A.J. Marian, M.D.

THE PERSONAL GENOME AND THE PRACTICE OF CARDIOVASCULAR MEDICINE

A.J. Marian, M.D.

The Brown Foundation Institute of Molecular Medicine, The University of Texas Health Science Center, Texas Heart Institute at St. Luke's Episcopal Hospital, Methodist DeBakey Heart & Vascular Center, Houston, Texas

Abstract

Recent advances in the DNA sequencing techniques have made it possible to sequence the entire protein coding regions of the genome and even the entire genome at a reasonable cost. Genetic discoveries, facilitated by these advances, have illustrated the enormous genetic diversity of the mankind. Accordingly, the genome of each individual has about 4 million sequences that are different from the general population, including a significant number of unique sequence variants. About 3 million of these variations affect 1 of the 3 billion bases in the genome, while the rest affect from 2 to several million base pairs of DNA. About 10,000 of these variants in each genome changes the amino acid sequence and hence, the protein structure. The biological and clinical significance of these DNA sequence variants follow a gradient that ranges from negligible to large. A small number of variants impart large effect sizes but those with negligible and small effect sizes are quite abundant in the genome. The most important contributions of these variants are likely to be in providing insights into the molecular mechanisms responsible for diseases, which is essential for the ultimate cure of the human disease. These genetic variants also influence susceptibility to diseases, responses to treatment and clinical outcome. Variants with large effect sizes have the potential to serve as diagnostic markers, prognosticators as well as individualization of therapy. The daunting challenge in the upcoming years is to identify the variants that have significant clinical impacts from those that impose no discernible effects. The landscape of the practice of medicine is expected to change with the incorporation of the information content of the DNA sequence variants into the routine practice of medicine. The medical community needs to be prepared to best utilize the information that will be available from "personal genomes" of the patients.

Introduction

June 26, 2010 is the tenth anniversary of the completion of the draft sequence of the human genome.[1, 2] The feat was accomplished using the ingenious DNA sequencing method that was developed by Dr. Frederick Sanger and Dr. Walter Gilbert, for which they received the 1980 Nobel Prize in Chemistry.[3] Sequencing of the human genome through "The Human Genome Project" took approximately 11 years and cost approximately $3 billion. It provided the first glimpses into the sequence of a haploid genome (1 copy of the chromosomes). Automation of DNA sequencing by the Sanger method dropped the cost of DNA sequencing from the initial price of $30 per base pair in 1990 to about $1 per 100 base pairs or $10,000 per 1 million base pairs. Despite the precipitous drop in cost, DNA sequencing remained relatively expensive, time consuming, and exclusive to research laboratories focused on sequencing specific genes of interest. Hence, the clinical applications of the genetic data had to await the development of faster and cheaper DNA sequencing technologies.

In 2005, the first "next-generation" DNA sequencing technique was introduced with the advent of array-based 454 pyrosequencing technology.[4] The machine could sequence hundreds of millions of base pairs in a single run in a few days. This was soon followed by the development of the sequencing by synthesis and

sequencing by ligation technologies that produce several Giga base pairs of DNA sequence (in which each base pair is sequenced multiple times) at a very low cost. These technologies, which are often referred to as massively parallel sequencing or deep resequencing, afforded the opportunity to sequence the entire human genome in weeks and the targeted genomic regions of interest in days. Accordingly, the estimated cost of sequencing 1 million base pairs of DNA has dropped to about $100. Today, the technology affords the ability to sequence the human genome in a relatively short period of time for less than $5,000.[5]

The DNA sequencing technology continues to evolve. Some companies now claim that they can sequence the entire human genome in a single day for less than $6,000. Likewise, single DNA molecule sequencing technologies have the potential to provide robust accuracy. Data analysis, however, remains a major challenge that is expected to be resolved within the upcoming years. Robust bioinformatics is necessary to decipher the terabytes of digital data that are generated through next-generation DNA sequencers. It is reasonable to expect that within a few years the complete sequence of the genome will be a commodity and available for the extraction of medical information.

