Cardiovascular Genomics, Personalized Medicine and the National Heart, Lung and Blood Institute: Part I, the Beginning of an Era

Christopher J. O'Donnell, M.D. M.P.H.[1,2,3] and **Elizabeth G. Nabel, M.D.**[3]

[1]Center for Population Studies in the Division of Intramural Research, National Heart, Lung, and Blood Institute

[2]Framingham Heart Study, National Heart, Lung, and Blood Institute

[3]Office of the Director, National Heart, Lung, and Blood Institute

The inaugural issue of *Circulation Genetics* arrives at a remarkable time in the history of genetic research and cardiovascular medicine. Despite the tremendous progress in knowledge, cardiovascular disease remains the leading cause of death in the United States[1], and it has overcome infectious diseases as the leading cause worldwide.[2] In addition, rates of cardiovascular disease (CVD) remain higher in African American and Hispanic populations in the US.[1] The recent Strategic Plan of the National Heart, Lung and Blood Institute (NHLBI) emphasizes research areas to fill the significant knowledge gaps needed to improve the diagnosis, treatment and control of known risk factors and clinically apparent disease. Simultaneously, the NHLBI Strategic Plan recognizes a tremendous opportunity that is available for use of genetic and genomic research to generate new knowledge that might reduce the morbidity and mortality from CVD in US populations.[3] Public availability of vast amounts of detailed sequence information about the human genome, completed sequence data on dozens of other animal genomes, and the private sector development of high throughput genetic technologies has transformed in a few short years the conduct of cardiovascular genetics and genomics research from a primary focus on Mendelian disorders to a current emphasis on genomewide association studies (GWAS) (Figure 1). In this review, we describe the rationale for the current emphasis on largescale genomic studies, we summarize the evolving approaches and progress to date, and we identify immediate-term research needs. The National Institutes of Health (NIH) and the NHLBI are supporting a portfolio of largescale genetic and genomic programs in diverse US populations with the longer term objective of translating knowledge into the prediction, prevention and pre-emption of CVD, as well as lung, sleep, and blood disorders. Underlying this portfolio is a strong commitment to make available participant level data and aggregate research results to the broad community of investigators, while protecting the privacy and confidentiality and respecting the informed consent of study participants.

Rare genetic variation in CVD

Much of our current understanding regarding the genetic predisposition of human diseases derives from detailed studies, during the "pre-genome" era, of large pedigrees harboring rare, Mendelian conditions. The critical role of clinical observation in such studies cannot be

overemphasized, since the identification of rare conditions usually begins with the assembly of pedigrees in which an autosomally dominant or recessive pattern of inheritance is identified. Pedigree studies using positional cloning approaches have led to a comprehensive body of information regarding the role of deleterious mutations, often in exons (the protein coding region of human genes) leading to altered protein function or number. These studies have uncovered rare (minor allele frequency [MAF] generally << 0.1%) genetic mutations in the sequence encoding key sarcomere and storage disorder proteins underlying hypertrophic and dilated cardiomyopathy, in ion channels underlying long-QT syndrome and ventricular tachycardia, and in *FBN1* underlying thoracic aortic aneurysms and aortic dissection.[4] Knowledge regarding causal mutations has contributed tremendously to our understanding of the pathophysiology of these conditions and in some cases, such as familial hypercholesterolemia, has led to the development of new treatments, such as lipid lowering drug therapies, that are now widely used in general populations. However, while genetic testing for some of the above conditions is now available in Clinical Laboratory Improvement Amendments (CLIA)-certified laboratories, in most cases there is not yet a clear role in clinical practice for the genetic testing of mutations in cardiovascular diagnosis or treatment.

One example of remarkable progress in the past several years has been our understanding of the molecular genetics of the rare, premature aging condition, Hutchinson-Gilford progeria syndrome (Progeria). Progeria results from a spontaneous, de novo point mutation in exon 11 of the LMNA gene, producing an alternatively spliced form of prelamin A, called progerin.[5] In normal cells, prelamin A undergoes a series of posttranslational modifications, including farnesylation and enzymatic cleavage of terminal 15 amino acids and the farnesyl group, allowing mature lamin A to migrate from the nuclear membrane to the nuclear lamina. In contrast, because of the alternate splicing which internally deletes 50 amino acids, progerin lacks the second cleavage site, remains permanently farnestylated, and adheres to the nuclear membrane, disrupting nuclear trafficking and lamina function. Children with progeria uniformly die of heart attack and stroke in their early to mid teens, due to a drop out of vascular smooth muscle cells (vsmcs) in the media of arteries and veins with maladaptive vascular remodeling, fibrosis and luminal narrowing.[6,7] Treatment of progerin cells in culture and transgenic progerin mice with a farnesyl transferase inhibitor (FTI) restores nuclear architecture and function, and prevents development and progression of the vascular disease in progerin mice.[8] Accordingly, a clinical trial has been initiated to study the effect of FTI treatment on the growth curve and CV disease in children with progeria. The role of progerin on vsmc cell cycle and proliferation, progenitor cell renewal and normal aging, all features of common CV disease, are under investigation.

