Review

Common and Rare Variant Association Study for Plasma Lipids and Coronary Artery Disease

Hayato Tada, Masa-aki Kawashiri, Tetsuo Konno, Masakazu Yamagishi and Kenshi Hayashi

Division of Cardiovascular Medicine, Kanazawa University Graduate School of Medicine, Kanazawa, Japan

Blood lipid levels are highly heritable and modifiable risk factors for coronary artery disease (CAD), and are the leading cause of death worldwide. These facts have motivated human genetic association studies that have the substantial potential to define the risk factors that are causal and to identify pathways and therapeutic targets for lipids and CAD.

The success of the HapMap project that provided an extensive catalog of human genetic variations and the development of microarray based genotyping chips (typically containing variations with allele frequencies >5%) facilitated common variant association study (CVAS; formerly termed genome-wide association study, GWAS) identifying disease-associated variants in a genome-wide manner. To date, 157 loci associated with blood lipids and 46 loci with CAD have been successfully identified, accounting for approximately 12%−14% of heritability for lipids and 10% of heritability for CAD. However, there is yet a major challenge termed "missing heritability problem," namely the observation that loci detected by CVAS explain only a small fraction of the inferred genetic variations. To explain such missing portions, focuses in genetic association studies have shifted from common to rare variants. However, it is challenging to apply rare variant association study (RVAS) in an unbiased manner because such variants typically lack the sufficient number to be identified statistically.

In this review, we provide a current understanding of the genetic architecture mostly derived from CVAS, and several updates on the progress and limitations of RVAS for lipids and CAD.

J Atheroscler Thromb, 2016; 23: 241-256.

Key words: Genome-wide association study; rare variant association study, Exome, Genome

Introduction

It has been shown that blood lipid levels are highly heritable and modifiable risk factors for coronary artery disease (CAD), which is the leading cause of death worldwide[1-4]. These facts have motivated human genetic association studies that have the substantial potential to define the risk factors that are causal and to identify pathways and therapeutic targets for lipids and CAD[5].

Common and rare variants across the full minor allele frequency (MAF) spectrum are important constituents of the genetic architecture of disease traits[6]. Using a common variant association study (CVAS; formerly called genome-wide association study, GWAS), 157 loci associated with blood lipids and 46 loci associated with CAD have been successfully identified, accounting for approximately12%−14% of heritability for lipids[7] and 10% of heritability for CAD[8]. In addition, it has been shown that variants with small effects can indicate pathways and therapeutic targets that enable clinically important changes in blood lipid levels and CAD[9]. Although several rare variant association studies (RVAS) using a candidate gene approach for complex traits have successfully identified novel associations, including lipid phenotypes[10, 11], it is yet quite challenging to apply RVAS in an unbiased manner because of the lack of statistical power.

In this review, we provide a current understanding of the genetic architecture mostly derived from CVAS and several updates on the progress and limitations of RVAS for lipids and CAD.

Address for correspondence: Hayato Tada, Division of Cardiovascular Medicine, Kanazawa University Graduate School of Medicine, 13-1 Takara-machi, Kanazawa 920-8641, Japan
E-mail: ht240z@sa3.so-net.ne.jp
Received: May 30, 2015
Accepted for publication: July 31, 2015

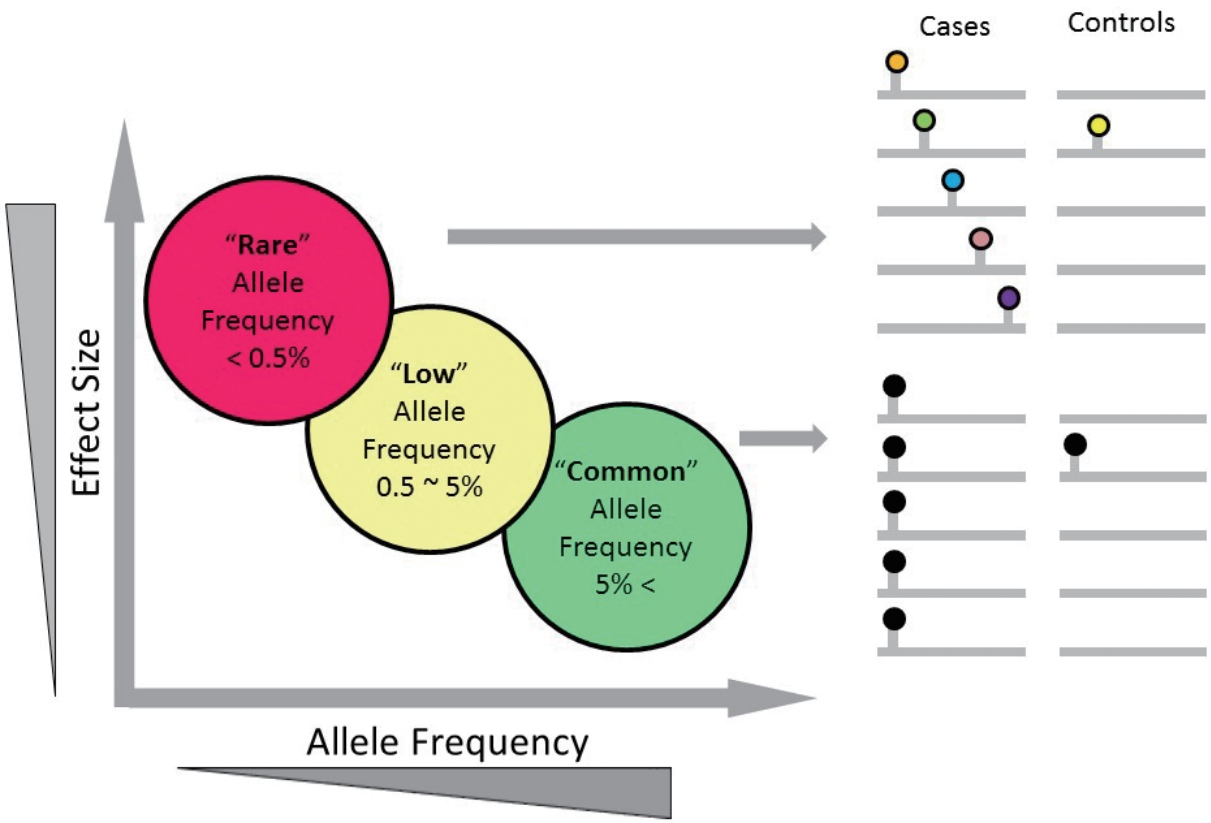

Fig. 1. Genetic architecture of disease traits and association testing using rare alleles.

(A) "Common" and "Low-frequency" variants can be typically assayed by genotyping and tested for association individually using single marker tests of association.
(B) "Rare" variants may be observed only in a single person (i.e., a singleton). Variants in this frequency range must be grouped together and tested in aggregate. Association signals from a burden of such rare variants are typically best validated using resequencing of the same genomic interval in additional individuals. This figure is adopted from Kathiresan *et al.* (2012).
Each line represents a DNA sequence from a different individual, and each circle represents a DNA sequence variant. Each color denotes a DNA sequence variant allele at a different nucleotide site.

Genetic Architecture for Complex Traits

What are Rare Variants?

With regard to the genetic architecture for a certain trait, it is useful to divide the variants according to their allele frequencies (**Fig. 1**) because this facilitates the determination of optimal research strategies[12]. Here, we define the threshold of "common" variants for MAF to be >5% of the population (greater than 1:10 people have a minor allele), and "rare" for MAF if <0.5% (less than 1:100 people have a minor allele). In **Table 1**, we divide potential causal variant frequencies and outline their implications for analytical strategies.

Why Rare Variants?

Although CVAS has identified thousands of common genetic variants associated with complex human traits, a large fraction of heritability is unexplained in most of these traits. In addition, large-scale surveys of human genetic variation using next-generation sequencing (NGS) technology have reported that human populations harbor an abundance of rare variants, several of which are deleterious because of the recent explosive population growth and weak purifying selection[13-15]. Those notions fortify the importance and relevance of exploring rare variants to understand human disease risk.

Common Variant/Complex Trait Hypothesis

The first concept of a genome-wide scan for association to detect variants that contribute to complex diseases was described in 1996[16]. Immediately after its development, the HapMap project provided an extensive catalog of SNPs that represented a significant fraction of all existing variations in the genomes

Table 1. Variants classification for association testing

Variant class	MAF	Implication for analysis
Common	>5%	Contemporary CVAS
Low-frequency	0.5%−5%	RVAS using exome array
		RVAS using exome array
Rare	<0.5%	RVAS using exome/genome sequencing
Private	Restricted to probands and immediate relatives	Still challenging

MAF: Minor allele frequency, CVAS: Common variant association study, RVAS: Rare variant association study

of multiple human populations[17]. With regard to this and microarray-based genotyping, using "GWAS chips," the Wellcome Trust Case Control Consortium (WTCCC) reported a "landmark paper," a modern CVAS strategy that demonstrated 24 independent associated signals across 7 different types of common complex traits using 2,000 cases and 3,000 controls[18]. Thereafter, several thousand genetic variants have been shown to be robustly associated with complex traits (http://www.genome.gov/gwastudies), providing valuable insight into the complexities of their genetic architecture[19]. These CVAS use common variations (MAF >5%), therefore examining only a portion of the genomic landscape of complex traits.

In addition, variants identified through CVAS are not likely to be the causal variant because selections of the variants are based on their distribution across the genome and not on known biological functions for a certain phenotype. However, those variants are likely to be associated only through linkage disequilibrium with the true, but unknown, causal variants. A large amount of common variants identified through CVAS associated with common traits are located outside of the genome's coding regions. This has complicated their biological interpretation and has resulted in enthusiasm for exome analysis targeting the coding regions.

Common Variant/Lipid and CAD Traits

In 2007, a common genetic variant in the glucokinase regulatory protein (GCKR) locus was shown to be associated with plasma triglyceride levels using the CVAS approach[20]. Since then, larger size meta-analyses using CVAS have successfully identified new loci associated with lipids and replicated the previously reported loci[21-23]. Currently, 157 loci have been shown to be associated with lipid traits, accounting for approximately12%-14% of the total variance in each lipid trait, as shown by the Framingham Heart Study, which employed the largest meta-analysis of CVAS using 188,577 individuals (**Fig. 2**). Furthermore, 40 among the 157 loci were shown to be associated with CAD traits, demonstrating the causal relevance of lipid traits to CAD (**Fig. 2**).

Similarly, the first robust meta-analysis of CVAS identified a locus (9p21) associated with CAD in 2007[18]. Since then, 46 loci have been identified to be associated with CAD traits, accounting for approximately 10% of the total variance through the largest meta-analysis of CVAS to date, using 63,746 CAD cases and 130,6811 controls[8]. Moreover, 12 among the 46 loci showed a significant association with the lipid traits, suggesting the importance of lipids (**Fig. 3**), particularly LDL cholesterol and triglyceride, as a causal CAD risk factor[24].

However, similar to most traits where independent estimates of the total contribution of genetics to variation in the trait are available, the variants identified by the largest meta-analysis of CVAS for lipids and CAD explain only a small fraction of the total genetic variance, resulting in the general question of how to account for the unexplained heritability[25]. Although several articles have indicated that much of that missing heritability could be explained by common variants than by rare variants[26-28], it is debated whether rare variants substantially contribute to the genetic architecture of complex traits. Several reasons are provided in the following section.