DNA Sequence Variations in the Genome

The presence of DNA sequence variants in the genome was known from the time when Sanger sequenced fragments of the human genome. However, the extent of sequence variations among the individuals was largely unknown. The Human Genome Project, which became available in 2001 and was completed in 2004, provided the sequence of the haploid genome and hence did not portray a complete profile of sequence variants in the genome.[1, 2, 6] In 2007, Dr. J. Craig Venter sequenced the first diploid genome and ushered in new insight into the genetic diversity of mankind.[7] The findings illustrated the complexity of the human genome in terms of the presence of extensive sequence variants. It was mesmerizing to learn that about 45% of Dr. Venter's genes were polymorphic and that his genome contained about 4 million sequence variants including more than 3 million single nucleotide polymorphisms (SNPs) and several hundred thousands structural variations. Some structural variations affected several million base pairs of DNA and multiple genes. Accordingly, structural variations involved three-fourths of the variants' nucleotides. Subsequent sequencing of additional "personal" genomes, including the genomes of Dr. James Watson, an African man, a Korean man, a Han Chinese man, as well as sequencing of several exomes further illustrated

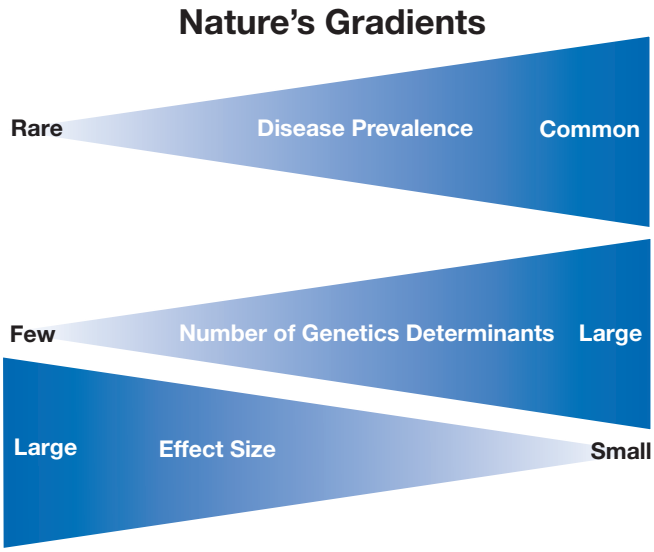

Figure 1. Genetic gradients. Single gene diseases are rare and caused by a single DNA sequence variant that imparts a large effect size, while a few others contribute to the phenotype as modifiers. The opposite end of the spectrum includes the complex and common disorders, which are the consequence of interactions of a very large number of DNA sequence variants, each imparting modest and often clinically indiscernible effect sizes.

the genetic diversity of humans.[7–11] Collectively, the data illustrates that each genome contains approximately 4 million sequence variants including more than 3 million SNPs, of which about 10,000 are nonsynonymous SNPs. By definition, non-synonymous SNPs change the amino acid sequence and hence have the potential to affect biological functions of the proteins and cause diseases. The genome of each individual also contains several hundred thousand structural variations, including about 200,000 small insertion/deletions and several hundred rearrangements and duplications, which can encompass several millions base pairs of DNA and several genes.[7, 12, 13] Overall, the genomes of different individuals differ in 1–3% of the sequence. The abundance of sequence variants in the genome accounts in part for the differences among the individuals in disease susceptibility, clinical outcome, and response to therapy.