In contrast to rare Mendelian conditions, common, "complex" conditions underlie the continued high rates of CVD morbidity and mortality. These common CVDs include coronary heart disease, heart failure and sudden cardiac death; the well established common risk factors for CVD include hypertension, dyslipidemia, and adult onset diabetes mellitus.[9, 10] While it appears that few cases of these conditions in the general population are explained by rare Mendelian causes, there is remarkably little reliable information on the actual proportion of risk for common CVDs (ie, the attributable risk) that is conferred by rare (eg, MAF <<0.1%) versus uncommon (eg, MAF <0.1-1%) mutations versus common variants (eg, MAF ≥1%).

Recently, investigators conducted sequencing studies in general populations to have begin to estimate the contribution to common CVD risk of variants in genes known to underlie Mendelian CVD or other candidate genes. For example, a recent sequencing study of genetic variation in exons of sarcomere protein genes and storage cardiomyopathy-causing genes was conducted for determinants of increased left ventricular wall thickness in ~1800 persons

from the community-based Framingham Heart Study. There were 7 mutations found in 5 sarcomere protein genes and 1 mutation in the alpha-galactosidase A (*GLA*) gene; at least one of these mutations was detected in 1% of all persons in the community and in 18% of person with increased left ventricular wall thickness.[11] In a separate community-based sequencing study, at least 1 out of 64 subjects was found to carry a functional mutation in one of three genes—*NCCT*, *NKCC2* and *ROMK*—previously known to cause rare recessive diseases featuring large reductions in blood pressure.[12] And in a sequencing study of the adipokine *ANGPTL4* gene in over 3500 subjects from the Dallas Heart Study, nonysnonymous (amino-acid altering) variants were found in 13 persons with low triglycerides versus 2 with high triglycerides, and more nonsynonymous variants were found in European Americans than in African Americans.[13] The role of human ANGPTL4 has been unknown, although in mice a genetic deletion of ANGPTL4 leads to lower plasma triglyceride levels, and overexpression of ANGPTL4 in the liver leads to hypertriglyceridemia.[13] A similar approach led to the discovery of variants of *PCSK9* occurring in ~3% of the population and associated with low LDL and low risk of incident coronary heart disease.[14, 15]

Reliable estimates regarding the overall contribution of rare and uncommon variants to the burden of common cardiovascular diseases awaits largescale sequencing studies of not only genes implicated in Mendelian forms of CVD but also all other known genes. Candidate gene sequencing studies are underway for CVD candidate genes supported by both the NHGRI Medical Resequencing Program[16] and the NHLBI Resequencing and Genotyping Program.[17] In a recent RFA, the NHLBI will support the development of technologies to enable the accurate, high throughput genomewide sequencing of all protein-coding gene exons and other important regulatory regions in large numbers of human subjects, with the ultimate goal of defining the association of such variants with cardiovascular, lung and blood diseases.

As we await results of sequencing studies, our estimates of the role of Mendelian versus more complex forms of inheritance in CVD derive primarily from available family history studies. One clue may be found in the results of a very large survey of 122,155 Utah families, which underscores the significant burden of familial risk to CHD and stroke but the relatively small contribution of rare conditions. A positive family history of coronary heart disease accounted for 14% of all families with CHD but nearly three quarters of persons with early onset CHD; a positive family history of stroke was noted in 11% of families and accounted for 86% of early onset strokes.[18] On the other hand, only 1% of families carry a very strongly positive family history although these families account for 17% of cases of early onset CHD and 19% of cases of early onset strokes.[18] These data generate the hypothesis that rare "private" mutations in families with a very strongly positive family history may account for a small but significant minority of the cases of early onset CVD.

Common genetic variation in CVD

There is substantial evidence that heritable factors underlie the variation in clinical CVD, subclinical CVD and its risk factors in human populations. However, while pedigree studies often provide strong evidence for a Mendelian mode of transmission for rare familial CVD, the impact of a familial predisposition and the mode of transmission is often not apparent from studies of individual pedigrees for most cases of common, "complex" CVD like myocardial infarction. Studies of large groups of families or observational studies of family history have provided evidence for a familial predisposition to myocardial infarction,[19, 20] atrial fibrillation,[21, 22] and congestive heart failure.[23] Heritability—the proportion of interindividual phenotypic variation that is attributed to genetic variation among individuals—can be estimated in groups of siblings or other related family members. There is

substantial evidence from multiple studies for moderate heritability (e.g., 30-50%) in the quantitative measurement of subclinical atherosclerosis, such as coronary artery calcification[24, 25] and carotid IMT.[26] Similarly, there is moderate heritability for traditional risk factor measures, such as systolic blood pressure, total cholesterol, and body mass index. Thus, the totality of evidence from studies of the familial predisposition for CVD and from heritability studies of the independent precursors of CVD suggests that genetic variation underlies a substantial proportion of CVD and its risk factors.

With the public availability of the complete human genome sequence came greater attention to the hypothesis that common genetic variants may underlie most common, complex diseases such as CVD.[26] Among the most common sources of genetic variation are single nucleotide polymorphisms (SNPs), occurring about every 1000 base pairs in the approximately 3 billion base pairs comprising the human genome sequence. A complete publicly available catalogue of SNPs has been made available and is being updated by the NIH-supported international HapMap Project (http://www.hapmap.org/). While there are likely >10,000,000 SNPs in the human genome, SNPs that reside nearby each other are highly correlated due to a lack of recombination (also known as 'linkage disequilibrium') within chromosomal segments. From detailed studies of linkage disequilibrium patterns, it appears that genotype scans of 300,000 to 500,000 SNPs spaced across the entire genome provide sufficient statistical power to define the majority of common genetic variation.[27] Greater numbers of SNPs are needed to define variation for populations of older ancestry, such as those of African ancestry, than for more recent populations, such those of European ancestry.[22] Technology development and competition has allowed the genotyping of thousands of SNPs using high throughput, multiplex SNP assays on genotyping chips.