Rare Variant/Complex Trait Hypothesis

A substantial number of rare variants are known to be responsible for monogenic, Mendelian diseases caused by single-gene mutation(s) with a major effect on a wide variety of traits[29, 30]. Initially, familial linkage analysis, followed by the publication of the Human Genome Project and the application of NGS, facilitated the identification of diseases that caused variations[31]. Now, >5,000 Mendelian diseases among a pool of 7,000 have been shown to be associated with a specific molecular defect[29]. However, such variants do not reach significant frequencies in human populations, with dominant variants mostly occurring because

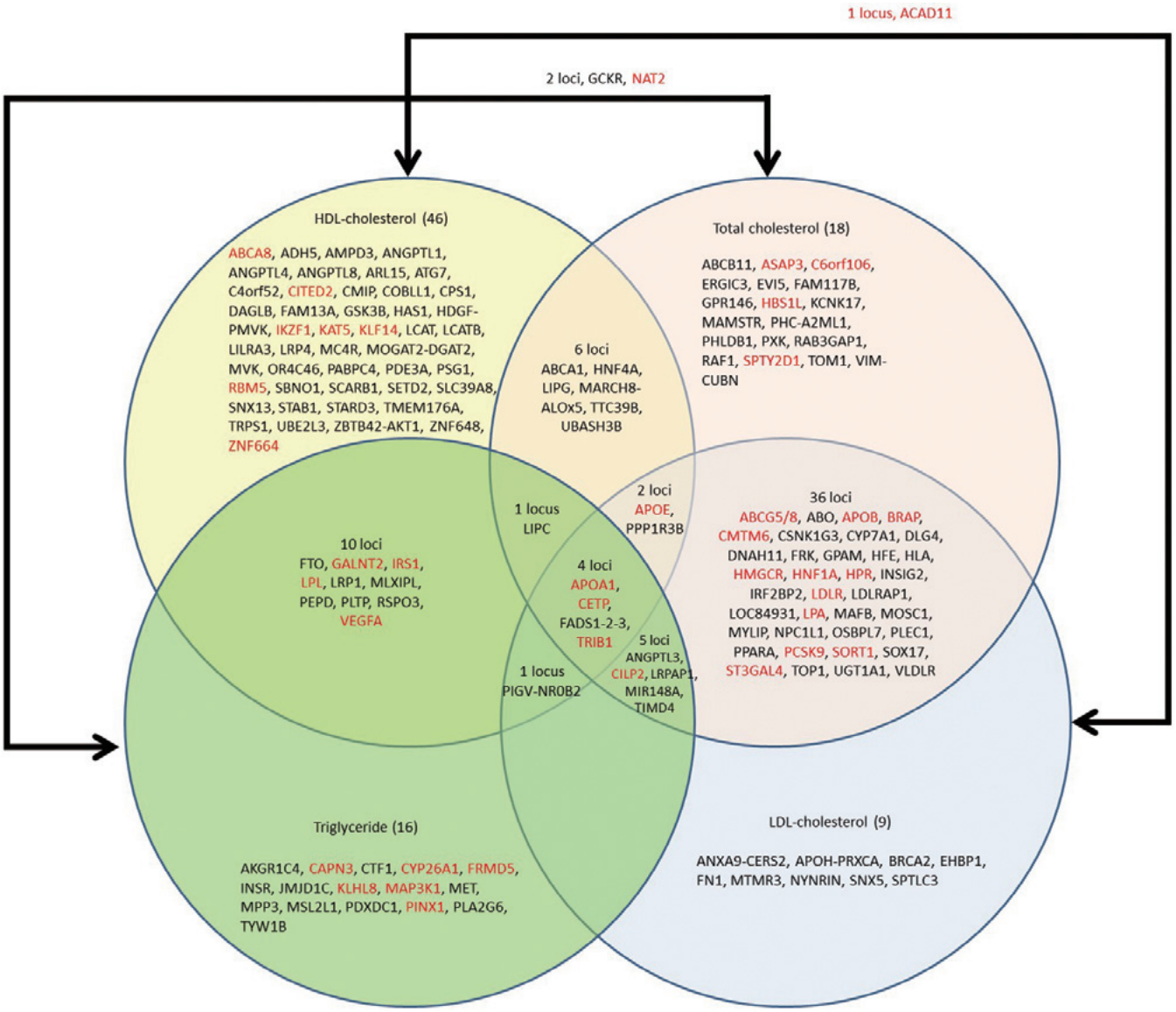

Fig. 2. 157 loci associated with plasma lipids detected through CVAS.

The number of loci primarily associated with only one trait is listed in parentheses after the trait name, and locus names are listed below. The loci that show an association with two or more traits are shown in the appropriate segment. Red colored genes have been shown to be associated with CAD. This figure was adopted and modified by reference 7.

of *de novo* mutations, whereas recessive variants are maintained at low frequencies by selection against homozygotes. Several massive studies using NGS technology have shown that the vast majority of variations are evolutionarily recent, rare, and enriched for deleterious alleles, resulting in the notion that those rare variations likely provide an important contribution to human phenotypic variation and disease susceptibility[13, 14, 32]. Similarly, it has been proposed that rare variants affecting common complex traits tend to have larger effect sizes than common variants, and both common variants with small effect sizes and rare variants with moderate effect sizes may provide significant contributions to variance in complex traits[33]. Therefore, it is reasonable to assume that a substantial proportion of trait heritability could be explained by those rare variants, although it is challenging to detect associations or to determine their effect.

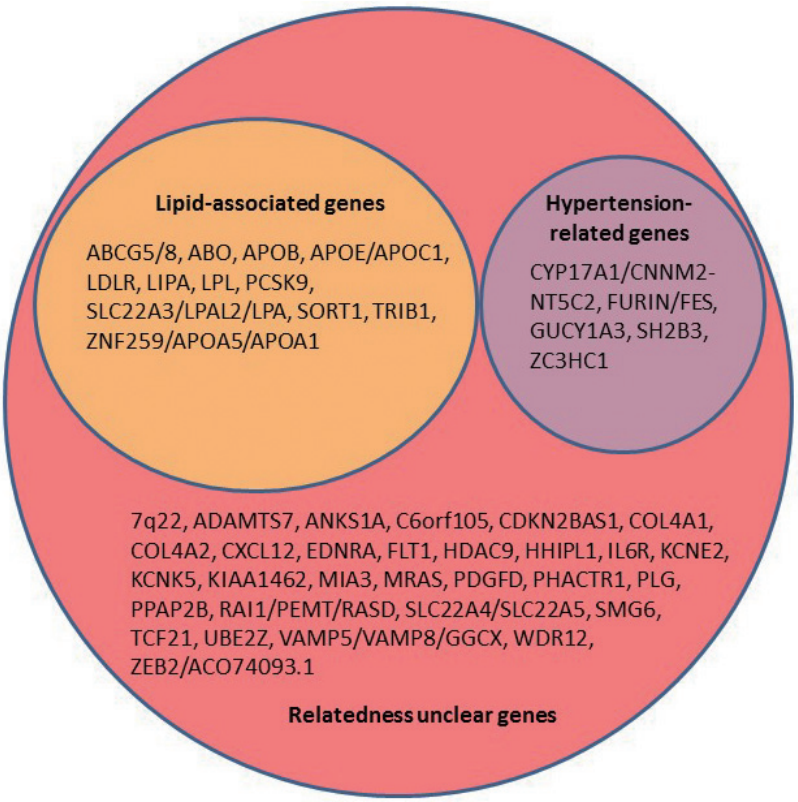

Fig. 3. 46 loci associated with CAD detected through CVAS.

Genes in the orange circle represent lipid-associated genes, hypertension-related genes are shown in purple, and genes that show unclear relatedness are shown in red circles. This figure was adopted and modified by reference 8.

Exome Array

Although exome or genome sequencing is increasingly affordable and facilitates the detection of rare variants in the human genome, genotyping arrays are a cost-effective approach when investigating genetic variants previously identified in large populations. Focusing on the exome and selecting variants likely to affect protein structure and function (e.g., nonsynonymous, stop codon, insertion/deletion, and splice site variants), a large number of genome variants in human populations could be cataloged into a small set of variants on a genotyping array with comprehensive coverage of the variants most likely to have functional consequences. This approach has recently resulted in the development of "exome arrays" on microarray genotyping platforms[34]. It is estimated that low-frequency to moderately rare variants (MAF approximately 0.1%) can be detected using these arrays. Currently, exome arrays are being widely applied to large cohorts, several of which already have CVAS data, to expand the allele frequency spectrum to low-frequency and moderately rare variants and to directly detect functionally significant variants. Using an exome array, Huyghe *et al.* reported rare variants associated with insulin processing and secretion in three novel genes and in two known genes[35]. More similar studies using these arrays are expected to be published within the next few years.

Exome Sequencing

Since 2007, great advances have been made in sequencing and capture technologies, enabling the accurate determination of nearly all protein-coding sequence variants in an individual[36-41]. These exome sequencing technologies have already accelerated genetic studies of Mendelian disorders, yielding approximately 20%−30% success in diagnosis, and there is great interest in extending them to complex traits[42-45]. To address this issue, a large number of strategies, analysis, and interpretation of exome sequencing studies have been proposed[46-49], and focused candidate gene sequencing studies have shown some promising results[9, 10, 50-54]. In complex trait genetics, exome sequencing studies can be used to

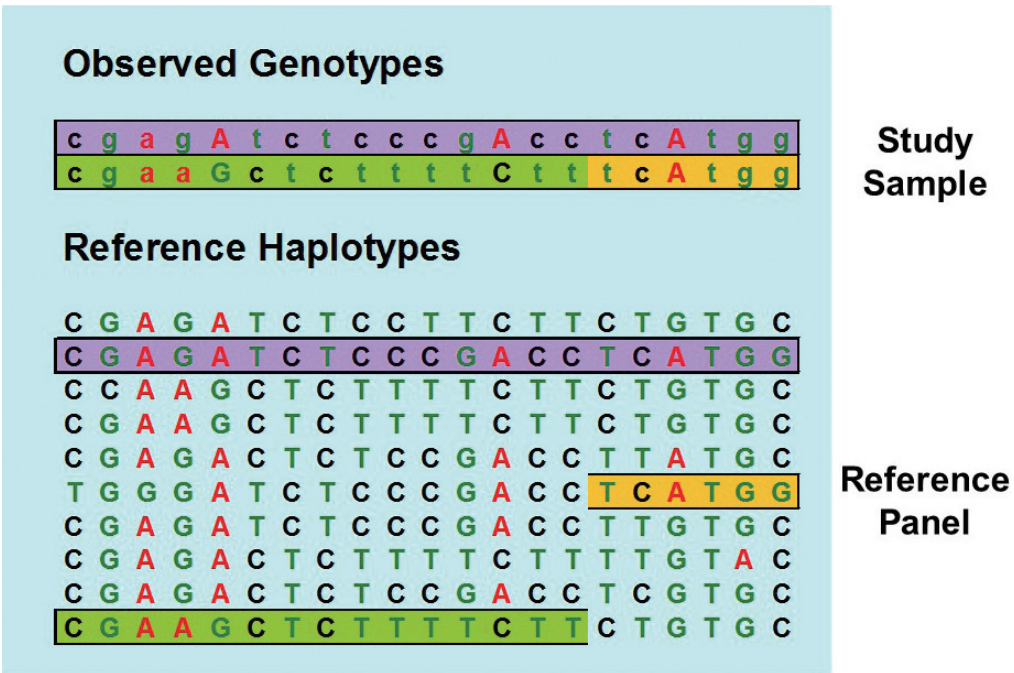

Fig. 4. Genotype imputation.