Effect Sizes of the Genetic Variants

The human genome is a complex structure regulated at multiple levels, the most basic being through variations in its sequence. To understand the potential clinical implications of the roughly 4 million sequence variants in the genome, it is important to realize that the functional and biological significance of the variants varies and follows a gradient that ranges from negligible to profound (Figure 1).[14] In general, on a given

phenotype, a small number of the variants impart major effect sizes and offer diagnostic and therapeutic values, whereas, a large number of variants impart modest and often negligible effect sizes and hence have minimal, if any, clinical utility. On one end of the spectrum are the single gene disorders, such as familial cardiomyopathies or familial hypercholesterolemia, wherein a single sequence variant (i.e., mutation) is sufficient and necessary to cause the phenotype, while a handful of others contribute to phenotypic expression of the disease, as the modifier alleles. On the end of the spectrum are the complex and common disorders such as coronary artery disease, wherein typically a very large number of DNA sequence variants contribute to the phenotype, each imparting a negligible to modest effect. In addition to the gradient of effect sizes, there is also a gradient in the frequencies of alleles or sequence variants in the population. The most common DNA sequence variants in the population typically impart modest or indiscernible effect sizes. In contrast, rare and uncommon variants typically exert major phenotypic effects. Consequently, the most apparent clinical implications of genetics are in uncommon and single gene disorders and primarily for rare and uncommon sequence variants. In contrast, for the complex phenotype, only a handful of DNA sequence variants would be expected to impart modest clinical significant effects.

Clinical Implications of Genetic Variants

Sequencing of every individual's genome soon is expected to become a readily available and relatively cheap commodity. Hence, physicians and patients alike are likely to be exposed to a flood of genetic information, with the expectations that the information will lead to the prevention of human disease and better medical care for the patients. However, various clinical hurdles need to be overcome, and robust phenotyping needs to be established before one could fully harness the wealth of genetic information.

DNA sequence variants and preclinical diagnosis. The phenotype, by definition, is recognized only after its expression. The expressed phenotype, however, is often not detected or recognized until the individual becomes symptomatic. In terms of coronary atherosclerosis, this is often the case as many patients with significant coronary or carotid atherosclerosis are asymptomatic and hence are not diagnosed until they develop angina, myocardial infarction, or stroke. In the case of familial diseases, such as familial dilated cardiomyopathy, approximately 20% of family members who have dilated cardiomyopathy are asymptomatic and not aware of the condition. The percentage is even higher for hypertrophic cardiomyopathy, which is a major cause of sudden cardiac death in the young and competitive athletes.[15] Tragically, sudden cardiac death is often the first manifestation of hypertrophic cardiomyopathy in this group.[15]

Because the DNA sequence variants or the mutations are present since the formation of single cell embryo, genetic tools could be used for preclinical diagnosis of those who carry the mutation in their genome and are at the risk of disease. For certain high-risk diseases, genetic-based screening could lead to selection of the healthy embryos prior to in-vitro fertilization and implantation. Likewise, genetic-based diagnosis could provide a reliable method to distinguish family members who have inherited or carry the genetic mutation and are at risk of the disease from those who do not, and therefore are not at risk. In the former situation, early genetic-based diagnosis could lead to intervention to prevent or attenuate expression of the disease, wherein genetic-based exclusion of family members who do not carry the mutation in a familial disease offers the best possible medical news that one can present to a family member.

Genetic-based accurate diagnosis. Clinical phenotyping, which forms the foundation of the practice of medicine today, is neither completely sensitive nor specific. Clinical phenotyping often is unable to distinguish the phenocopy conditions, as their presentations often resemble the disease that they mimic. For example, more than a dozen diseases clinically mimic cardiomyopathies since their primary presentation is either cardiac hypertrophy or dilatation. Based on the prevalence of phenocopy conditions, it is estimated that approximately 10 to 15% of the clinically diagnosed hypertrophic cardiomyopathy are phenocopy conditions, such as glycogen storage diseases, Fabry disease, and Noonan syndrome, among others.[16] The prevalence of phenocopy conditions for dilated cardiomyopathy is probably even higher. Genetic-based diagnosis could lead to distinction of true cardiomyopathies from the phenocopy conditions. The distinction has important clinical implications as the management of patients with cardiomyopathies and phenocopy conditions differs significantly.

