OH, Hofman A, Uitterlinden AG, Koenig W, Meisinger C, Peters A, Thorand B, Jukema JW, Eriksen BO, Toft I, Wilsgaard T, Onland-Moret NC, van der Schouw YT, Debette S, Kumari M, Svensson P, van der Harst P, Kivimaki M, Keating BJ, Sattar N, Dehghan A, Reiner AP, Ingelsson E, den Ruijter HM, de Bakker PIW, Pasterkamp G, Ärnlöv J, Holmes MV, Asselbergs FW. Cystatin C and cardiovascular disease: a Mendelian Randomization Study. *J Am Coll Cardiol* 2016;**68**:934–945.
87. Prins BP, Abbasi A, Wong A, Vaez A, Nolte I, Franceschini N, Stuart PE, Guterriez Achury J, Mistry V, Bradfield JP, Valdes AM, Bras J, Shatunov A, Consortium P, International Stroke Genetics C, Systemic Sclerosis C, Treat OAc Consortium D, Consortium CAD, Consortium ALS, International Parkinson's Disease Genomics C, Autism Spectrum Disorder Working Group Of The Psychiatric Genomics C, Consortium CK, Consortium G, International Consortium For Blood P, Schizophrenia Working Group Of The Psychiatric Genomics C, Inflammation Working Group Of The CC, Lu C, Han B, Raychaudhuri S, Bevan S, Mayes MD, Tsoi LC, Evangelou E, Nair RP, Grant SF, Polychronakos C, Radstake TR, van Heel DA, Dunstan ML, Wood NW, Al-Chalabi A, Dehghan A, Hakonarson H, Markus HS, Elder JT, Knight J, Arking DE, Spector TD, Koeleman BP, van Duijn CM, Martin J, Morris AP, Weersma RK, Wijmenga C, Munroe PB, Perry JR, Pouget JG, Jamshidi Y, Snieder H, Alizadeh BZ. Investigating the causal relationship of C-reactive protein with 32 complex somatic and psychiatric outcomes: a large-scale cross-consortium Mendelian Randomization Study. *PLoS Med* 2016;**13**:e1001976.
88. Bis JC, Kavousi M, Franceschini N, Isaacs A, Abecasis GR, Schminke U, Post WS, Smith AV, Cupples LA, Markus HS, Schmidt R, Huffman JE, Lehtimaki T, Baumert J, Munzel T, Heckbert SR, Dehghan A, North K, Oostra B, Bevan S, Stoegerer EM, Hayward C, Raitakari O, Meisinger C, Schillert A, Sanna S, Volzke H, Cheng YC, Thorsson B, Fox CS, Rice K, Rivadeneira F, Nambi V, Halperin E, Petrovic KE, Peltonen L, Wichmann HE, Schnabel RB, Dorr M, Parsa A, Aspelund T, Demissie S, Kathiresan S, Reilly MP, Taylor K, Uitterlinden A, Couper DJ, Sitzer M, Kahonen M, Illig T, Wild PS, Orru M, Ludemann J, Shuldiner AR, Eiriksdottir G, White CC, Rotter JI, Hofman A, Seissler J, Zeller T, Usala G, Ernst F, Launer LJ, D'Agostino RB Sr, O'Leary DH, Ballantyne C, Thiery J, Ziegler A, Lakatta EG, Chilukoti RK, Harris TB, Wolf PA, Psaty BM, Polak JF, Li X, Rathmann W, Uda M, Boerwinkle E, Klopp N, Schmidt H, Wilson JF, Viikari J, Koenig W, Blankenberg S, Newman AB, Witteman J, Heiss G, Duijn C, Scuteri A, Homuth G, Mitchell BD, Gudnason V, O'Donnell CJ, Consortium CA. Meta-analysis of genome-wide association studies from the CHARGE consortium identifies common variants associated with carotid intima media thickness and plaque. *Nat Genet* 2011;**43**:940–947.