Uppercase indicates "actually typed" genotypes, and lowercase indicates "imputed" genotypes.

identify rare coding variants that are not detected by microarray-based CVAS or exome array-based RVAS. The potential of exome sequencing studies of complex traits is based on the success of candidate gene studies and has firm roots in population genetic theory, motivating researchers to continue pursuing the collection of much larger sample sizes that definitely bring fruitful results.

Statistical Methods for the Analysis of Rare Variants

Candidate gene sequencing, and deep resequencing of genes associated with complex traits detected through CVAS, has identified rare variants that explain additional trait variance, suggesting the importance of aggregating information of multiple rare variants to produce adequately powered tests for association[9, 10, 50].

However, there are several complicating factors associated with the construction of well-powered gene-based association testing. Only a small fraction of variants within a gene may affect gene function, and strategies for classification of variants could be misleading with false-positive or false-negative results. Different variants may affect gene function in different ways, and simple aggregation methods may inappropriately cancel the association signals. Furthermore, different functional variants will have different effect sizes and should be appropriately weighed in an aggregate test. Among an enormous amount of papers focusing on statistical strategies, there are four different types of approaches: burden tests, mixed-effects models, data-driven approaches, and hybrid methods[55]. Burden tests basically test whether there is an excess of rare variants in cases or controls, aggregating multiple rare variants in a gene. However, when a portion of causal variants has effects in opposite directions, aggregated variants could weaken the association signal. Mixed-effects models have been developed to deal with this issue, providing it with more power compared with burden tests in the presence of opposite directions of genetic effects (e.g., sequence kernel association test (SKAT))[56]. A data-driven approaches perform several tests to identify the optimal cutoff (such as MAF) and uses permutation to correct multiple testing to obtain empirical p values (e.g., the variable threshold)[57]. Although all of these approaches are weaker when non-causal variants are included in the analysis, hybrid methods have been proposed to combine methods that are powerful for variants with the same effects and methods that are robust when either non-causal variants or variants with opposite effects are present (e.g., SKAT-O)[58]. Because of these pros and cons, it appears reasonable to adopt several different types of analytical strategies listed above to under-

Table 2. Novel associations identified through RVAS

Year	Trait	Gene	Variants	MAF	Effect size [mg/dl (s.e.) or Odds ratio]	Genotyping	Number of samples	Population samplings	Ref.
2013	Triglyceride HDL cholesterol	APOC3	rs76353203	3.80%	−45.5 (6.6) 12.3 (2.1)	exome array	1267	isolated population-based cohort	71
2014	LDL cholesterol	ABCA6	rs77542162	3.40%	5.1 (0.7)	imputation-based CVAS	36,000	population-based cohort (Meta-analysis)	59
2014	LDL cholesterol	PNPLA5	multiple variants	<5% (gene-based)	41.7 (7.8)	exome sequencing	2005	population-based cohort/extremes	60
2014	Total cholesterol CAD	TM6SF2	rs58542926	8.90%	−5.6 (1.4) 0.87	exome array	5643	population-based cohort	68
2015	Early onset MI	LDLR/APOA5	multiple variants	<5% (gene-based)	2.4/2.0	exome sequencing	9793	case-control (extremes)	61
2015	Total cholesterol	ADAMTS3	rs117087731	1%	11.9 (2.0)	imputation-based CVAS	62166	population-based cohort (Meta-analysis)	62
2015	Total cholesterol	MTHFD2L	rs182616603	1%	14.5 (1.7)	imputation-based CVAS	62166	population-based cohort (Meta-analysis)	62
2015	LDL cholesterol	MTHFD2L	rs182616603	1%	12.1 (1.7)	imputation-based CVAS	62166	population-based cohort (Meta-analysis)	62
2015	Triglyceride	TM4SF5	rs193042029	1%	−15.0 (2.6)	imputation-based CVAS	62166	population-based cohort (Meta-analysis)	62

MAF: minor allele frequency, CAD: coronary artery disease, APOC3: apolipoprotein C3, ABCA6: ATP-binding cassette, sub-family A (ABC1), member 6, PNPLA5: patatin-like phospholipase domain containing 5, TM6SF2: transmembrane 6 superfamily 2, LDLR: LDL receptor, APOA5: apolipoprotein A5, ADAMTS3: ADAM metallopeptidase with thrombospondin type 1 motif, 3, MTHFD2L: methylenetetrahydrofolate dehydrogenase ($NADP^+$ dependent) 2-like, TM4SF5: transmembrane 4 L6 family member 5, CVAS: common variant association study

stand the genetic architecture of rare variants for certain traits.

Genotype-Imputation

Because of the accumulation of a deep catalogue of human genotypes by 1000 genome project and the HapMap Project, we can estimate the missing genotypes based on haplotypes in a certain population (**Fig. 4**), expanding the opportunity for identifying signals in several CVAS, including lipids and CAD[7, 8]. Recently, the accumulation of genotypes across ethnicities enables us to use population-specific reference panel, resulting in identifications of novel associations in rare variants including lipids[59].

Rare Variant/Lipid and CAD Traits

In **Table 2**, genes identified through RVAS are summarized. In 2014, a novel association between LDL cholesterol and the burden of rare or low-frequency variants in patatin-like phospholipase domain containing 5 (PNPLA5), encoding phospholipase-domain-containing protein through exome sequencing for the subjects with extreme LDL cholesterol phenotype, replicated by sequencing study in independent samples[60]. It is reasonable to adopt such a strategy (using extreme phenotype) to identify signals from rare variants. This is because is the strategy is more likely to identify signals in such cases with extreme phenotype under the condition in which a greater sample size is required to observe associations in RVAS because of its frequency. Similarly, Kathiresan and co-workers investigated associations between rare coding variants using exome sequencing and early onset myocardial infarction, and they identified that the burden of rare variants in LDL receptor, and apolipoprotein A-V (APOA5) genes contributed to the development of such extreme cases[61]. This is quite impressive because an unbiased manner with a reasonable sample size revealed that approximately 3% of such cases could potentially have familial hypercholesterolemia caused by LDL receptor mutations, and approximately 1% of such cases could potentially have familial hypertriglyceridemia caused by APOA5 mutations. Moreover, a genome-wide imputation and association analysis revealed several novel loci associated with lipids (ADAM metallopeptidase with thrombospondin type 1 motif, 3; ADAMTS3, methylenetetrahydrofolate dehydrogenase ($NADP^+$ dependent)

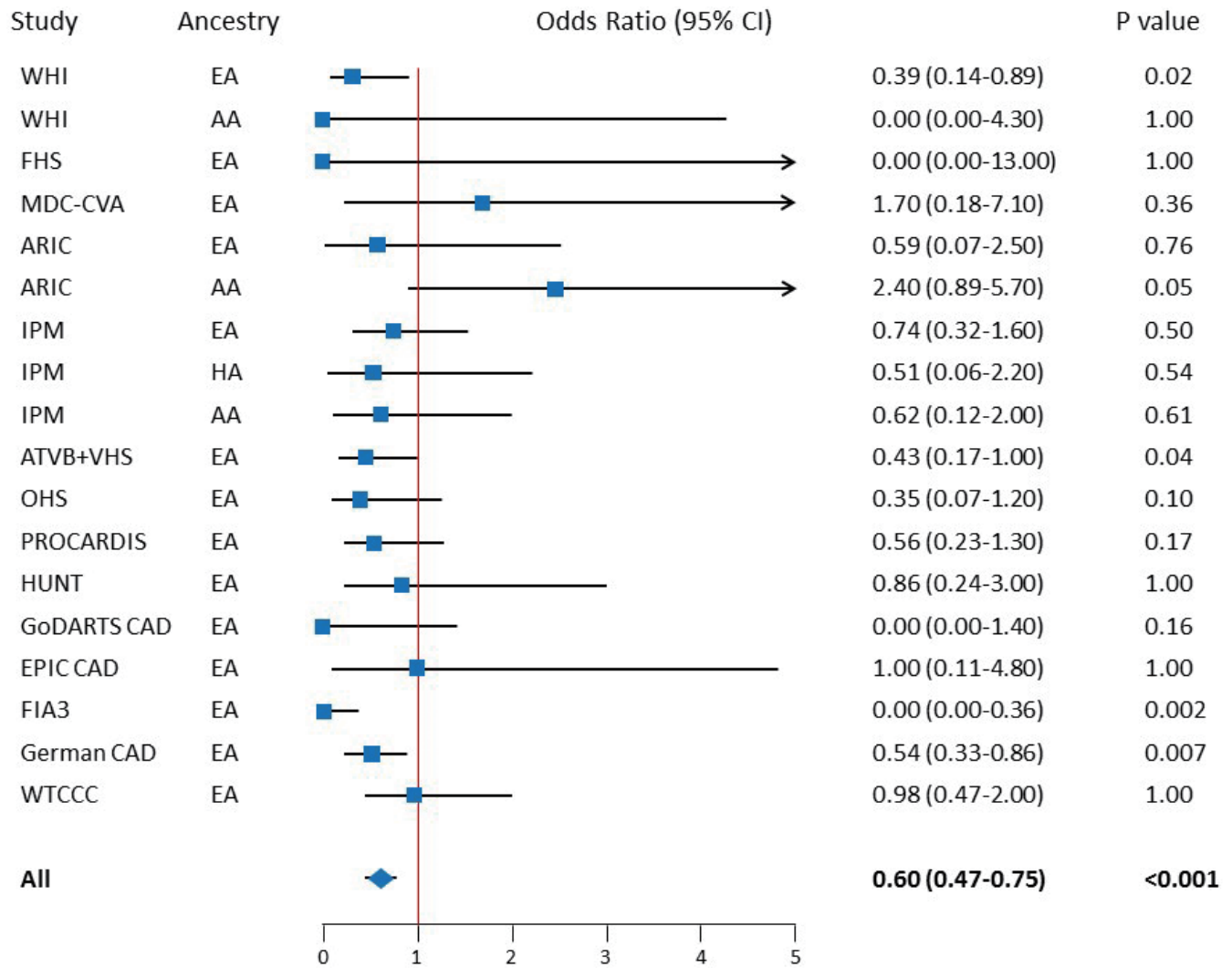

Fig. 5. Association of APOC3 loss-of-function mutations with risk of coronary heart disease among 110,970 participants in 15 studies.

The diamond indicates a combined estimate of the odds ratio and the corresponding 95% confidence interval. HA denotes Hispanic ancestry. The full study names are as follows: ARIC, Atherosclerosis Risk in Communities; ATVB, Italian Atherosclerosis, Thrombosis, and Vascular Biology Study; EPIC, European Prospective Study into Cancer and Nutrition; FHS, Framingham Heart Study; FIA3, First Myocardial Infarction, in AC County 3; GoDARTS, Genetics of Diabetes Audit and Research Tayside Study; HUNT, Nord−Trøndelag Health Study; IPM, Mt. Sinai Institute for Personalized Medicine Biobank; MDC−CVA, Malmö Diet and Cancer Study Cardiovascular Cohort; OHS, Ottawa Heart Study; PROCARDIS, Precocious Coronary Artery Disease Study; VHS, Verona Heart Study; WHI, Women's Health Initiative; WTCCC, Wellcome Trust Case Control Consortium. EA, European ancestry; AA, African ancestry.