Genetic data is also expected to lead to identification of those who have the susceptibility alleles for complex conditions as well as those who are at risk of developing drug toxicity. Likewise, the clinical diagnosis of hypertrophic cardiomyopathy is often difficult in patients with essential hypertension, as per convention, the presence of hypertension excludes the diagnosis of hypertrophic cardiomyopathy. However, often the 2

Gene	Symbol	Drug	Biological Effect of Minor Allele	Clinical Effects of Minor Allele
P450 Isozyme 2C9	*CYP2C9*	Coumadin	Principle enzyme Coumadin metabolism	Inadequate or excess anticogulation
Vitamin K epoxide reductase complex, subunit 1	*VKORC1*		Reduces inactive vitamin K to the active form	
P450 isozyme 2C19	*CYP2C19*	Clopidogrel	Inadequate platelet inhibition due to slow conversion of pro-active to active drug	Increased risk of thrombotic events
Solute carrier organic anion transporter family, member 1B1	*SLCO1B1*	Statins	Reduces liver uptake of statins	Skeletal myopathy
Apolioprotein E	*APOE*	Statins	Poor uptake LDL-C	Response of LDL-C to treatment with statins
Adrenergic receptor β1	*ADRB1*	β Blockers	Ineffective β blockade	Response to treatment with β blockers
Angiotension converting enzyme	*ACE1*	ACE Inhibitors	Higher plasma and tissue ACE Levels	Response of treatment with ACE inhibitors

Table 1. Examples of DNA sequence variations influencing response to cardiovascular drug therapy.

conditions exist concomitantly. Concomitant presence of coronary artery disease and primary dilated cardiomyopathy is also difficult to discern. Moreover, the clinical diagnosis of hypertrophic cardiomyopathy is often difficult to distinguish from physiological hypertrophy observed in competitive athletes. The clinical significance of this distinction is noteworthy, as hypertrophic cardiomyopathy is the most common cause of SCD in competitive athletes.[17] A genetic-based approach could lead to identification of true hypertrophic cardiomyopathy in those with hypertension and in competitive athletes or concomitant dilated cardiomyopathy and coronary artery disease. Thus, a genetic-based diagnosis is expected to provide superior resolution for an accurate identification of individuals with specific phenotype of interest. One could anticipate that genetic-based diagnosis will supplant phenotype-based diagnosis and hence advance the practice of medicine significantly.

DNA sequence variants and risk stratification. The prognostic utility of the DNA sequence variants depends on the effect size of the specific variant and the remoteness of the phenotype from the genome. The larger the effect size, the greater the potential prognostic impact. Likewise, the impact of the DNA sequence variants are much larger on the proximal phenotypes, such as mRNA or protein levels, than on the distant phenotypes, such as the clinical events, wherein a large number of competing factors dilute the impact of the sequence variants. Accordingly, in single gene disorders such as familial cardiomyopathies and long QT interval, mutations offer some insight into the prognosis even though alone they are not sufficient to reliably predict the outcome. In contrast, in complex disorders such as coronary atherosclerosis, the prognostic impact of DNA sequence variants is very small and, for the vast majority if not all variants, negligible. Nonetheless, a small fraction of sequence variants with large effect sizes could provide valuable prognostic information under certain clinical circumstances.

The limited utility of DNA sequence variants in prognostication and risk stratification largely reflects the complexity of the clinical phenotypes that typically result from intertwined, non-linear, and often random and dynamic interactions among the various constituents of the phenotypes, whether genetic or non-genetic. Moreover, the outcome of such complex interactions is often context dependent. The genome contributes to the phenotype through copy number variants, intronic sequence variants, microRNAs, alternatively spliced mRNA species, and the epigenetics regulation of gene expression. Likewise, post-translational modifications of proteins such as phosphorylation and ubiquitination are important biological regulators of the clinical phenotype. The complexity of the factors that influence the clinical phenotype restricts the diagnostic utility of DNA sequence variants, particularly for the common cardiovascular disorders. The genome in this regard provides the platform for the various constituents to contribute to phenotypic expression of the clinical phenotype.