Genomewide association studies

A genome-wide association study (GWAS) is a study of genetic variation across the entire human genome designed to identify genetic associations with observable clinical traits or with the presence or absence of a disease. A first demonstration that a GWAS could be used to identify genes predisposing to common disease was reported in 2005, when a GWAS using a first generation chip of 100,000 SNPs was used in a case-control study of fewer than 150 subjects to identify variation in the complement factor H (*CFH*) predisposing to age-related macular degeneration.[28] The pooled estimate of multiple subsequent replication studies demonstrated unequivocally a consistent association with AMD of *CFH* variant. Since 2005, technology development and price reductions have led to several generations of DNA chips with higher densities, ranging from 300,000 to over 1,000,000 SNPs. Between 2006 and 2007, about ten published GWAS reports had appeared. Since 2006 to the present, over 150 GWAS reports have been published in the peer reviewed literature, reporting associations with a wide range of disease phenotypes of >300 SNP variants.

While many GWAS's for CVD are in progress and have yet to be reported, there are already several remarkable discoveries, through GWAS reports, of genetic variants associated with CVD and its risk factors. CVD outcomes studied by GWAS have included peripheral arterial disease, MI, stroke and atrial fibrillation.[29] Other GWAS phenotypes have included traditional CVD risk factors (including BMI and obesity, type 2 diabetes mellitus and fasting glucose, hypertension and systolic BP, nicotine dependence, as well as levels of LDL, HDL and triglycerides), QT interval length on ECG, inflammatory and hemostatic biomarkers, and imaging measures of subclinical atherosclerosis (including coronary artery calcium, carotid IMT) and left ventricular hypertrophy.[29] A few selected GWAS reports highlight some surprising findings. In one example, QT interval length on ECG was among the first CVD phenotypes to be studied by GWAS, revealing a strong association with SNPs in the *CAPON* gene in a staged study involving two independent cohorts[30] and subsequently

independently replicated in another large cohort.[31] Unlike most genes known to be associated with QT interval length, *CAPON* does not encode an ion channel but rather is a regulator of nitric oxide synthase; its mechanism of prolonging the QT interval is unknown. In a second example, a recent GWAS report was undertaken to identify genetic determinants of alterations in LDL cholesterol, HDL cholesterol or triglycerides in three independent cohorts in which collaborators had recently identified new loci underlying type 2 diabetes mellitus.[32-34] Using an *in silico* meta-analysis of GWAS data from all three cohorts, common SNPs from 18 loci were associated with at least one of the lipid phenotypes. Of these loci, six were unexpected and had not been previously implicated in lipid physiology.[35] A risk score, using nine of previously identified LDL or HDL related SNPs, was tested for association with incident CVD; the risk score discriminated as well as but not better than clinical risk factors.[36] Finally, three independent GWAS studies recently reported strong evidence for an association of a region in chromosome 9p21 near the cyclin-dependent kinase inhibitor genes *CDKN2A* and *CDKN2B*[33-35]. However, the signal for association is in a region free of known protein-coding sequence, and, interestingly, one of the three new loci associated with type 2 diabetes mellitus is also in chr 9p21.[37-39] Several other loci are implicated but have somewhat lower strength of evidence for association with myocardial infarction.[40]

The number of GWAS reports is increasing in an exponential manner and several common themes are emerging. First, because of the risk of false positive results from multiple testing, SNP associations must meet rigorous statistical thresholds to be considered "significant on a genomewide level". Remarkably, many of the initial GWAS association findings have been replicated, in a convincing manner, using modern, stringent guidelines.[41] Second, the magnitude of most SNP associations in GWAS is modest (relative risks of 1.2-1.4), as is the proportion of variation explained by individual SNPs, emphasizing the need for large sample sizes to discover true positive associations. Third, while a growing number of GWAS associations have confirmed previously reported associations using other methods, many loci discovered in GWAS are neither located in or near protein-coding genes, nor are they near genes previously known to be associated with the disease of interest. Thus, it appears that some genetic associations may be attributed to genetic mechanisms other than alterations in the sequence of protein-coding genes. Of interest, in a recent analysis of human gene expression, it was estimated that the majority of bases in the human genome are found in gene transcripts, including non-protein coding transcripts.[42] Fourth, it should be remembered that genetic associations from GWAS do not constitute the "causal" variant but rather are a marker of a probable nearby variant. Further resequencing and/or fine mapping of associated regions are indicated to isolate the functional variant(s). Fifth, a number of loci identified in GWAS are associated with several distinct diseases, suggesting pleiotropy. Finally, to date there have been very few studies that convincingly describe the biological function or deleterious effect of SNPs and/or genes identified through GWAS.