2-Like; MTHFD2L, transmembrane 4 L6 family member 5; TM4SF5), including low-frequency variants[62].

In addition, using targeted gene sequencing strategies, several rare variants have been shown to be associated with lipids, as stated in the previous sections. The identifications of such rare variants, regardless of their effect sizes, could successfully result in the development of novel types of drugs[63-66]. Consistent with this paradigm, the particular rare apolipoprotein C-III (APOC3) variant was verified to be associated with a lowered triglyceride level coupled with a reduction of CAD through exome sequencing (**Fig. 5**)[67], resulting in the development of an APOC3 inhibitor. Moreover, another novel signal in transmembrane 6 superfamily 2 (TM6SF2, p.Glu167Lys) has been identified through the exome array study for the population-based cohort[68]. This particular variant was associated

with reduced total cholesterol level and a reduced risk for myocardial infarction.

However, several rare coding variants with large effect sizes on HDL cholesterol and/or triglycerides detected through the recent RVAS using 56,000 individuals were not associated with CAD, suggesting that examples of rare variants with robust effects on both lipids and CAD may be limited[69]. Now, researchers, including our group, are collecting much larger sample sizes (>100,000 in exome arrays and >10,000 in exome sequencing) to detect robust associations with lipid and CAD traits using rare variants in an exome-wide strategy.

Population Sampling Strategies

We must select the appropriate population sampling strategies, such as population-based cohort, case-control study, or extreme phenotypes, acknowledging the pros and cons of them in terms of the phenotypes of interest, effect sizes, and allele frequencies. For example, contemporary CVAS using case-control design were suitable to identify disease associated genes in common traits[8]. In contrast, recent RVAS for extreme phenotypes successfully identified novel signals based on the assumption that rarer variants with larger effect sizes should be identified in such extreme cases[60, 61]. In addition, several other studies used a specific population, particularly "isolates", as the frequency of rare variants may increase in certain populations because of drift and founder effects[70]. Furthermore, these studies identified robust associations in a variant in APOC3 (R19X) and triglyceride/HDL cholesterol[71] and a variant in ATP-binding cassette, subfamily A (ABC1), member 6 (ABCA6) (rs77542162), and LDL cholesterol[59].

Future Directions

Whole-Genome-Sequencing

Until recently, whole-genome sequencing-based studies have been used to identify mutations underlying rare Mendelian diseases[72-74], but not to explore the genetic architecture of common complex traits in population-based samples. In 2013, the initial step for analyzing whole-genome sequence data to characterize the genetic architecture of a HDL cholesterol trait was described[75]. In addition to highlighting the value of whole-genome sequencing beyond that afforded by exome sequencing, those authors estimated that 62% of heritability in the HDL cholesterol trait could be attributed to common variants (MAF >1%) and 8% to rare variants (MAF <1%). It is reasonable that the overall contributions by rare variants are smaller than those by common variants simple because of their difference in frequencies. However, identifications of novel associations in rare variants, although their overall contributions are relatively small, could result in the development of novel drugs.

Improved technologies for assaying rare variants combined with novel analytical methods for performing aggregate, gene-based tests of association will facilitate rare-variant association studies feasible on a genome-wide scale in well-powered but reasonably sized cohorts.

Refinement of Genetic Variation Information

With regard to the huge number of variants provided by current NGS technologies, it is an important challenge to understand the variants that are functionally significant to incorporate them into gene-based tests of association, particularly in RVAS. This requires improvements to bioinformatics methods for predicting the functions of protein-coding variants and an improved understanding of the regulatory architecture of the genome (non-coding variants). For both objectives, various efficient tools have already been developed, including recent advances by the Encyclopedia of DNA Elements (ENCODE) consortium, involving the distribution of potential regulatory sequences across the genome[76-80], the NIH Roadmap Epigenetic Project producing a public resource of human epigenomic data[81], and the FANTOM project led by RIKEN to develop effective system for functional gene annotation by designing appropriate rules and methods[82]. Dr. Shinya Yamanaka at Kyoto University established induced pluripotent stem cells using the FANTOM database to select 24 candidates of transcription factors[83]. However, variants predicted as being "damaging" by these tools are frequently annotation artifacts[84], suggesting that refining variants in annotations is yet a challenge.

Conclusions

Novel genotyping and sequencing technology, extensive catalogs of human genetic variation, and novel analytical methods have been developed for the feasibility of RVAS. It is indisputable that current rigorous efforts could reveal the contribution of rare variants to the overall genetic architecture of plasma lipids and CAD in the next few years.

Acknowledgments and Notice of Grant Support

None.

Conflict of Interest Disclosures

None.

References

1) Go AS, Mozaffarian D, Roger VL, Benjamin EJ, Berry JD, Borden WB, Bravata DM, Dai S, Ford ES, Fox CS, Franco S, Fullerton HJ, Gillespie C, Hailpern SM, Heit JA, Howard VJ, Huffman MD, Kissela BM, Kittner SJ, Lackland DT, Lichtman JH, Lisabeth LD, Magid D, Marcus GM, Marelli A, Matchar DB, McGuire DK, Mohler ER, Moy CS, Mussolino ME, Nichol G, Paynter NP, Schreiner PJ, Sorlie PD, Stein J, Turan TN, Virani SS, Wong ND, Woo D, Turner MB; American Heart Association Statistics Committee and Stroke Statistics Subcommittee Circulation, 2013; 127: 143-152
2) Arsenault BJ, Boekholdt SM, Kastelein JJ. Lipid parameters for measuring risk of cardiovascular disease. Nat Rev Cardiol, 2011; 8: 197-206
3) Emerging Risk Factors Collaboration, Di Angelantonio E, Sarwar N, Perry P, Kaptoge S, Ray KK, Thompson A, Wood AM, Lewington S, Sattar N, Packard CJ, Collins R, Thompson SG, Danesh J. Major lipids, apolipoproteins, and risk of vascular disease. JAMA, 2009; 302: 1993-2000
4) Weiss LA, Pan L, Abney M, Ober C. The sex-specific genetic architecture of quantitative traits in humans. Nat Genet, 2006; 38: 218-222
5) Kathiresan S, Srivastava D. Genetics of human cardiovascular disease. Cell, 2012; 148: 1242-1257
6) Gibson G. Rare and common variants: twenty arguments. Nat Rev Genet, 2012; 13: 135-145
7) Global Lipids Genetics Consortium, Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, Ganna A, Chen J, Buchkovich ML, Mora S, Beckmann JS, Bragg-Gresham JL, Chang HY, Demirkan A, Den Hertog HM, Do R, Donnelly LA, Ehret GB, Esko T, Feitosa MF, Ferreira T, Fischer K, Fontanillas P, Fraser RM, Freitag DF, Gurdasani D, Heikkilä K, Hyppönen E, Isaacs A, Jackson AU, Johansson A, Johnson T, Kaakinen M, Kettunen J, Kleber ME, Li X, Luan J, Lyytikäinen LP, Magnusson PK, Mangino M, Mihailov E, Montasser ME, Müller-Nurasyid M, Nolte IM, O'Connell JR, Palmer CD, Perola M, Petersen AK, Sanna S, Saxena R, Service SK, Shah S, Shungin D, Sidore C, Song C, Strawbridge RJ, Surakka I, Tanaka T, Teslovich TM, Thorleifsson G, Van den Herik EG, Voight BF, Volcik KA, Waite LL, Wong A, Wu Y, Zhang W, Absher D, Asiki G, Barroso I, Been LF, Bolton JL, Bonnycastle LL, Brambilla P, Burnett MS, Cesana G, Dimitriou M, Doney AS, Döring A, Elliott P, Epstein SE, Eyjolfsson GI, Gigante B, Goodarzi MO, Grallert H, Gravito ML, Groves CJ, Hallmans G, Hartikainen AL, Hayward C, Hernandez D, Hicks AA, Holm H, Hung YJ, Illig T, Jones MR, Kaleebu P, Kastelein JJ, Khaw KT, Kim E, Klopp N, Komulainen P, Kumari M, Langenberg C, Lehtimäki T, Lin SY, Lindström J, Loos RJ, Mach F, McArdle WL, Meisinger C, Mitchell BD, Müller G, Nagaraja R, Narisu N, Nieminen TV, Nsubuga RN, Olafsson I, Ong KK, Palotie A, Papamarkou T, Pomilla C, Pouta A, Rader DJ, Reilly MP, Ridker PM, Rivadeneira F, Rudan I, Ruokonen A, Samani N, Scharnagl H, Seeley J, Silander K, Stancáková A, Stirrups K, Swift AJ, Tiret L, Uitterlinden AG, van Pelt LJ, Vedantam S, Wainwright N, Wijmenga C, Wild SH, Willemsen G, Wilsgaard T, Wilson JF, Young EH, Zhao JH, Adair LS, Arveiler D, Assimes TL, Bandinelli S, Bennett F, Bochud M, Boehm BO, Boomsma DI, Borecki IB, Bornstein SR, Bovet P, Burnier M, Campbell H, Chakravarti A, Chambers JC, Chen YD, Collins FS, Cooper RS, Danesh J, Dedoussis G, de Faire U, Feranil AB, Ferrières J, Ferrucci L, Freimer NB, Gieger C, Groop LC, Gudnason V, Gyllensten U, Hamsten A, Harris TB, Hingorani A, Hirschhorn JN, Hofman A, Hovingh GK, Hsiung CA, Humphries SE, Hunt SC, Hveem K, Iribarren C, Järvelin MR, Jula A, Kähönen M, Kaprio J, Kesäniemi A, Kivimaki M, Kooner JS, Koudstaal PJ, Krauss RM, Kuh D, Kuusisto J, Kyvik KO, Laakso M, Lakka TA, Lind L, Lindgren CM, Martin NG, März W, McCarthy MI, McKenzie CA, Meneton P, Metspalu A, Moilanen L, Morris AD, Munroe PB, Njølstad I, Pedersen NL, Power C, Pramstaller PP, Price JF, Psaty BM, Quertermous T, Rauramaa R, Saleheen D, Salomaa V, Sanghera DK, Saramies J, Schwarz PE, Sheu WH, Shuldiner AR, Siegbahn A, Spector TD, Stefansson K, Strachan DP, Tayo BO, Tremoli E, Tuomilehto J, Uusitupa M, van Duijn CM, Vollenweider P, Wallentin L, Wareham NJ, Whitfield JB, Wolffenbuttel BH, Ordovas JM, Boerwinkle E, Palmer CN, Thorsteinsdottir U, Chasman DI, Rotter JI, Franks PW, Ripatti S, Cupples LA, Sandhu MS, Rich SS, Boehnke M, Deloukas P, Kathiresan S, Mohlke KL, Ingelsson E, Abecasis GR. Discovery and refinement of loci associated with lipid levels. Nat Genet, 2013; 45: 1274-1283
8) CARDIoGRAMplusC4D Consortium, Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR, Ingelsson E, Saleheen D, Erdmann J, Goldstein BA, Stirrups K, König IR, Cazier JB, Johansson A, Hall AS, Lee JY, Willer CJ, Chambers JC, Esko T, Folkersen L, Goel A, Grundberg E, Havulinna AS, Ho WK, Hopewell JC, Eriksson N, Kleber ME, Kristiansson K, Lundmark P, Lyytikäinen LP, Rafelt S, Shungin D, Strawbridge RJ, Thorleifsson G, Tikkanen E, Van Zuydam N, Voight BF, Waite LL, Zhang W, Ziegler A, Absher D, Altshuler D, Balmforth AJ, Barroso I, Braund PS, Burgdorf C, Claudi-Boehm S, Cox D, Dimitriou M, Do R; DIAGRAM Consortium; CARDIOGENICS Consortium, Doney AS, El Mokhtari N, Eriksson P, Fischer K, Fontanillas P, Franco-Cereceda A, Gigante B, Groop L, Gustafsson S, Hager J, Hallmans G, Han BG, Hunt SE, Kang HM, Illig T, Kessler T, Knowles JW, Kolovou G, Kuusisto J, Langenberg C, Langford C, Leander K, Lokki ML, Lundmark A, McCarthy MI, Meisinger C, Melander O, Mihailov E, Maouche S, Morris AD, Müller-Nurasyid M; MuTHER Consortium, Nikus K, Peden JF, Rayner NW, Rasheed A, Rosinger S, Rubin D, Rumpf MP, Schäfer A, Sivananthan M, Song C, Stewart AF, Tan ST, Thorgeirsson G, van der Schoot CE, Wagner PJ; Wellcome Trust Case Control Consortium, Wells GA, Wild PS, Yang TP, Amouyel P, Arveiler D, Basart H, Boehnke M, Boerwinkle E, Bram-