89. Verhaaren BFJ, Debette S, Bis JC, Smith JA, Ikram MK, Adams HH, Beecham AH, Rajan KB, Lopez LM, Barral S, van Buchem MA, van der Grond J, Smith AV, Hegenscheid K, Aggarwal NT, de Andrade M, Atkinson EJ, Beekman M, Beiser AS, Blanton SH, Boerwinkle E, Brickman AM, Bryan RN, Chauhan G, Chen CPLH, Chouraki V, de Craen AJM, Crivello F, Deary IJ, Deelen J, De Jager PL, Dufouil C, Elkind MSV, Evans DA, Freudenberger P, Gottesman RF, Guðnason V, Habes M, Heckbert SR, Heiss G, Hilal S, Hofer E, Hofman A, Ibrahim-Verbaas CA, Knopman DS, Lewis CE, Liao J, Liewald DCM, Luciano M, van der Lugt A, Martinez OO, Mayeux R, Mazoyer B, Nalls M, Nauck M, Niessen WJ, Oostra BA, Psaty BM, Rice KM, Rotter JI, von Sarnowski B, Schmidt H, Schreiner PJ, Schuur M, Sidney SS, Sigurdsson S, Slagboom PE, Stott DJM, van Swieten JC, Teumer A, Töglhofer AM, Traylor M, Trompet S, Turner ST, Tzourio C, Uh H-W, Uitterlinden AG, Vernooij MW, Wang JJ, Wong TY, Wardlaw JM, Windham BG, Wittfeld K, Wolf C, Wright CB, Yang Q, Zhao W, Zijdenbos A, Jukema JW, Sacco RL, Kardia SLR, Amouyel P, Mosley TH, Longstreth WT, DeCarli CC, van Duijn CM, Schmidt R, Launer LJ, Grabe HJ, Seshadri SS, Ikram MA, Fornage M. Multiethnic genome-wide association study of cerebral white matter hyperintensities on MRI. *Circ Cardiovasc Genet* 2015;**8**:398–409.
90. AFGen Consortium, Metastroke Consortium of the ISGC, Neurology Working Group of the Charge Consortium. Large-scale analyses of common and rare variants identify 12 new loci associated with atrial fibrillation. *Nat Genet* 2017;**49**:946–952.
91. Lindgren A, Maguire J. Stroke recovery genetics. *Stroke* 2016;**47**:2427–2434.
92. Sachdev PS, Lo JW, Crawford JD, Mellon L, Hickey A, Williams D, Bordet R, Mendyk AM, Gele P, Deplanque D, Bae HJ, Lim JS, Brodtmann A, Werden E, Cumming T, Kohler S, Verhey FR, Dong YH, Tan HH, Chen C, Xin X, Kalaria RN, Allan LM, Akinyemi RO, Ogunniyi A, Klimkowicz-Mrowiec A, Dichgans M, Wollenweber FA, Zietemann V, Hoffmann M, Desmond DW, Linden T, Blomstrand C, Fagerberg B, Skoog I, Godefroy O, Barbay M, Roussel M, Lee BC, Yu KH, Wardlaw J, Makin SJ, Doubal FN, Chappell FM, Srikanth VK, Thrift AG, Donnan GA, Kandiah N, Chander RJ, Lin X, Cordonnier C, Moulin S, Rossi C, Sabayan B, Stott DJ, Jukema JW, Melkas S, Jokinen H, Erkinjuntti T, Mok VC, Wong A, Lam BY, Leys D, Henon H, Bombois S, Lipnicki DM, Kochan NA. STROKOG (stroke and cognition consortium): an international consortium to examine the epidemiology, diagnosis, and treatment of neurocognitive disorders in relation to cerebrovascular disease. *Alzheimer Dement* 2017;**7**:11–23.
93. Traylor M, Rutten-Jacobs LC, Holliday EG, Malik R, Sudlow C, Rothwell PM, Maguire JM, Koblar SA, Bevan S, Boncoraglio G, Dichgans M, Levi C, Lewis CM, Markus HS. Differences in common genetic predisposition to ischemic stroke by age and sex. *Stroke* 2015;**46**:3042–3047.