NHLBI genome programs and new research models

The NHLBI supports a robust portfolio of GWAS programs that provide population diversity and great depth and breadth of clinically relevant phenotypes (Table 1). The NHLBI's position is that the greatest public benefit will be realized if data from GWAS are made available, under terms and conditions consistent with the informed consent provided by individual participants, in a timely manner to the largest possible number of investigators. Accordingly, the NHLBI has played a leadership role in the development of a trans-NIH data sharing policy for GWAS studies, effective January 25, 2008 (http://grants.nih.gov/grants/gwas/), and the NHLBI adheres strongly to this policy. Concurrent with development of the policy, the National Center for Biotechnology (NCBI) created a genotype-phenotype NIH database called dbGaP (database of genotype and

phenotype).[43] The SNP Health Association Resource (SHARe) Program is the first NHLBI genomic program that now has data available for biomedical researchers.[44] The SHARe scientific resource draws genotyping of the three generations of participations in the NHLBI Framingham Heart Study, including 550,000 SNPs in over 9,400 consenting participants. The phenotype component includes >16,800 clinical variables, including risk factors, biomarkers, subclinical imaging measures, and clinical CVD as well as other chronic disease areas, such as bone density and dementia. Additional cohorts are planned for SHARe, including the SHARe Asthma Research Project (SHARP).

STAMPEED (SNP Typing for Association with Multiple Phenotypes from Existing Epidemiologic Data) consists of 13 grants awarded to support GWAS in a wide range of cardiovascular, lung and related diseases. Nearly 40,000 male and female participants are included with substantial representation from minority populations. In the Candidate Gene Association Resource (CARe), a GWAS is now underway in over 10,000 African American subjects from five NHLBI supported population cohorts to study a range of cardiovascular phenotypes. The ENDGAME (Enhancing Development of Genome-Wide Association Methods) program supports eleven teams of investigators to develop innovative strategies and tools for planning, initiating, conducting, and analyzing genome-wide association studies. All strategies and tools developed through these awards will be made available to the scientific community. In its totality, with ongoing or planned GWAS, the NHLBI portfolio will study well over 100,000 research participants from multiple, diverse and well-phenotyped disease cohorts and population cohorts.

Observations from GWAS studies to date and the robust portfolio of ongoing GWAS suggest a number of immediate next steps for genomewide association studies (Figure 2). First, the availability of data for sharing will increase the opportunities for a broad diversity of scientists to make new discoveries. Second, collaboration between multiple GWAS will be necessary to convincingly replicate genetic associations and exploit the power of individual GWAS. The aforementioned consortia of researchers studying the genomics of type 2 diabetes mellitus and the Wellcome Trust Case Control Consortium *(*WTCCC) represent two successful models of genomics collaboration focused on one or a few diseases. In one newly formed consortium, the Cohorts for Heart and Aging Research in Genome Epidemiology (CHARGE) Consortium, GWAS results are being systematically pooled in a series of *in silico* meta-analyses from five prospective, population-based cohorts to study GWAS associations across over twenty different phenotype areas in CVD, aging and multiple other chronic diseases.[45] Third, the functional significance of GWAS associations must be defined, affording great opportunity for collaboration between bench scientists, statisticians, epidemiologists, and clinical investigators. Moreover, functional genomics requires animal and human tissue models and systems biology approaches that utilize proteomics, metabolomics, and gene expression profiling. Finally, as sequencing cost decreases and accuracy improves, deep sequencing of ever larger proportions of the human genome with follow providing the opportunity to define the totality of common, less common and rare variation underlying CVD. The scientific opportunities have never been greater to unravel the mystery of genetic predisposition to CVD.

Conclusion

The NHLBI supports a large portfolio of population based genetic and genomic programs in diverse US populations. We are strongly committed to making participant level data and aggregate research results available to a broad community of investigators, while protecting the privacy and confidentiality of study participants. An abundance of knowledge regarding the Mendelian forms of CVD exists; yet, the current explosion of genomewide association studies is leading to a remarkable pace of discovery regarding highly prevalent

cardiovascular risk factors and common, clinically apparent CVD. These discoveries represent only "the end of the beginning". The next phase of discovery will require deep sequencing of the human genome, functional research in animal and human models, and application to clinical medicine.

References

1. The Fact Book. National Institutes of Health, National Heart, Lung, and Blood Institute; 2007.

2. The maladies of affluence. The Economist. August 7.2007

3. NHLBI Strategic Plan; 2007. http://apps.nhlbi.nih.gov/strategicplan/StrategicPlan_Appendix.pdf

4. Maron BJ, Moller JH, Seidman CE, Vincent M, Dietz HC, Moss AJ, Towbin JA, Sondheimer HM, Pyeritz RE, McGee G, Epstein AE. Impact of laboratory molecular diagnosis on contemporary diagnostic criteria for genetically transmitted cardiovascular diseases: Hypertrophic cardiomyopathy, Long-QT Syndrome, and Marfan Syndrome. Circulation. 1998; 98:1460–1471.

5. Eriksson M, Brown WT, Gordon LB, Glynn MW, Singer J, Scott L, Erdos MR, Robbins CM, Moses TY, Berglund P, Dutra A, Pak E, Durkin S, Csoka AB, Boehnke M, Glover TW, Collins FS. Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome. Nature. 2003:293–298. [PubMed: 12714972]

6. Merideth MA, Gordon LB, Clauss S, Sachdev V, Smith ACM, Perry MB, Brewer CC, Zalewski C, Kim HJ, Solomon B, Brooks BP, Gerber LH, Turner ML, Domingo DL, Hart TC, Graf J, Reynolds JC, Gropan A, Yanovski JA, Gerhard-Herman M, Collins FS, Nabel EG, Cannon RO, Gahl WA, Introne WJ. Phenotype and course of Hutchinson-Gilford Progeria Syndrome. N Engl J Med. 2008; 358:592–604. [PubMed: 18256394]