billa P, Cambien F, Cupples AL, de Faire U, Dehghan A, Diemert P, Epstein SE, Evans A, Ferrario MM, Ferrières J, Gauguier D, Go AS, Goodall AH, Gudnason V, Hazen SL, Holm H, Iribarren C, Jang Y, Kähönen M, Kee F, Kim HS, Klopp N, Koenig W, Kratzer W, Kuulasmaa K, Laakso M, Laaksonen R, Lee JY, Lind L, Ouwehand WH, Parish S, Park JE, Pedersen NL, Peters A, Quertermous T, Rader DJ, Salomaa V, Schadt E, Shah SH, Sinisalo J, Stark K, Stefansson K, Trégouët DA, Virtamo J, Wallentin L, Wareham N, Zimmermann ME, Nieminen MS, Hengstenberg C, Sandhu MS, Pastinen T, Syvänen AC, Hovingh GK, Dedoussis G, Franks PW, Lehtimäki T, Metspalu A, Zalloua PA, Siegbahn A, Schreiber S, Ripatti S, Blankenberg SS, Perola M, Clarke R, Boehm BO, O'Donnell C, Reilly MP, März W, Collins R, Kathiresan S, Hamsten A, Kooner JS, Thorsteinsdottir U, Danesh J, Palmer CN, Roberts R, Watkins H, Schunkert H, Samani NJ. Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet, 2013; 45: 25-33

9) Cohen JC, Boerwinkle E, Mosley TH Jr, Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. N Engl J Med, 2006; 354: 1264-1272
10) Cohen JC, Kiss RS, Pertsemlidis A, Marcel YL, McPherson R, Hobbs HH. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. Science, 2004; 305: 869-872
11) Johansen CT, Wang J, McIntyre AD, Martins RA, Ban MR, Lanktree MB, Huff MW, Péterfy M, Mehrabian M, Lusis AJ, Kathiresan S, Anand SS, Yusuf S, Lee AH, Glimcher LH, Cao H, Hegele RA. Excess of rare variants in non-genome-wide association study candidate genes in patients with hypertriglyceridemia. Circ Cardiovasc Genet, 2012; 5: 66-72
12) Wagner MJ. Rare-variant genome-wide association studies: a new frontier in genetic analysis of complex traits. Pharmacogenomics, 2013; 14: 413-424
13) Nelson MR, Wegmann D, Ehm MG, Kessner D, St Jean P, Verzilli C, Shen J, Tang Z, Bacanu SA, Fraser D, Warren L, Aponte J, Zawistowski M, Liu X, Zhang H, Zhang Y, Li J, Li Y, Li L, Woollard P, Topp S, Hall MD, Nangle K, Wang J, Abecasis G, Cardon LR, Zöllner S, Whittaker JC, Chissoe SL, Novembre J, Mooser V. An abundance of rare functional variants in 202 drug target genes sequenced in 14,002 people. Science, 2012; 337: 100-104
14) Tennessen JA1, Bigham AW, O'Connor TD, Fu W, Kenny EE, Gravel S, McGee S, Do R, Liu X, Jun G, Kang HM, Jordan D, Leal SM, Gabriel S, Rieder MJ, Abecasis G, Altshuler D, Nickerson DA, Boerwinkle E, Sunyaev S, Bustamante CD, Bamshad MJ, Akey JM; Broad GO; Seattle GO; NHLBI Exome Sequencing Project. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. Science, 2012; 337: 64-69
15) Keinan A, Clark AG. Recent explosive human population growth has resulted in an excess of rare genetic variants. Science, 2012; 336: 740-743
16) Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science, 1996; 273: 1516-1517
17) International HapMap Consortium. The International HapMap Project. Nature, 2003 18; 426: 789-796
18) Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature, 2007; 447: 661-678
19) Manolio TA. Bringing genome-wide association findings into clinical use. Nat Rev Genet, 2013; 14: 549-558
20) Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical Research, Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, Hughes TE, Groop L, Altshuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Bengtsson Boström K, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nilsson P, Orho-Melander M, Råstam L, Speliotes EK, Taskinen MR, Tuomi T, Guiducci C, Berglund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjögren M, Sterner M, Surti A, Svensson M, Svensson M, Tewhey R, Blumenstiel B, Parkin M, Defelice M, Barry R, Brodeur W, Camarata J, Chia N, Fava M, Gibbons J, Handsaker B, Healy C, Nguyen K, Gates C, Sougnez C, Gage D, Nizzari M, Gabriel SB, Chirn GW, Ma Q, Parikh H, Richardson D, Ricke D, Purcell S. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science, 2007; 316: 1331-1336
21) Kathiresan S, Melander O, Guiducci C, Surti A, Burtt NP, Rieder MJ, Cooper GM, Roos C, Voight BF, Havulinna AS, Wahlstrand B, Hedner T, Corella D, Tai ES, Ordovas JM, Berglund G, Vartiainen E, Jousilahti P, Hedblad B, Taskinen MR, Newton-Cheh C, Salomaa V, Peltonen L, Groop L, Altshuler DM, Orho-Melander M. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. Nat Genet, 2008; 40: 189-197
22) Kathiresan S, Willer CJ, Peloso GM, Demissie S, Musunuru K, Schadt EE, Kaplan L, Bennett D, Li Y, Tanaka T, Voight BF, Bonnycastle LL, Jackson AU, Crawford G, Surti A, Guiducci C, Burtt NP, Parish S, Clarke R, Zelenika D, Kubalanza KA, Morken MA, Scott LJ, Stringham HM, Galan P, Swift AJ, Kuusisto J, Bergman RN, Sundvall J, Laakso M, Ferrucci L, Scheet P, Sanna S, Uda M, Yang Q, Lunetta KL, Dupuis J, de Bakker PI, O'Donnell CJ, Chambers JC, Kooner JS, Hercberg S, Meneton P, Lakatta EG, Scuteri A, Schlessinger D, Tuomilehto J, Collins FS, Groop L, Altshuler D, Collins R, Lathrop GM, Melander O, Salomaa V, Peltonen L, Orho-Melander M, Ordovas JM, Boehnke M, Abecasis GR, Mohlke KL, Cupples LA. Common variants at 30 loci contribute to polygenic dyslipidemia. Nat Genet, 2009; 41: 56-65
23) Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, Pirruccello JP, Ripatti S, Chasman DI, Willer CJ, Johansen CT, Fouchier SW, Isaacs A, Peloso GM, Barbalic M, Ricketts SL, Bis JC, Aulchenko YS, Thorleifsson G, Feitosa MF, Chambers J, Orho-Melander M, Melander O, Johnson T, Li X, Guo X, Li M, Shin Cho Y, Jin Go M, Jin Kim Y, Lee JY, Park T, Kim K, Sim X, Twee-Hee Ong R, Croteau-Chonka DC, Lange LA, Smith JD, Song K, Hua Zhao J, Yuan X, Luan J, Lamina C, Ziegler A, Zhang W, Zee RY, Wright AF, Wit-

teman JC, Wilson JF, Willemsen G, Wichmann HE, Whitfield JB, Waterworth DM, Wareham NJ, Waeber G, Vollenweider P, Voight BF, Vitart V, Uitterlinden AG, Uda M, Tuomilehto J, Thompson JR, Tanaka T, Surakka I, Stringham HM, Spector TD, Soranzo N, Smit JH, Sinisalo J, Silander K, Sijbrands EJ, Scuteri A, Scott J, Schlessinger D, Sanna S, Salomaa V, Saharinen J, Sabatti C, Ruokonen A, Rudan I, Rose LM, Roberts R, Rieder M, Psaty BM, Pramstaller PP, Pichler I, Perola M, Penninx BW, Pedersen NL, Pattaro C, Parker AN, Pare G, Oostra BA, O'Donnell CJ, Nieminen MS, Nickerson DA, Montgomery GW, Meitinger T, McPherson R, McCarthy MI, McArdle W, Masson D, Martin NG, Marroni F, Mangino M, Magnusson PK, Lucas G, Luben R, Loos RJ, Lokki ML, Lettre G, Langenberg C, Launer LJ, Lakatta EG, Laaksonen R, Kyvik KO, Kronenberg F, König IR, Khaw KT, Kaprio J, Kaplan LM, Johansson A, Jarvelin MR, Janssens AC, Ingelsson E, Igl W, Kees Hovingh G, Hottenga JJ, Hofman A, Hicks AA, Hengstenberg C, Heid IM, Hayward C, Havulinna AS, Hastie ND, Harris TB, Haritunians T, Hall AS, Gyllensten U, Guiducci C, Groop LC, Gonzalez E, Gieger C, Freimer NB, Ferrucci L, Erdmann J, Elliott P, Ejebe KG, Döring A, Dominiczak AF, Demissie S, Deloukas P, de Geus EJ, de Faire U, Crawford G, Collins FS, Chen YD, Caulfield MJ, Campbell H, Burtt NP, Bonnycastle LL, Boomsma DI, Boekholdt SM, Bergman RN, Barroso I, Bandinelli S, Ballantyne CM, Assimes TL, Quertermous T, Altshuler D, Seielstad M, Wong TY, Tai ES, Feranil AB, Kuzawa CW, Adair LS, Taylor HA Jr, Borecki IB, Gabriel SB, Wilson JG, Holm H, Thorsteinsdottir U, Gudnason V, Krauss RM, Mohlke KL, Ordovas JM, Munroe PB, Kooner JS, Tall AR, Hegele RA, Kastelein JJ, Schadt EE, Rotter JI, Boerwinkle E, Strachan DP, Mooser V, Stefansson K, Reilly MP, Samani NJ, Schunkert H, Cupples LA, Sandhu MS, Ridker PM, Rader DJ, van Duijn CM, Peltonen L, Abecasis GR, Boehnke M, Kathiresan S. Biological, clinical and population relevance of 95 loci for blood lipids. Nature, 2010; 466: 707-713