DNA sequence variants and individualized therapy. The inter-individual variation in response to drugs, both efficacy as well as toxicity, has been recognized since the dawning of medicine. Sir William Osler, the father of modern medicine, advocated indi-

vidualized therapy emphasizing that no 2 patients are identical in their diseases and responses to treatment. Advances in molecular genetics have led to partial elucidation of DNA sequence variants that influence response to and toxicity from drugs (pharmacogenetics). Table 1 summarizes partial lists of pharmacogenetics of cardiovascular drugs.

A well-recognized and frequently encountered example in the practice of medicine is response to warfarin therapy, which could vary by 10-fold among individuals. Inter-individual variations in response to Coumadin have been attributed to SNPs in genes encoding cytochrome P450 isoform 2C9 (CYP2C9) and vitamin K epoxide reductase (VKORC1). Genetic variations in CYP2C9 and VKORC1 together are responsible for about 30% of the inter-individual variations in Coumadin dose. A genotype-based approach to Coumadin dosing could help achieve the desirable therapeutic range while avoiding the risks of potential excess or inadequate anticoagulation, including hospitalization rate.[18] The response to anti-platelet agent clopidogrel, an inactive pro-drug that is converted to an active metabolite in the liver by the P-450 enzymes CYP3A4, CYP3A5, and CYP2C19, is in part genetically determined. CYP2C19, which is a polymorphic enzyme, catalyzes metabolism of a large number of drugs including omeprazole, fluvastatin, and clopidogrel. Non-synonymous SNPs, which reduce or abolish the enzymatic function of CYP2C19 (loss-of-function alleles), metabolize pro-drug clopidogrel less efficiently and generate a lower amount of the active metabolite. Hence, they exhibit reduced platelet inhibition with clopidogrel and would be expected to experience a high rate of thrombotic events, including death, stent thrombosis, and recurrent ischemic events.[19, 20]

Genetic variations in the solute carrier organic anion transporter 1B1 gene *(SLCO1B1)* have been associated with statin-induced myopathy.[21, 22] Likewise, SNPs in β1-adrenergic receptors can influence the response of patients with systolic heart failure to treatment with β blockers.[23] Moreover, SNPs in genes involved in lipid metabolism such APOE, PCSK9, and HMGCR have been implicated in response to statins.[24] Finally, SNPs in genes responsible for congenital long QT syndrome are associated with drug-induced cardiac arrhythmias.[25] A very large number of genetic variants influence response to various pharmacological agents, both in terms of efficacy as well as toxicity. The potential clinical implications and financial burden of pharmacogenetics are quite remarkable. With the availability of the sequence data of every individual's genome and increasing knowledge about the functional significance of DNA sequence variants, pharmacogenetics is likely to become rapidly incorporated into the daily practice of medicine.

Conclusion

The era of "personal genomes" has arrived. It is now possible to sequence the entire genome of every individual at a reasonable cost. Whole genome sequencing is becoming a readily available commodity that needs to be incorporated into the daily practice of medicine. The challenge facing physicians, patients, and researchers alike is to decipher those DNA sequence variants that are pathogenic, which are a handful, from those that have no discernible effects, which are the vast majority of variants in the genome, while realizing that there is a gradient of effect sizes. There should be no doubt that DNA sequence variants contribute to the pathogenesis of the clinical phenotype and are an important determinant of expression and severity of the disease, clinical outcomes, and the response to therapy. However, the clinical utility of these variants is best served in the pre-clinical identification of those at risk prior to and independent of the clinical phenotype, thereby increasing the potential for disease prevention. The greatest impact of molecular genetics is in providing fundamental insight into the pathogenesis of human diseases and offering the opportunity to identify new diagnostic and prognostic markers and preventive or therapeutic targets. The clinical impact of sequence variants is less for risk stratification and prognostication, as the clinical phenotypes arise from complex, non-linear, and often stochastic interactions among multiple constituents of the phenotype. Equally important is the clinical utility of sequence variants in identifying the responders from the poor responders and identifying those who are at risk of developing drug toxicity, hence fulfilling the promise of *Primum non nocere*.