7. Varga R, Eriksson M, Erdos MR, Oliver M, Harten I, Kolodgie F, Capell S, Gordon LB, Virmani R, Wight TN, Nabel EG, Collins FS. Progressive vascular smooth muscle cell defects in a mouse model of Hutchinson-Gilford progeria syndrome. Proc Natl Acad Sci USA. 2006; 103:3250–3255. [PubMed: 16492728]

8. Capell BC, Olive M, Erdos MR, Ao K, Faddah DA, Tavarz, Conneely KN, Qu X, San H, Ganesh SK, Chen X, Avallone H, Kolodgie FD, Virmani R, Nabel EG, Collins FS. A farnesyltransferase inhibitor prevents both the onset and late progression of cardiovascular disease in a progeria mouse model. Proc Natl Acad Sci USA. 2008 (In press).

9. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, Kannel WB. General cardiovascular risk profile for use in primary care. Circulation. 2008; 117:743–753. [PubMed: 18212285]

10. Wilson PWF, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation. 1998; 97:1837–1847. [PubMed: 9603539]

11. Morita H, Larson MG, Barr SC, Vasan RS, O'Donnell CJ, Hirschhorn JN, Levy D, Corey D, Seidman CE, Seidman JG, Benjamin EJ. Single-gene mutations and increased left ventricular wall thickness in the community. Circulation. 2006; 113:2697–2705. [PubMed: 16754800]

12. 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. 40:592–599. 200. [PubMed: 18391953]

13. Romeo S, Pennacchio LA, Fu Y, Boerwinkle E, Tybjaerg-Hansen, Hobbs HH, Cohen JC. Population-based resequencing of ANGPTL4 uncovers variations that reduce triglycerides and increase HDL. Nat Genet. 2007; 39:513–516. [PubMed: 17322881]

14. Cohen JD, Boerwinkle E, Mosley TH, Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. N Engl J Med. 2006; 354:1264–1272. [PubMed: 16554528]

15. Cohen J, Pertsemlidis A, Kotowski IK, Graham R, Garcia CK, Hobbs HH. Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. Nat Genet. 2005; 37:161–165. [PubMed: 15654334]

16. NHGRI Medical Resequencing Program; 2008. www.genome.gov/15014882

17. NHLBI Resequencing and Genotyping Service; 2008. www.rsng.nhlbi.gov/scripts/index.cfm

18. William RR, Hunt SC, Heiss G, Province MA, Bensen JT, Higgins M, Chamberlain RM, Ware J, Hopkins PN. Usefulness of cardiovascular family history data for population-based preventive medicine and medical research (the Health Family Tree Study and the NHLBI Family Heart Study). Am J Cardiol. 2001; 87:129–135. [PubMed: 11152826]

19. Lloyd-Jones DM, Nam BH, D;Agostinio RB Sr, Levy D, Murabito JM, Wang TJ, Wilson PW, O'Donnell CJ. Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring. JAMA. 2004; 291:2204–2211. [PubMed: 15138242]

20. Marenberg ME, Risch N, Berkman LF, de Faire U. Genetic susceptibility to death from coronary heart disease in a study of twins. N Eng J Med. 1994; 330:1041–1046.

21. Arnar DO, Thorvaldsson S, Manolio TA, Thorgeirsson G, Kristjansson K, Hakonarson H, Stefansson K. Familial aggregation of atrial fibrillation in Iceland. Eur Heart J. 2006; 27:708–712. [PubMed: 16428254]

22. Fox CS, Parise H, D'Agostino RB Sr, Lloyd-Jones DM, Vasan RS, Wang TJ, Levy D, Wolf PA, Benjamin EJ. Parental atrial fibrillation as a risk factor for atrial fibrillation in offspring. JAMA. 2004; 292:1174–1175.

23. Lee DS, Pencina MJ, Benjamin EJ, Wang J, Levy D, O'Donnell CJ, Nam B-H, Larson MG, D'Agostino RB, Vasan RS. Association of parental heart failure with risk of heart failure in offspring. N Eng J Med. 2006; 355:138–147.

24. Peyser PA, Bielak LF, Chu JS, Turner ST, Ellsworth DL, Boerwinkle E, Sheedy PF. Heritability of coronary artery calcium quantity measured by electron beam computed tomography in asymptomatic adults. Circulation. 2002; 106:304. [PubMed: 12119244]

25. O'Donnell CJ, Chzaro I, Wilson PWF, Fox C, Hannan MT, Kiel DP, Cupples LA. Evidence for heritability of abdominal aortic calcific deposits in the Framingham Heart Study. Circulation. 2002; 106:337. [PubMed: 12119250]

26. Manolio TA, Boerwinkle E, O'Donnell CJ, Wilson AF. Genetics of ultrasonographic carotid atherosclerosis. Arterioscler Thromb Vasc Biol. 2004; 24:1567–1577. [PubMed: 15256397]

27. Pe'er I, de Bakker PIW, Maller J, Yelensky R, Altshuler D, Daly MJ. Evaluating and improving power in whole-genome association studies using fixed marker sets. Nat Genet. 2006; 38:663–667. [PubMed: 16715096]

28. Kelin RJ, eiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, Henning AK, SanGiovanni JP, Mane SM, Mayne ST, Bracken MB, Ferris FL, Ott J, Barnstable C, Hoh J. Complement factor H polymorphism in age-related macular degeneration. Science. 2005; 308:362–364. [PubMed: 15769856]