24) Do R, Willer CJ, Schmidt EM, Sengupta S, Gao C, Peloso GM, Gustafsson S, Kanoni S, Ganna A, Chen J, Buchkovich ML, Mora S, Beckmann JS, Bragg-Gresham JL, Chang HY, Demirkan A, Den Hertog HM, Donnelly LA, Ehret GB, Esko T, Feitosa MF, Ferreira T, Fischer K, Fontanillas P, Fraser RM, Freitag DF, Gurdasani D, Heikkilä K, Hyppönen E, Isaacs A, Jackson AU, Johansson A, Johnson T, Kaakinen M, Kettunen J, Kleber ME, Li X, Luan J, Lyytikäinen LP, Magnusson PK, Mangino M, Mihailov E, Montasser ME, Müller-Nurasyid M, Nolte IM, O'Connell JR, Palmer CD, Perola M, Petersen AK, Sanna S, Saxena R, Service SK, Shah S, Shungin D, Sidore C, Song C, Strawbridge RJ, Surakka I, Tanaka T, Teslovich TM, Thorleifsson G, Van den Herik EG, Voight BF, Volcik KA, Waite LL, Wong A, Wu Y, Zhang W, Absher D, Asiki G, Barroso I, Been LF, Bolton JL, Bonnycastle LL, Brambilla P, Burnett MS, Cesana G, Dimitriou M, Doney AS, Döring A, Elliott P, Epstein SE, Eyjolfsson GI, Gigante B, Goodarzi MO, Grallert H, Gravito ML, Groves CJ, Hallmans G, Hartikainen AL, Hayward C, Hernandez D, Hicks AA, Holm H, Hung YJ, Illig T, Jones MR, Kaleebu P, Kastelein JJ, Khaw KT, Kim E, Klopp N, Komulainen P, Kumari M, Langenberg C, Lehtimäki T, Lin SY, Lindström J, Loos RJ, Mach F, McArdle WL, Meisinger C, Mitchell BD, Müller G, Nagaraja R, Narisu N, Nieminen TV, Nsubuga RN, Olafsson I, Ong KK, Palotie A, Papamarkou T, Pomilla C, Pouta A, Rader DJ, Reilly MP, Ridker PM, Rivadeneira F, Rudan I, Ruokonen A, Samani N, Scharnagl H, Seeley J, Silander K, Stančáková A, Stirrups K, Swift AJ, Tiret L, Uitterlinden AG, van Pelt LJ, Vedantam S, Wainwright N, Wijmenga C, Wild SH, Willemsen G, Wilsgaard T, Wilson JF, Young EH, Zhao JH, Adair LS, Arveiler D, Assimes TL, Bandinelli S, Bennett F, Bochud M, Boehm BO, Boomsma DI, Borecki IB, Bornstein SR, Bovet P, Burnier M, Campbell H, Chakravarti A, Chambers JC, Chen YD, Collins FS, Cooper RS, Danesh J, Dedoussis G, de Faire U, Feranil AB, Ferrières J, Ferrucci L, Freimer NB, Gieger C, Groop LC, Gudnason V, Gyllensten U, Hamsten A, Harris TB, Hingorani A, Hirschhorn JN, Hofman A, Hovingh GK, Hsiung CA, Humphries SE, Hunt SC, Hveem K, Iribarren C, Järvelin MR, Jula A, Kähönen M, Kaprio J, Kesäniemi A, Kivimaki M, Kooner JS, Koudstaal PJ, Krauss RM, Kuh D, Kuusisto J, Kyvik KO, Laakso M, Lakka TA, Lind L, Lindgren CM, Martin NG, März W, McCarthy MI, McKenzie CA, Meneton P, Metspalu A, Moilanen L, Morris AD, Munroe PB, Njølstad I, Pedersen NL, Power C, Pramstaller PP, Price JF, Psaty BM, Quertermous T, Rauramaa R, Saleheen D, Salomaa V, Sanghera DK, Saramies J, Schwarz PE, Sheu WH, Shuldiner AR, Siegbahn A, Spector TD, Stefansson K, Strachan DP, Tayo BO, Tremoli E, Tuomilehto J, Uusitupa M, van Duijn CM, Vollenweider P, Wallentin L, Wareham NJ, Whitfield JB, Wolffenbuttel BH, Altshuler D, Ordovas JM, Boerwinkle E, Palmer CN, Thorsteinsdottir U, Chasman DI, Rotter JI, Franks PW, Ripatti S, Cupples LA, Sandhu MS, Rich SS, Boehnke M, Deloukas P, Mohlke KL, Ingelsson E, Abecasis GR, Daly MJ, Neale BM, Kathiresan S. Common variants associated with plasma triglycerides and risk for coronary artery disease. Nat Genet, 2013; 45: 1345-1352

25) Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM. Finding the missing heritability of complex diseases. Nature, 2009; 461: 747-753

26) Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. Nat Genet, 2008; 40: 695-701

27) Gibson G. Rare and common variants: twenty arguments. Nat Rev Genet, 2012; 13: 135-145

28) Liu DJ, Leal SM. Estimating genetic effects and quantifying missing heritability explained by identified rare-variant associations. Am J Hum Genet, 2012; 91: 585-596

29) Online Mendelian Inheritance in Man, OMIM®. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD), {7 November 2014}. World Wide Web URL: http: //omim.org/

30) Kawashiri MA, Hayashi K, Konno T, Fujino N, Ino H, Yamagishi M. Current perspectives in genetic cardiovascular disorders: from basic to clinical aspects. Heart Vessels, 2014; 29: 129-141
31) Brunham LR, Hayden MR. Hunting human disease genes: lessons from the past, challenges for the future. Hum Genet, 2013; 132: 603-617
32) Fu W, O'Connor TD, Jun G, Kang HM, Abecasis G, Leal SM, Gabriel S, Rieder MJ, Altshuler D, Shendure J, Nickerson DA, Bamshad MJ; NHLBI Exome Sequencing Project, Akey JM. Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. Nature, 2013; 493: 216-220
33) Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. Nat Genet, 2008; 40: 695-701
34) http: //genome.sph.umich.edu/wiki/Exome_Chip_Design
35) Huyghe JR, Jackson AU, Fogarty MP, Buchkovich ML, Stančáková A, Stringham HM, Sim X, Yang L, Fuchsberger C, Cederberg H, Chines PS, Teslovich TM, Romm JM, Ling H, McMullen I, Ingersoll R, Pugh EW, Doheny KF, Neale BM, Daly MJ, Kuusisto J, Scott LJ, Kang HM, Collins FS, Abecasis GR, Watanabe RM, Boehnke M, Laakso M, Mohlke KL. Exome array analysis identifies new loci and low-frequency variants influencing insulin processing and secretion. Nat Genet, 2013; 45: 197-201
36) Dahl F, Stenberg J, Fredriksson S, Welch K, Zhang M, Nilsson M, Bicknell D, Bodmer WF, Davis RW, Ji H. Multigene amplification and massively parallel sequencing for cancer mutation discovery. Proc Natl Acad Sci U S A, 2007; 104: 9387-9392
37) Okou DT, Steinberg KM, Middle C, Cutler DJ, Albert TJ, Zwick ME. Microarray-based genomic selection for high-throughput resequencing. Nat Methods, 2007; 4: 907-909
38) Porreca GJ, Zhang K, Li JB, Xie B, Austin D, Vassallo SL, LeProust EM, Peck BJ, Emig CJ, Dahl F, Gao Y, Church GM, Shendure J. Multiplex amplification of large sets of human exons. Nat Methods, 2007; 4: 931-936
39) Gnirke A, Melnikov A, Maguire J, Rogov P, LeProust EM, Brockman W, Fennell T, Giannoukos G, Fisher S, Russ C, Gabriel S, Jaffe DB, Lander ES, Nusbaum C. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nat Biotechnol, 2009; 27: 182-189
40) Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, Shaffer T, Wong M, Bhattacharjee A, Eichler EE, Bamshad M, Nickerson DA, Shendure J. Targeted capture and massively parallel sequencing of 12 human exomes. Nature, 2009; 461: 272-276
41) Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, Howard E, Shendure J, Turner DJ. Target-enrichment strategies for next-generation sequencing. Nat Methods, 2010; 7: 111-118
42) Stitziel NO, Peloso GM, Abifadel M, Cefalu AB, Fouchier S, Motazacker MM, Tada H, Larach DB, Awan Z, Haller JF, Pullinger CR, Varret M, Rabès JP, Noto D, Tarugi P, Kawashiri MA, Nohara A, Yamagishi M, Risman M, Deo R, Ruel I, Shendure J, Nickerson DA, Wilson JG, Rich SS, Gupta N, Farlow DN, Neale BM, Daly MJ, Kane JP, Freeman MW, Genest J, Rader DJ, Mabuchi H, Kastelein JJ, Hovingh GK, Averna MR, Gabriel S, Boileau C, Kathiresan S. Exome sequencing in suspected monogenic dyslipidemias. Circ Cardiovasc Genet, 2015; 8: 343-350
43) Albrechtsen A, Grarup N, Li Y, Sparsø T, Tian G, Cao H, Jiang T, Kim SY, Korneliussen T, Li Q, Nie C, Wu R, Skotte L, Morris AP, Ladenvall C, Cauchi S, Stančáková A, Andersen G, Astrup A, Banasik K, Bennett AJ, Bolund L, Charpentier G, Chen Y, Dekker JM, Doney AS, Dorkhan M, Forsen T, Frayling TM, Groves CJ, Gui Y, Hallmans G, Hattersley AT, He K, Hitman GA, Holmkvist J, Huang S, Jiang H, Jin X, Justesen JM, Kristiansen K, Kuusisto J, Lajer M, Lantieri O, Li W, Liang H, Liao Q, Liu X, Ma T, Ma X, Manijak MP, Marre M, Mokrosiński J, Morris AD, Mu B, Nielsen AA, Nijpels G, Nilsson P, Palmer CN, Rayner NW, Renström F, Ribel-Madsen R, Robertson N, Rolandsson O, Rossing P, Schwartz TW; D.E.S.I.R. Study Group, Slagboom PE, Sterner M; DIAGRAM Consortium, Tang M, Tarnow L, Tuomi T, van't Riet E, van Leeuwen N, Varga TV, Vestmar MA, Walker M, Wang B, Wang Y, Wu H, Xi F, Yengo L, Yu C, Zhang X, Zhang J, Zhang Q, Zhang W, Zheng H, Zhou Y, Altshuler D, 't Hart LM, Franks PW, Balkau B, Froguel P, McCarthy MI, Laakso M, Groop L, Christensen C, Brandslund I, Lauritzen T, Witte DR, Linneberg A, Jørgensen T, Hansen T, Wang J, Nielsen R, Pedersen O. Exome sequencing-driven discovery of coding polymorphisms associated with common metabolic phenotypes. Diabetologia, 2013; 56: 298-310
44) Kiezun A, Garimella K, Do R, Stitziel NO, Neale BM, McLaren PJ, Gupta N, Sklar P, Sullivan PF, Moran JL, Hultman CM, Lichtenstein P, Magnusson P, Lehner T, Shugart YY, Price AL, de Bakker PI, Purcell SM, Sunyaev SR. Exome sequencing and the genetic basis of complex traits. Nat Genet, 2012; 44: 623-630
45) Tada H, Kawashiri MA, Nohara A, Saito R, Tanaka Y, Nomura A, Konno T, Sakata K, Fujino N, Takamura T, Inazu A, Mabuchi H, Yamagishi M, Hayashi K. Whole exome sequencing combined with integrated variant annotation prediction identifies asymptomatic Tangier disease with compound heterozygous mutations in ABCA1 gene. Atherosclerosis, 2015; 240: 324-329
46) Li, B. and Leal, S.M. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. Am J Hum Genet, 2008; 83: 311-321
47) Neale BM1, Rivas MA, Voight BF, Altshuler D, Devlin B, Orho-Melander M, Kathiresan S, Purcell SM, Roeder K, Daly MJ. Testing for an unusual distribution of rare variants. PLoS Genet, 2011; 7: e1001322
48) Morris, A.P, Zeggini, E. An evaluation of statistical approaches to rare variant analysis in genetic association studies. Genet Epidemiol, 2010; 34: 188-193
49) Do R, Kathiresan S, Abecasis GR. Exome sequencing and complex disease: practical aspects of rare variant association studies. Hum Mol Genet, 2012; 21: R1-9
50) Kotowski IK, Pertsemlidis A, Luke A, Cooper RS, Vega GL, Cohen JC, Hobbs HH. A spectrum of PCSK9 alleles contributes to plasma levels of low-density lipoprotein