The practice of medicine is changing, and a vital component of its success in the 21st century is the incorporation of genetic information. It has become essential that the medical community arm itself with the appropriate expertise and resources to handle the upcoming flood of genetic information.

Acknowlegements

Supported by grants from the NHLBI, Clinical Scientist Award in Translational Research from the Burroughs Wellcome Fund, and The TexGen Fund from the Greater Houston Community Foundation.

References

1. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferriera S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigó R, Campbell MJ, Sjolander KV, Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A, Zhu X. The sequence of the human genome. Science. 2001 Feb 16;291(5507):1304-51.

2. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blöcker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kaspryzk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ; International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature. 2001 Feb 15;409(6822):860-921.

3. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A. 1977 Dec;74(12):5463-7.

4. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM. Genome sequencing in microfabricated high-density picolitre reactors. Nature. 2005 Sep 15;437(7057):376-80.

5. Drmanac R, Sparks AB, Callow MJ, Halpern AL, Burns NL, Kermani BG, Carnevali P, Nazarenko I, Nilsen GB, Yeung G, Dahl F, Fernandez A, Staker B, Pant KP, Baccash J, Borcherding AP, Brownley A, Cedeno R, Chen L, Chernikoff D, Cheung A, Chirita R, Curson B, Ebert JC, Hacker CR, Hartlage R, Hauser B, Huang S, Jiang Y, Karpinchyk V, Koenig M, Kong C, Landers T, Le C, Liu J, McBride CE, Morenzoni M, Morey RE, Mutch K, Perazich H, Perry K, Peters BA, Peterson J, Pethiyagoda CL, Pothuraju K, Richter C, Rosenbaum AM, Roy S, Shafto J, Sharanhovich U, Shannon KW, Sheppy CG, Sun M, Thakuria JV, Tran A, Vu D, Zaranek AW, Wu X, Drmanac S, Oliphant AR, Banyai WC, Martin B, Ballinger DG, Church GM, Reid CA. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. Science. 2010 Jan 1;327(5961):78-81.

6. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. Nature. 2004 Oct 21;431(7011):931-45.

7. Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, Walenz BP, Axelrod N, Huang J, Kirkness EF, Denisov G, Lin Y, Macdonald JR, Pang AW, Shago M, Stockwell TB, Tsiamouri A, Bafna V, Bansal V, Kravitz SA, Busam DA, Beeson KY, McIntosh TC, Remington KA, Abril JF, Gill J, Borman J, Rogers YH, Frazier ME, Scherer SW, Strausberg RL, Venter JC. The diploid genome sequence of an individual human. PLoS Biol. 2007 Sep 4;5(10):e254.

8. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, Boutell JM, Bryant J, Carter RJ, Keira Cheetham R, Cox AJ, Ellis DJ, Flatbush MR, Gormley NA, Humphray SJ, Irving LJ, Karbelashvili MS, Kirk SM, Li H, Liu X, Maisinger KS, Murray LJ, Obradovic B, Ost T, Parkinson ML, Pratt MR, Rasolonjatovo IM, Reed MT, Rigatti R, Rodighiero C, Ross MT, Sabot A, Sankar SV, Scally A, Schroth GP, Smith ME, Smith VP, Spiridou A, Torrance PE, Tzonev SS, Vermaas EH, Walter K, Wu X, Zhang L, Alam MD, Anastasi C, Aniebo IC, Bailey DM, Bancarz IR, Banerjee S, Barbour SG, Baybayan PA, Benoit VA, Benson KF, Bevis C, Black PJ, Boodhun A, Brennan JS, Bridgham JA, Brown RC, Brown AA, Buermann DH, Bundu AA, Burrows JC, Carter NP, Castillo N, Chiara ECM, Chang S, Neil Cooley R, Crake NR, Dada OO, Diakoumakos KD, Dominguez-Fernandez B, Earnshaw DJ, Egbujor UC, Elmore DW, Etchin SS, Ewan MR, Fedurco M, Fraser LJ, Fuentes Fajardo KV, Scott Furey W, George D, Gietzen KJ, Goddard CP, Golda GS, Granieri PA, Green DE, Gustafson DL, Hansen NF, Harnish K, Haudenschild CD, Heyer NI, Hims MM, Ho JT, Horgan AM, Hoschler K, Hurwitz S, Ivanov DV, Johnson MQ, James T, Huw Jones TA, Kang GD, Kerelska TH, Kersey AD, Khrebtukova I, Kindwall AP, Kingsbury Z, Kokko-Gonzales PI, Kumar A, Laurent MA, Lawley CT, Lee SE, Lee X, Liao AK, Loch JA, Lok M, Luo S, Mammen RM, Martin JW, McCauley PG, McNitt P, Mehta P, Moon KW, Mullens JW, Newington T, Ning Z, Ling Ng B, Novo SM, O'Neill MJ, Osborne MA, Osnowski A, Ostadan O, Paraschos LL, Pickering L, Pike AC, Chris Pinkard D, Pliskin DP, Podhasky J, Quijano VJ, Raczy C, Rae VH, Rawlings SR, Chiva Rodriguez A, Roe PM, Rogers J, Rogert Bacigalupo MC, Romanov N, Romieu A, Roth RK, Rourke NJ, Ruediger ST, Rusman E, Sanches-Kuiper RM, Schenker MR, Seoane JM, Shaw RJ, Shiver MK, Short SW, Sizto NL, Sluis JP, Smith MA, Ernest Sohna Sohna J, Spence EJ, Stevens K, Sutton N, Szajkowski L, Tregidgo CL, Turcatti G, Vandevondele S, Verhovsky Y, Virk SM, Wakelin S, Walcott GC, Wang J, Worsley GJ, Yan J, Yau L, Zuerlein M, Mullikin JC, Hurles ME, McCooke NJ, West JS, Oaks FL, Lundberg PL, Klenerman D, Durbin R, Smith AJ. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 2008 Nov 6;456(7218):53-9.

9. Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen YJ, Makhijani V, Roth GT, Gomes X, Tartaro K, Niazi F, Turcotte CL, Irzyk GP, Lupski JR, Chinault C, Song Xz, Liu Y, Yuan Y, Nazareth L, Qin X, Muzny DM, Margulies M, Weinstock GM, Gibbs RA, Rothberg JM. The complete genome of an individual by massively parallel DNA sequencing. Nature. 2008 Apr 17;452(7189):872-6.

10. Kim JI, Ju YS, Park H, Kim S, Lee S, Yi JH, Mudge J, Miller NA, Hong D, Bell CJ, Kim HS, Chung IS, Lee WC, Lee JS, Seo SH, Yun JY, Woo HN, Lee H, Suh D, Kim HJ, Yavartanoo M, Kwak M, Zheng Y, Lee MK, Kim JY, Gokcumen O, Mills RE, Zaranek AW, Thakuria J, Wu X, Kim RW, Huntley JJ, Luo S, Schroth GP, Wu TD, Kim H, Yang KS, Park WY, Church GM, Lee C, Kingsmore SF, Seo JS. A highly annotated whole-genome sequence of a Korean individual. Nature. 2009 Aug 20;460(7258):1011-5.

11. Wang J, Wang W, Li R, Li Y, Tian G, Goodman L, Fan W, Zhang J, Li J, Zhang J, Guo Y, Feng B, Li H, Lu Y, Fang X, Liang H, Du Z, Li D, Zhao Y, Hu Y, Yang Z, Zheng H, Hellmann I, Inouye M, Pool J, Yi X, Zhao J, Duan J, Zhou Y, Qin J, Ma L, Li G, Yang Z, Zhang G, Yang B, Yu C, Liang F, Li W, Li S, Li D, Ni P, Ruan J, Li Q, Zhu H, Liu D, Lu Z, Li N, Guo G, Zhang J, Ye J, Fang L, Hao Q,

Chen Q, Liang Y, Su Y, San A, Ping C, Yang S, Chen F, Li L, Zhou K, Zheng H, Ren Y, Yang L, Gao Y, Yang G, Li Z, Feng X, Kristiansen K, Wong GK, Nielsen R, Durbin R, Bolund L, Zhang X, Li S, Yang H, Wang J. The diploid genome sequence of an Asian individual. Nature. 2008 Nov 6;456(7218):60-5.