94. Randall JC, Winkler TW, Kutalik Z, Berndt SI, Jackson AU, Monda KL, Kilpelainen TO, Esko T, Magi R, Li S, Workalemahu T, Feitosa MF, Croteau-Chonka DC, Day FR, Fall T, Ferreira T, Gustafsson S, Locke AE, Mathieson I, Scherag A, Vedantam S, Wood AR, Liang L, Steinthorsdottir V, Thorleifsson G, Dermitzakis ET, Dimas AS, Karpe F, Min JL, Nicholson G, Clegg DJ, Person T, Krohn JP, Bauer S, Buechler C, Eisinger K, Consortium D, Bonnefond A, Froguel P, Investigators M, Hottenga JJ, Prokopenko I, Waite LL, Harris TB, Smith AV, Shuldiner AR, McArdle WL, Caulfield MJ, Munroe PB, Gronberg H, Chen YD, Li G, Beckmann JS, Johnson T, Thorsteinsdottir U, Teder-Laving M, Khaw KT, Wareham NJ, Zhao JH, Amin N, Oostra BA, Kraja AT, Province MA, Cupples LA, Heard-Costa NL, Kaprio J, Ripatti S, Surakka I, Collins FS, Saramies J, Tuomilehto J, Jula A, Salomaa V, Erdmann J, Hengstenberg C, Loley C, Schunkert H, Lamina C, Wichmann HE, Albrecht E, Gieger C, Hicks AA, Johansson A, Pramstaller PP, Kathiresan S, Speliotes EK, Penninx B, Hartikainen AL, Jarvelin MR, Gyllensten U, Boomsma DI, Campbell H, Wilson JF, Chanock SJ, Farrall M, Goel A, Medina-Gomez C, Rivadeneira F, Estrada K, Uitterlinden AG, Hofman A, Zillikens MC, den Heijer M, Kiemeney LA, Maschio A, Hall P, Tyrer J, Teumer A, Volzke H, Kovacs P, Tonjes A, Mangino M, Spector TD, Hayward C, Rudan I, Hall AS, Samani NJ, Attwood AP, Sambrook JG, Hung J, Palmer LJ, Lokki ML, Sinisalo J, Boucher G, Huikuri H, Lorentzon M, Ohlsson C, Eklund N, Eriksson JG, Barlassina C, Rivolta C, Nolte IM, Snieder H, Van der Klauw MM, Van Vliet-Ostaptchouk JV, Gejman PV, Shi J, Jacobs KB, Wang Z, Bakker SJ, Mateo Leach I, Navis G, van der Harst P, Martin NG, Medland SE, Montgomery GW, Yang J, Chasman DI, Ridker PM, Rose LM, Lehtimaki T, Raitakari O, Absher D, Iribarren C, Basart H, Hovingh KG, Hypponen E, Power C, Anderson D, Beilby JP, Hui J, Jolley J, Sager H, Bornstein SR, Schwarz PE, Kristiansson K, Perola M, Lindstrom J, Swift AJ, Uusitupa M, Atalay M, Lakka TA, Rauramaa R, Bolton JL, Fowkes G, Fraser RM, Price JF, Fischer K, Krjuta Kov K, Metspalu A, Mihailov E, Langenberg C, Luan J, Ong KK, Chines PS, Keinanen-Kiukaanniemi SM, Saaristo TE, Edkins S, Franks PW, Hallmans G, Shungin D, Morris AD, Palmer CN, Erbel R, Moebus S, Nothen MM, Pechlivanis S, Hveem K, Narisu N, Hamsten A, Humphries SE, Strawbridge RJ, Tremoli E, Grallert H, Thorand B, Illig T, Koenig W, Muller-Nurasyid M, Peters A, Boehm BO, Kleber ME, Marz W, Winkelmann BR, Kuusisto J, Laakso M, Arveiler D, Cesana G, Kuulasmaa K, Virtamo J, Yarnell JW, Kuh D, Wong A, Lind L, de Faire U, Gigante B, Magnusson PK, Pedersen NL, Dedoussis G, Dimitriou M, Kolovou G, Kanoni S, Stirrups K, Bonnycastle LL, Njolstad I, Wilsgaard T, Ganna A, Rehnberg E, Hingorani A, Kivimaki M, Kumari M, Assimes TL, Barroso I, Boehnke M, Borecki IB, Deloukas P, Fox CS, Frayling T, Groop LC, Haritunians T, Hunter D, Ingelsson E, Kaplan R, Mohlke KL, O'Connell JR, Schlessinger D, Strachan DP, Stefansson K, van Duijn CM, Abecasis GR, McCarthy MI, Hirschhorn JN, Qi L, Loos RJ, Lindgren CM, North KE, Heid IM. Sex-stratified genome-wide association studies including 270, 000 individuals show sexual dimorphism in genetic loci for anthropometric traits. *PLoS Genet* 2013;**9**:e1003500.