cholesterol. Am J Hum Genet, 2006; 78: 410-422
51) Morita H, Larson MG, Barr SC, Vasan RS, O'Donnell CJ, Hirschhorn JN. Single-gene mutations and increased left ventricular wall thickness in the community: the Framingham Heart Study. Circulation, 2006; 113: 2697-2705
52) Ji W, Foo JN, O'Roak BJ, Zhao H, Larson MG, Simon DB, Newton-Cheh C, State MW, Levy D, Lifton RP. Rare independent mutations in renal salt handling genes contribute to blood pressure variation. Nat Genet, 2008; 40: 592-599
53) Nejentsev S, Walker N, Riches D, Egholm M, Todd JA. Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. Science, 2009; 324: 387-389
54) Weeke P, Parvez B, Blair M, Short L, Ingram C, Kucera G. Candidate Gene Approach to Identifying Rare Genetic Variants Associated with Lone Atrial Fibrillation. Heart Rhythm, 2014; 11: 46-52
55) Li B, Liu DJ, Leal SM. Identifying rare variants associated with complex traits via sequencing. Curr Protoc Hum Genet, 2013 Jul; Chapter 1: Unit 1.26
56) Wu MC, Kraft P, Epstein MP, Taylor DM, Chanock SJ, Hunter DJ, Lin X. Powerful SNP-set analysis for case-control genome-wide association studies. Am J Hum Genet, 2010; 86: 929-942
57) Price AL, Kryukov GV, de Bakker PI, Purcell SM, Staples J, Wei LJ, Sunyaev SR. Pooled association tests for rare variants in exon-resequencing studies. Am J Hum Genet, 2010; 86: 832-848
58) Lee S, Emond MJ, Bamshad MJ, Barnes KC, Rieder MJ, Nickerson DA; NHLBI GO Exome Sequencing Project—ESP Lung Project Team, Christiani DC, Wurfel MM, Lin X. Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. Am J Hum Genet, 2012; 91: 224-237
59) van Leeuwen EM, Karssen LC, Deelen J, Isaacs A, Medina-Gomez C, Mbarek H, Kanterakis A, Trompet S, Postmus I, Verweij N, van Enckevort DJ, Huffman JE, White CC, Feitosa MF, Bartz TM, Manichaikul A, Joshi PK, Peloso GM, Deelen P, van Dijk F, Willemsen G, de Geus EJ, Milaneschi Y, Penninx BW, Francioli LC, Menelaou A, Pulit SL, Rivadeneira F, Hofman A, Oostra BA, Franco OH, Mateo Leach I, Beekman M, de Craen AJ, Uh HW, Trochet H, Hocking LJ, Porteous DJ, Sattar N, Packard CJ, Buckley BM, Brody JA, Bis JC, Rotter JI, Mychaleckyj JC, Campbell H, Duan Q, Lange LA, Wilson JF, Hayward C, Polasek O, Vitart V, Rudan I, Wright AF, Rich SS, Psaty BM, Borecki IB, Kearney PM, Stott DJ, Adrienne Cupples L; Genome of The Netherlands Consortium, Jukema JW, van der Harst P, Sijbrands EJ, Hottenga JJ, Uitterlinden AG, Swertz MA, van Ommen GJ, de Bakker PI, Eline Slagboom P, Boomsma DI, Wijmenga C, van Duijn CM. Genome of The Netherlands population-specific imputations identify an ABCA6 variant associated with cholesterol levels. Nat Commun, 2015; 6: 6065
60) Lange LA, Hu Y, Zhang H, Xue C, Schmidt EM, Tang ZZ, Bizon C, Lange EM, Smith JD, Turner EH, Jun G, Kang HM, Peloso G, Auer P, Li KP, Flannick J, Zhang J, Fuchsberger C, Gaulton K, Lindgren C, Locke A, Manning A, Sim X, Rivas MA, Holmen OL, Gottesman O, Lu Y, Ruderfer D, Stahl EA, Duan Q, Li Y, Durda P, Jiao S, Isaacs A, Hofman A, Bis JC, Correa A, Griswold ME, Jakobsdottir J, Smith AV, Schreiner PJ, Feitosa MF, Zhang Q, Huffman JE, Crosby J, Wassel CL, Do R, Franceschini N, Martin LW, Robinson JG, Assimes TL, Crosslin DR, Rosenthal EA, Tsai M, Rieder MJ, Farlow DN, Folsom AR, Lumley T, Fox ER, Carlson CS, Peters U, Jackson RD, van Duijn CM, Uitterlinden AG, Levy D, Rotter JI, Taylor HA, Gudnason V Jr, Siscovick DS, Fornage M, Borecki IB, Hayward C, Rudan I, Chen YE, Bottinger EP, Loos RJ, Sætrom P, Hveem K, Boehnke M, Groop L, McCarthy M, Meitinger T, Ballantyne CM, Gabriel SB, O'Donnell CJ, Post WS, North KE, Reiner AP, Boerwinkle E, Psaty BM, Altshuler D, Kathiresan S, Lin DY, Jarvik GP, Cupples LA, Kooperberg C, Wilson JG, Nickerson DA, Abecasis GR, Rich SS, Tracy RP, Willer CJ; NHLBI Grand Opportunity Exome Sequencing Project. Whole-exome sequencing identifies rare and low-frequency coding variants associated with LDL cholesterol. Am J Hum Genet, 2014; 94: 233-245
61) Do R, Stitziel NO, Won HH, Jørgensen AB, Duga S, Angelica Merlini P, Kiezun A, Farrall M, Goel A, Zuk O, Guella I, Asselta R, Lange LA, Peloso GM, Auer PL; NHLBI Exome Sequencing Project, Girelli D, Martinelli N, Farlow DN, DePristo MA, Roberts R, Stewart AF, Saleheen D, Danesh J, Epstein SE, Sivapalaratnam S, Hovingh GK, Kastelein JJ, Samani NJ, Schunkert H, Erdmann J, Shah SH, Kraus WE, Davies R, Nikpay M, Johansen CT, Wang J, Hegele RA, Hechter E, Marz W, Kleber ME, Huang J, Johnson AD, Li M, Burke GL, Gross M, Liu Y, Assimes TL, Heiss G, Lange EM, Folsom AR, Taylor HA, Olivieri O, Hamsten A, Clarke R, Reilly DF, Yin W, Rivas MA, Donnelly P, Rossouw JE, Psaty BM, Herrington DM, Wilson JG, Rich SS, Bamshad MJ, Tracy RP, Cupples LA, Rader DJ, Reilly MP, Spertus JA, Cresci S, Hartiala J, Tang WH, Hazen SL, Allayee H, Reiner AP, Carlson CS, Kooperberg C, Jackson RD, Boerwinkle E, Lander ES, Schwartz SM, Siscovick DS, McPherson R, Tybjaerg-Hansen A, Abecasis GR, Watkins H, Nickerson DA, Ardissino D, Sunyaev SR, O'Donnell CJ, Altshuler D, Gabriel S, Kathiresan S. Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. Nature, 2015; 518: 102-106
62) Surakka I, Horikoshi M, Mägi R, Sarin AP, Mahajan A, Lagou V, Marullo L, Ferreira T, Miraglio B, Timonen S, Kettunen J, Pirinen M, Karjalainen J, Thorleifsson G, Hägg S, Hottenga JJ, Isaacs A, Ladenvall C, Beekman M, Esko T, Ried JS, Nelson CP, Willenborg C, Gustafsson S, Westra HJ, Blades M, de Craen AJ, de Geus EJ, Deelen J, Grallert H, Hamsten A, Havulinna AS, Hengstenberg C, Houwing-Duistermaat JJ, Hyppönen E, Karssen LC, Lehtimäki T, Lyssenko V, Magnusson PK, Mihailov E, Müller-Nurasyid M, Mpindi JP, Pedersen NL, Penninx BW, Perola M, Pers TH, Peters A, Rung J, Smit JH, Steinthorsdottir V, Tobin MD, Tsernikova N, van Leeuwen EM, Viikari JS, Willems SM, Willemsen G, Schunkert H,

Erdmann J, Samani NJ, Kaprio J, Lind L, Gieger C, Metspalu A, Slagboom PE, Groop L, van Duijn CM, Eriksson JG, Jula A, Salomaa V, Boomsma DI, Power C, Raitakari OT, Ingelsson E, Järvelin MR, Thorsteinsdottir U, Franke L, Ikonen E, Kallioniemi O, Pietiäinen V, Lindgren CM, Stefansson K, Palotie A, McCarthy MI, Morris AP, Prokopenko I, Ripatti S; ENGAGE Consortium. The impact of low-frequency and rare variants on lipid levels. Nat Genet, 2015; 47: 589-597