29. Manolio TA, Brooks LD, Collins FS. A HapMap harvest of insights into the genetics of common disease. J Clin Invest. 2008; 118:1590–1595. [PubMed: 18451988]

30. Arking DE, Pfeufer A, Post W, Kao WHL, Newton-Cheh C, Ikeda M, West K, Kashuk C, Akyol M, Perz S, Jalilzadeh S, Illig T, Gieger C, Guo G-Y, Larson MG, Wichmann HE, Marban E, O'Donnell CJ, Hirschhorn JN, Kaab S, Spooner PM, Meitinger T, Chakravarti A. A common genetic variant in the NOS1 regulator NOS1AP modulates cardiac repolarization. Nat Genet. 2006; 38:644–651. [PubMed: 16648850]

31. Aarnoudse A-JLHJ, Newton-Cheh C, de Bakker PIW, Straus SMJM, Kors JA, Hofman A, Uitterlinden AG, Witteman JCM, Stricker BHC. Common NOS1AP variants are associated with a prolonged QTc interval in the Rotterdam Study. Circulation. 2007; 116:10–16. [PubMed: 17576865]

32. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliot KS, Lango H, Timpson NJ, Perry JR, Rayner NW, Freathy RM, Barrett JC, Shields B, Morris AP, Ellard S, Groves CJ, Harries LW, Marchini JL, Owen KR, Knight B, Cardon LR, Walker M, Hitman GA, Morris AD, Doney AS, Wellcome Trust Case Control Consortium (WTCCC). McCarthy MI, Hattersley AT. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science. 2007; 317:1035–1036. [PubMed: 17717169]

33. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, Erdos MR, Stringham HM, Chines PS, Jackson AU, Prokunina-Olsson L, Ding CJ, Swift AJ, Narisu N, Hu T, Pruim R, Xiao R, Li XY, Conneely KN, Riebow NL, Sprau AG, tong M, White PP, Hetrick KN, Barnhart MW, Bark CW, Goldstein JL, Watkins L, Xiang F, Saramies J, Buchanan TA, Watanabe RM, Valle tT,

Kinnunen L, abecasis GR, Pugh EW, Doheny KF, Bergman RN, Tuomilehto J, Collins FS, Boehnke M. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibililty variants. Science. 2007; 316:1341–1345. [PubMed: 17463248]

34. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University; 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, Altschuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Borstrom K Bengtsson, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nillson P, Orho-Melander M, Rastam L, Speliotes EK, Taskinen MR, Tuomi T, Guiducci C, Bergllund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjogren 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. [PubMed: 17463246]

35. Kathiresan S, Melander O, Guiducci C, Surti A, Brutt 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. [PubMed: 18193044]

36. Kathiresan S, Melander O, Anevski D, Guiducci C, Burtt NP, Roos C, Hirschhorn JN, Berglund G, Hedblad B, Groop L, Altshuler DM, Newton-Cheh C, Orho-Melander M. Polymorphisms associated with cholesterol and risk of cardiovascular events. N Engl J Med. 2008; 358:1240–1249. [PubMed: 18354102]

37. The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common disease controls. Nature. 2007; 447:661–678. [PubMed: 17554300]

38. McPherson R, Pertsemlidis A, Kavaslar N, Stweart A, Roberts R, Cox DR, Hinds DA, Pennacchio LA, Tybjaerg-Hansen A, Folsom AR, Boerwinkle E, Hobbs HH, Cohen JC. A common allele on chromosome 9 associated with coronary heart disease. Scienc. 2007; 316:1488–1491.

39. Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, Jonasdottir A, Jonasdottir A, Sirgurdsson A, Baker A, Palsson A, Masson G, Gudbjartsson DF, Magnusson KP, Andersen K, Levley AI, Backman VM, Matthiasdotir S, Jonsdottir T, Palsson S, Einarsdotir H, Gunnarsdottir S, Gylfason A, Vaccarino V, Hoopoer WC, Reilly MP, Granger CB, Austin H, Rader DJ, Shah SH, Quyyumi AA, Gulcher JR, Thorgeirsson G, Thorsteinsdottir U, Kong A, Stefansson K. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science. 2007; 316:1491–1493. [PubMed: 17478679]

40. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, Dixon RJ, Meitinger T, Braund P, Wichmann HE, Barret JH, Konig IR, Stevens SE, Szymczak S, Tregouet DA, Iles MM, Pahlke F, Pollard H, Lieb W, Cambien F, Fischer M, Ouwehand W, Blankenberg S, Galmforth AJ, Baessler A, Ball SG, Strom TM, Braenne I, Gieger C, Deloukas P, Tobin MD, Ziegler A, Thompson JR, Schunkert H, WTCCC and the Cardiogenics Consortium. Genomewide association analysis of coronary artery disease. N Engl J Med. 2007; 357:43–453.