95. Konig IR, Loley C, Erdmann J, Ziegler A. How to include chromosome X in your genome-wide association study. *Genet Epidemiol* 2014;**38**:97–103.
96. Loley C, Alver M, Assimes TL, Bjonnes A, Goel A, Gustafsson S, Hernesniemi J, Hopewell JC, Kanoni S, Kleber ME, Lau KW, Lu Y, Lyytikainen LP, Nelson CP, Nikpay M, Qu L, Salfati E, Scholz M, Tukiainen T, Willenborg C, Won HH, Zeng L, Zhang W, Anand SS, Beutner F, Bottinger EP, Clarke R, Dedoussis G, Do R, Esko T, Eskola M, Farrall M, Gauguier D, Giedraitis V, Granger CB, Hall AS, Hamsten A, Hazen SL, Huang J, Kahonen M, Kyriakou T, Laaksonen R, Lind L, Lindgren C, Magnusson PK, Marouli E, Mihailov E, Morris AP, Nikus K, Pedersen N, Rallidis L, Salomaa V, Shah SH, Stewart AF, Thompson JR, Zalloua PA, Chambers JC, Collins R, Ingelsson E, Iribarren C, Karhunen PJ, Kooner JS, Lehtimaki T, Loos RJ, Marz W, McPherson R, Metspalu A, Reilly MP, Ripatti S, Sanghera DK, Thiery J, Watkins H, Deloukas P, Kathiresan S, Samani NJ, Schunkert H, Erdmann J, Konig IR. No association of coronary artery disease with x-chromosomal variants in comprehensive international meta-analysis. *Sci Rep* 2016;**6**:35278.

97. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, ReproGen C, Psychiatric Genomics C, Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control C, Duncan L, Perry JR, Patterson N, Robinson EB, Daly MJ, Price AL, Neale BM. An atlas of genetic correlations across human diseases and traits. *Nat Genet* 2015;**47**:1236–1241.
98. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C, Patterson N, Daly MJ, Price AL, Neale BM. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 2015;**47**:291–295.
99. Pickrell JK, Berisa T, Liu JZ, Segurel L, Tung JY, Hinds DA. Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet* 2016;**48**:709–717.
100. Zheng J, Erzurumluoglu AM, Elsworth BL, Kemp JP, Howe L, Haycock PC, Hemani G, Tansey K, Laurin C, Early G, Lifecourse Epidemiology Eczema C, Pourcain BS, Warrington NM, Finucane HK, Price AL, Bulik-Sullivan BK, Anttila V, Paternoster L, Gaunt TR, Evans DM, Neale BM. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* 2017;**33**:272–279.
101. Denny JC, Ritchie MD, Basford MA, Pulley JM, Bastarache L, Brown-Gentry K, Wang D, Masys DR, Roden DM, Crawford DC. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics* 2010;**26**:1205–1210.
102. Hebbring SJ. The challenges, advantages and future of phenome-wide association studies. *Immunology* 2014;**141**:157–165.
103. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, Liu B, Matthews P, Ong G, Pell J, Silman A, Young A, Sprosen T, Peakman T, Collins R. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015; **12**:e1001779.