63) Cohen JC. Emerging LDL therapies: Using human genetics to discover new therapeutic targets for plasma lipids. J Clin Lipidol, 2013; 7(3 Suppl): S1-5
64) Wetterau JR, Aggerbeck LP, Bouma ME, Eisenberg C, Munck A, Hermier M, Schmitz J, Gay G, Rader DJ, Gregg RE. Absence of microsomal triglyceride transfer protein in individuals with abetalipoproteinemia. Science, 1992; 258: 999-1001
65) Abifadel M, Varret M, Rabès JP, Allard D, Ouguerram K, Devillers M, Cruaud C, Benjannet S, Wickham L, Erlich D, Derré A, Villéger L, Farnier M, Beucler I, Bruckert E, Chambaz J, Chanu B, Lecerf JM, Luc G, Moulin P, Weissenbach J, Prat A, Krempf M, Junien C, Seidah NG, Boileau C. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. Nat Genet, 2003; 34: 154-156
66) Inazu A, Brown ML, Hesler CB, Agellon LB, Koizumi J, Takata K, Maruhama Y, Mabuchi H, Tall AR. Increased high-density lipoprotein levels caused by a common cholesteryl-ester transfer protein gene mutation. N Engl J Med, 1990; 323: 1234-1238
67) TG and HDL Working Group of the Exome Sequencing Project, National Heart, Lung, and Blood Institute, Crosby J, Peloso GM, Auer PL, Crosslin DR, Stitziel NO, Lange LA, Lu Y, Tang ZZ, Zhang H, Hindy G, Masca N, Stirrups K, Kanoni S, Do R, Jun G, Hu Y, Kang HM, Xue C, Goel A, Farrall M, Duga S, Merlini PA, Asselta R, Girelli D, Olivieri O, Martinelli N, Yin W, Reilly D, Speliotes E, Fox CS, Hveem K, Holmen OL, Nikpay M, Farlow DN, Assimes TL, Franceschini N, Robinson J, North KE, Martin LW, DePristo M, Gupta N, Escher SA, Jansson JH, Van Zuydam N, Palmer CN, Wareham N, Koch W, Meitinger T, Peters A, Lieb W, Erbel R, Konig IR, Kruppa J, Degenhardt F, Gottesman O, Bottinger EP, O'Donnell CJ, Psaty BM, Ballantyne CM, Abecasis G, Ordovas JM, Melander O, Watkins H, Orho-Melander M, Ardissino D, Loos RJ, McPherson R, Willer CJ, Erdmann J, Hall AS, Samani NJ, Deloukas P, Schunkert H, Wilson JG, Kooperberg C, Rich SS, Tracy RP, Lin DY, Altshuler D, Gabriel S, Nickerson DA, Jarvik GP, Cupples LA, Reiner AP, Boerwinkle E, Kathiresan S. Loss-of-Function Mutations in APOC3, Triglycerides, and Coronary Disease. N Engl J Med, 2014; 371: 22-31
68) Holmen OL, Zhang H, Fan Y, Hovelson DH, Schmidt EM, Zhou W, Guo Y, Zhang J, Langhammer A, Løchen ML, Ganesh SK, Vatten L, Skorpen F, Dalen H, Zhang J, Pennathur S, Chen J, Platou C, Mathiesen EB, Wilsgaard T, Njølstad I, Boehnke M, Chen YE, Abecasis GR, Hveem K, Willer CJ. Systematic evaluation of coding variation identifies a candidate causal variant in TM6SF2 influencing total cholesterol and myocardial infarction risk. Nat Genet, 2014; 46: 345-351
69) Peloso GM, Auer PL, Bis JC, Voorman A, Morrison AC, Stitziel NO, Brody JA, Khetarpal SA, Crosby JR, Fornage M, Isaacs A, Jakobsdottir J, Feitosa MF, Davies G, Huffman JE, Manichaikul A, Davis B, Lohman K, Joon AY, Smith AV, Grove ML, Zanoni P, Redon V, Demissie S, Lawson K, Peters U, Carlson C, Jackson RD, Ryckman KK, Mackey RH, Robinson JG, Siscovick DS, Schreiner PJ, Mychaleckyj JC, Pankow JS, Hofman A, Uitterlinden AG, Harris TB, Taylor KD, Stafford JM, Reynolds LM, Marioni RE, Dehghan A, Franco OH, Patel AP, Lu Y, Hindy G, Gottesman O, Bottinger EP, Melander O, Orho-Melander M, Loos RJ, Duga S, Merlini PA, Farrall M, Goel A, Asselta R, Girelli D, Martinelli N, Shah SH, Kraus WE, Li M, Rader DJ, Reilly MP, McPherson R, Watkins H, Ardissino D; NHLBI GO Exome Sequencing Project, Zhang Q, Wang J, Tsai MY, Taylor HA, Correa A, Griswold ME, Lange LA, Starr JM, Rudan I, Eiriksdottir G, Launer LJ, Ordovas JM, Levy D, Chen YD, Reiner AP, Hayward C, Polasek O, Deary IJ, Borecki IB, Liu Y, Gudnason V, Wilson JG, van Duijn CM, Kooperberg C, Rich SS, Psaty BM, Rotter JI, O'Donnell CJ, Rice K, Boerwinkle E, Kathiresan S, Cupples LA. Association of low-frequency and rare coding-sequence variants with blood lipids and coronary heart disease in 56,000 whites and blacks. Am J Hum Genet, 2014; 94: 223-232
70) Pardo LM, MacKay I, Oostra B, van Duijn CM, Aulchenko YS. The effect of genetic drift in a young genetically isolated population. Ann Hum Genet, 2005; 69: 288-295
71) Tachmazidou I, Dedoussis G, Southam L, Farmaki AE, Ritchie GR, Xifara DK, Matchan A, Hatzikotoulas K, Rayner NW, Chen Y, Pollin TI, O'Connell JR, Yerges-Armstrong LM, Kiagiadaki C, Panoutsopoulou K, Schwartzentruber J, Moutsianas L; UK10K consortium, Tsafantakis E, Tyler-Smith C, McVean G, Xue Y, Zeggini E. A rare functional cardioprotective APOC3 variant has risen in frequency in distinct population isolates. Nat Commun, 2013; 4: 2872
72) Roach JC, Glusman G, Smit AF, Huff CD, Hubley R, Shannon PT, Rowen L, Pant KP, Goodman N, Bamshad M, Shendure J, Drmanac R, Jorde LB, Hood L, Galas DJ. Analysis of genetic inheritance in a family quartet by whole-genome sequencing. Science, 2010; 328: 636-639
73) Sobreira NL, Cirulli ET, Avramopoulos D, Wohler E, Oswald GL, Stevens EL, Ge D, Shianna KV, Smith JP, Maia JM, Gumbs CE, Pevsner J, Thomas G, Valle D, Hoover-Fong JE, Goldstein DB. Whole-genome sequencing of a single proband together with linkage analysis identifies a Mendelian disease gene. PLoS Genet, 2010; 6: e1000991
74) McMillin MJ, Below JE, Shively KM, Beck AE, Gildersleeve HI, Pinner J, Gogola GR, Hecht JT, Grange DK, Harris DJ, Earl DL, Jagadeesh S, Mehta SG, Robertson SP, Swanson JM, Faustman EM, Mefford HC, Shendure J, Nickerson DA, Bamshad MJ; University of Washington Center for Mendelian Genomics. Mutations in ECEL1 cause distal arthrogryposis type 5D. Am J Hum Genet, 2013; 92: 150-156
75) Morrison AC, Voorman A, Johnson AD, Liu X, Yu J, Li A, Muzny D, Yu F, Rice K, Zhu C, Bis J, Heiss G,

O'Donnell CJ, Psaty BM, Cupples LA, Gibbs R, Boerwinkle E; Cohorts for Heart and Aging Research in Genetic Epidemiology (CHARGE) Consortium. Whole-genome sequence-based analysis of high-density lipoprotein cholesterol. Nat Genet, 2013; 45: 899-901

76) Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res, 2010; 38: e164
77) Yandell M, Huff C, Hu H, Singleton M, Moore B, Xing J, Jorde LB, Reese MG. A probabilistic disease-gene finder for personal genomes. Genome Res, 2011; 21: 1529-1542
78) Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin), 2012; 6: 80-92
79) Flicek P, Ahmed I, Amode MR, Barrell D, Beal K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fairley S, Fitzgerald S, Gil L, García-Girón C, Gordon L, Hourlier T, Hunt S, Juettemann T, Kähäri AK, Keenan S, Komorowska M, Kulesha E, Longden I, Maurel T, McLaren WM, Muffato M, Nag R, Overduin B, Pignatelli M, Pritchard B, Pritchard E, Riat HS, Ritchie GR, Ruffier M, Schuster M, Sheppard D, Sobral D, Taylor K, Thormann A, Trevanion S, White S, Wilder SP, Aken BL, Birney E, Cunningham F, Dunham I, Harrow J, Herrero J, Hubbard TJ, Johnson N, Kinsella R, Parker A, Spudich G, Yates A, Zadissa A, Searle SM. Ensembl 2013. Nucleic Acids Res, 2013; 41(Database issue): D48-55
80) Hoffman MM, Ernst J, Wilder SP, Kundaje A, Harris RS, Libbrecht M, Giardine B, Ellenbogen PM, Bilmes JA, Birney E, Hardison RC, Dunham I, Kellis M, Noble WS. Integrative annotation of chromatin elements from ENCODE data. Nucleic Acids Res, 2013; 41: 827-841
81) Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A, Kellis M, Marra MA, Beaudet AL, Ecker JR, Farnham PJ, Hirst M, Lander ES, Mikkelsen TS, Thomson JA. The NIH Roadmap Epigenomics Mapping Consortium. Nat Biotechnol, 2010; 28: 1045-1048
82) Kawai J, Shinagawa A, Shibata K, Yoshino M, Itoh M, Ishii Y, Arakawa T, Hara A, Fukunishi Y, Konno H, Adachi J, Fukuda S, Aizawa K, Izawa M, Nishi K, Kiyosawa H, Kondo S, Yamanaka I, Saito T, Okazaki Y, Gojobori T, Bono H, Kasukawa T, Saito R, Kadota K, Matsuda H, Ashburner M, Batalov S, Casavant T, Fleischmann W, Gaasterland T, Gissi C, King B, Kochiwa H, Kuehl P, Lewis S, Matsuo Y, Nikaido I, Pesole G, Quackenbush J, Schriml LM, Staubli F, Suzuki R, Tomita M, Wagner L, Washio T, Sakai K, Okido T, Furuno M, Aono H, Baldarelli R, Barsh G, Blake J, Boffelli D, Bojunga N, Carninci P, de Bonaldo MF, Brownstein MJ, Bult C, Fletcher C, Fujita M, Gariboldi M, Gustincich S, Hill D, Hofmann M, Hume DA, Kamiya M, Lee NH, Lyons P, Marchionni L, Mashima J, Mazzarelli J, Mombaerts P, Nordone P, Ring B, Ringwald M, Rodriguez I, Sakamoto N, Sasaki H, Sato K, Schönbach C, Seya T, Shibata Y, Storch KF, Suzuki H, Toyo-oka K, Wang KH, Weitz C, Whittaker C, Wilming L, Wynshaw-Boris A, Yoshida K, Hasegawa Y, Kawaji H, Kohtsuki S, Hayashizaki Y; RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection. Nature, 2001; 409: 685-690
83) Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell, 2006; 126: 663-676
84) MacArthur DG, Balasubramanian S, Frankish A, Huang N, Morris J, Walter K, Jostins L, Habegger L, Pickrell JK, Montgomery SB, Albers CA, Zhang ZD, Conrad DF, Lunter G, Zheng H, Ayub Q, DePristo MA, Banks E, Hu M, Handsaker RE, Rosenfeld JA, Fromer M, Jin M, Mu XJ, Khurana E, Ye K, Kay M, Saunders GI, Suner MM, Hunt T, Barnes IH, Amid C, Carvalho-Silva DR, Bignell AH, Snow C, Yngvadottir B, Bumpstead S, Cooper DN, Xue Y, Romero IG; 1000 Genomes Project Consortium, Wang J, Li Y, Gibbs RA, McCarroll SA, Dermitzakis ET, Pritchard JK, Barrett JC, Harrow J, Hurles ME, Gerstein MB, Tyler-Smith C. A systematic survey of loss-of-function variants in human protein-coding genes. Science, 2012; 335: 823-828