41. NCI-NHGRI Working Group on Replication in Association Studies. Chanock SJ, Manolio T, boehnke M, Boerwinkle E, Hunter DJ, Thomas G, Hirschhorn JN, Agecasis G, Altshuler D, Bailey-Wilson JE, Brooks LD, Cardon LR, Daly M, Donnelly P, Fraumeni JF Jr, Freimer NB, Gerhard DS, Gunter C, Guttmacher AE, Guyer MS, Harris EL, Hoh J, Hoover r, Kong CA, Merikangas KR, Morton CC, Plamer LJ, Phimister EG, Rice JP, Robert J, Rotimi C, Tucker MA, vogan KJ, Wacholder S, Wijsman EM, Winn DM, Collins FS. Replicating genotype-phenotype associations. Nature. 2007; 447:655–660. [PubMed: 17554299]

42. ENCODE Project Consortium; Birney E, Stamatoyannopoulos JA, Dutta A, Guigó R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, Thurman RE, Kuehn MS, Taylor CM, Neph S, Koch CM, Asthana S, Malhotra A, Adzhubei I, Greenbaum JA, Andrews RM, Flicek P, Boyle PJ, Cao H, Carter NP, Clelland GK, Davis S, Day N, Dhami P, Dillon SC, Dorschner MO, Fiegler H, Giresi PG, Goldy J, Hawrylycz M, Haydock A, Humbert R, James KD, Johnson BE, Johnson EM, Frum TT, Rosenzweig ER, Karnani N, Lee K, Lefebvre GC, Navas PA, Neri F,

Parker SC, Sabo PJ, Sandstrom R, Shafer A, Vetrie D, Weaver M, Wilcox S, Yu M, Collins FS, Dekker J, Lieb JD, Tullius TD, Crawford GE, Sunyaev S, Noble WS, Dunham I, Denoeud F, Reymond A, Kapranov P, Rozowsky J, Zheng D, Castelo R, Frankish A, Harrow J, Ghosh S, Sandelin A, Hofacker IL, Baertsch R, Keefe D, Dike S, Cheng J, Hirsch HA, Sekinger EA, Lagarde J, Abril JF, Shahab A, Flamm C, Fried C, Hackermüller J, Hertel J, Lindemeyer M, Missal K, Tanzer A, Washietl S, Korbel J, Emanuelsson O, Pedersen JS, Holroyd N, Taylor R, Swarbreck D, Matthews N, Dickson MC, Thomas DJ, Weirauch MT, Gilbert J, Drenkow J, Bell I, Zhao X, Srinivasan KG, Sung WK, Ooi HS, Chiu KP, Foissac S, Alioto T, Brent M, Pachter L, Tress ML, Valencia A, Choo SW, Choo CY, Ucla C, Manzano C, Wyss C, Cheung E, Clark TG, Brown JB, Ganesh M, Patel S, Tammana H, Chrast J, Henrichsen CN, Kai C, Kawai J, Nagalakshmi U, Wu J, Lian Z, Lian J, Newburger P, Zhang X, Bickel P, Mattick JS, Carninci P, Hayashizaki Y, Weissman S, Hubbard T, Myers RM, Rogers J, Stadler PF, Lowe TM, Wei CL, Ruan Y, Struhl K, Gerstein M, Antonarakis SE, Fu Y, Green ED, Karaöz U, Siepel A, Taylor J, Liefer LA, Wetterstrand KA, Good PJ, Feingold EA, Guyer MS, Cooper GM, Asimenos G, Dewey CN, Hou M, Nikolaev S, Montoya-Burgos JI, Löytynoja A, Whelan S, Pardi F, Massingham T, Huang H, Zhang NR, Holmes I, Mullikin JC, Ureta-Vidal A, Paten B, Seringhaus M, Church D, Rosenbloom K, Kent WJ, Stone EA, NISC Comparative Sequencing Program; Baylor College of Medicine Human Genome Sequencing Center; Washington University Genome Sequencing Center; Broad Institute; Children's Hospital Oakland Research Institute. Batzoglou S, Goldman N, Hardison RC, Haussler D, Miller W, Sidow A, Trinklein ND, Zhang ZD, Barrera L, Stuart R, King DC, Ameur A, Enroth S, Bieda MC, Kim J, Bhinge AA, Jiang N, Liu J, Yao F, Vega VB, Lee CW, Ng P, Shahab A, Yang A, Moqtaderi Z, Zhu Z, Xu X, Squazzo S, Oberley MJ, Inman D, Singer MA, Richmond TA, Munn KJ, Rada-Iglesias A, Wallerman O, Komorowski J, Fowler JC, Couttet P, Bruce AW, Dovey OM, Ellis PD, Langford CF, Nix DA, Euskirchen G, Hartman S, Urban AE, Kraus P, Van Calcar S, Heintzman N, Kim TH, Wang K, Qu C, Hon G, Luna R, Glass CK, Rosenfeld MG, Aldred SF, Cooper SJ, Halees A, Lin JM, Shulha HP, Zhang X, Xu M, Haidar JN, Yu Y, Ruan Y, Iyer VR, Green RD, Wadelius C, Farnham PJ, Ren B, Harte RA, Hinrichs AS, Trumbower H, Clawson H, Hillman-Jackson J, Zweig AS, Smith K, Thakkapallayil A, Barber G, Kuhn RM, Karolchik D, Armengol L, Bird CP, de Bakker PI, Kern AD, Lopez-Bigas N, Martin JD, Stranger BE, Woodroffe A, Davydov E, Dimas A, Eyras E, Hallgrímsdóttir IB, Huppert J, Zody MC, Abecasis GR, Estivill X, Bouffard GG, Guan X, Hansen NF, Idol JR, Maduro VV, Maskeri B, McDowell JC, Park M, Thomas PJ, Young AC, Blakesley RW, Muzny DM, Sodergren E, Wheeler DA, Worley KC, Jiang H, Weinstock GM, Gibbs RA, Graves T, Fulton R, Mardis ER, Wilson RK, Clamp M, Cuff J, Gnerre S, Jaffe DB, Chang JL, Lindblad-Toh K, Lander ES, Koriabine M, Nefedov M, Osoegawa K, Yoshinaga Y, Zhu B, de Jong PJ. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature. 2007; 447:799–816. [PubMed: 17571346]

43. Mailman MD, Feolo M, Jin Y, Kimura M, Tryka K, Bagoutdinov R, Hao L, Kiang A, Paschall J, Phan L, Poova N, Pretel S, Ziyabari L, Lee M, Shao Y, Wang ZY, Sirotkin K, Ward M, Kholodov M, Zbicz K, Beck J, Kimelman M, Shevelev S, Preuss D, Yaschenko E, Graeff A, Ostell J, Sherry ST. The NCBI dbGaP database of enotypes and phenotypes. Nat Genet. 2007; 39:118–1186.

44. Russo G. NIH offers free access to wealth of disease data. Nature. 2006; 444:982. [PubMed: 17183282]

45. Cohorts for heart and aging research in genome epidemiology (CHARGE). Consortium members include the Netherland's Rotterdam Study; the NHLBI-supported Atherosclerosis Risk in Communities (ARIC) Study, Cardiovascular Health Study (CHS) and Iceland Age, Gene/Environment Susceptibility (AGES) Study. 2008.
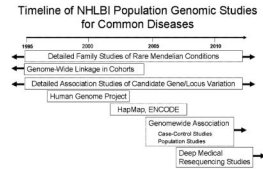
Figure 1.
Timeline of NHLBI genetic and genomic investigations for rare and common diseases, taking advantages of advances in the Human Genome Project and other technologies.
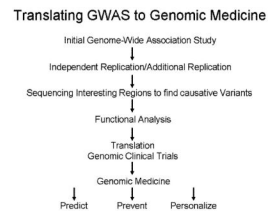
Figure 2.
NHLBI Strategic Plan for translating genomic discoveries to clinical medicine. The NHLBI goal is to capitalize on the rich scientific knowledge to predict risk for disease, prevent its onset, and personalize diagnostic and therapeutic care.

Table 1

NHLBI Programs in Genomic Research

Program	Objective	Subjects and Population(s)	Selected CVD, Lung, and Blood Phenotype(s)
Candidate Gene Association Resource (CARe)	Genomewide association study in over 10,000 African American subjects. Candidate gene association Study of ~50,000 SNPs in ~ 2,000 genes implicated in CVD.	GWAS in >10,000 African American subjects from five cohorts. Candidate gene studies in >48,000 subjects from nine cohorts. Cohorts include: ARIC, CARDIA, CFS, CHS, CSSCD, FHS, JHS, MESA and SHHS	Aging, anthropometry, atrial fibrillation, blood biomarkers, blood pressure/hypertension, coronary heart disease/myocardial infarction, cardiac ultrasound measures/heart failure, diabetes, kidney disease, lipids, pulmonary function, sleep, stroke, subclinicla atherosclerosis, peripheral arterial disease.
SNP Typing for Association with Multiple Phenotypes from Existing Epidemiologic Data (STAMPEED)	13 GWAS projects to identify genetic variants related to heart, lung, and blood disorders and their risk factors by utilizing genome-wide association studies in existing human studies	GWAS studies in multiple existing human studies totaling over 38,000 subjects.	Aging, asthma, coronary heart disease/myocardial infarction, stroke, platelet function, hematopoietic cell transplant outcomes, subclinical atherosclerosis, sickle cell anemia outcomes, aging
SNP Health Association Resource (SHARe)	GWAS project in large NHLBI cohorts studies to identify genetic variants related to heart, lung and blood disorders and their risk factors.	Framingham Heart Study SHARe: GWAS in over 9,400 subjects from three generations of the Framingham Heart Study SHARe Asthma Research Project (SHARP): GWAS in ~5,000 subjects from asthma cohorts.	Aging, anthropometry, atrial fibrillation, blood biomarkers, blood pressure/hypertension, coronary heart disease/myocardial infarction, cardiac ultrasound measures/heart failure, diabetes, kidney disease, lipids, pulmonary function, sleep, stroke, subclinical atherosclerosis, peripheral arterial disease. Asthma and asthma-treatment response.
Enhancing Development of Genome-wide Association Methods (ENDGAME)	Consortium of 11 research groups to develop and test innovative, informative, and cost-effective study designs and analytical strategies for performing GWAS on complex diseases.		
Women's Health Genome Study	GWAS to identify genetic variants related to cardiovascular disease, cancer and selected risk factors.	Up to 28,000 women participants in the Women's Health Study, a study cohort which has been followed for over a decade as part of a randomized controlled study for prevention of cardiovascular disease and cancer.	Blood biomarkers, hypertension, coronary heart disease/myocardial infarction, diabetes, lipids, stroke.

Abbreviations: ARIC: Atherosclerosis Risk in Communities, CARDIA: Coronary Artery Risk Development in Young Adults, CFS: Cleveland Family Study, CHS: Cardiovascular Health Study, CSSCD: Cooperative Study of Sickle Cell Disease, FHS: Framingham Heart Study, JHS: Jackson Heart Study, MESA: Multi-Ethnic Study of Atherosclerosis, SHHS: Sleep Heart Health Study.