12. Korbel JO, Urban AE, Affourtit JP, Godwin B, Grubert F, Simons JF, Kim PM, Palejev D, Carriero NJ, Du L, Taillon BE, Chen Z, Tanzer A, Saunders AC, Chi J, Yang F, Carter NP, Hurles ME, Weissman SM, Harkins TT, Gerstein MB, Egholm M, Snyder M. Paired-end mapping reveals extensive structural variation in the human genome. Science. 2007 Oct 19;318(5849):420-6.
13. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, Gonzalez JR, Gratacos M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Zhang J, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurles ME. Global variation in copy number in the human genome. Nature. 2006 Nov 23;444(7118):444-54.
14. Marian AJ. Nature's genetic gradients and the clinical phenotype. Circ Cardiovasc Genet. 2009 Dec;2(6):537-9.
15. Maron BJ, Doerer JJ, Haas TS, Tierney DM, Mueller FO. Sudden deaths in young competitive athletes: analysis of 1866 deaths in the United States, 1980-2006. Circulation. 2009 Mar 3;119(8):1085-92.
16. Marian AJ. Hypertrophic cardiomyopathy: from genetics to treatment. Eur J Clin Invest. 2010 Apr;40(4):360-9.
17. Maron BJ. Hypertrophic cardiomyopathy: a systematic review. JAMA. 2002 Mar 13;287(10):1308-20.
18. Epstein RS, Moyer TP, Aubert RE, O Kane DJ, Xia F, Verbrugge RR, Gage BF, Teagarden JR. Warfarin genotyping reduces hospitalization rates results from the MM-WES (Medco-Mayo Warfarin Effectiveness study). J Am Coll Cardiol. 2010 Jun 22;55(25):2804-12.
19. Simon T, Verstuyft C, Mary-Krause M, Quteineh L, Drouet E, Méneveau N, Steg PG, Ferrières J, Danchin N, Becquemont L; French Registry of Acute ST-Elevation and Non-ST-Elevation Myocardial Infarction (FAST-MI) Investigators. Genetic determinants of response to clopidogrel and cardiovascular events. N Engl J Med. 2009 Jan 22;360(4):363-75.
20. Mega JL, Close SL, Wiviott SD, Shen L, Hockett RD, Brandt JT, Walker JR, Antman EM, Macias W, Braunwald E, Sabatine MS. Cytochrome p-450 polymorphisms and response to clopidogrel. N Engl J Med. 2009 Jan 22;360(4):354-62.
21. Voora D, Shah SH, Spasojevic I, Ali S, Reed CR, Salisbury BA, Ginsburg GS. The SLCO1B1*5 genetic variant is associated with statin-induced side effects. J Am Coll Cardiol. 2009 Oct 20;54(17):1609-16.
22. SEARCH Collaborative Group, Link E, Parish S, Armitage J, Bowman L, Heath S, Matsuda F, Gut I, Lathrop M, Collins R. SLCO1B1 variants and statin-induced myopathy — a genomewide study. N Engl J Med. 2008 Aug 21;359(8):789-99.
23. Liggett SB, Mialet-Perez J, Thaneemit-Chen S, Weber SA, Greene SM, Hodne D, Nelson B, Morrison J, Domanski MJ, Wagoner LE, Abraham WT, Anderson JL, Carlquist JF, Krause-Steinrauf HJ, Lazzeroni LC, Port JD, Lavori PW, Bristow MR. A polymorphism within a conserved beta(1)-adrenergic receptor motif alters cardiac function and beta-blocker response in human heart failure. Proc Natl Acad Sci U S A. 2006 Jul 25;103(30):11288-93.
24. Thompson JF, Hyde CL, Wood LS, Paciga SA, Hinds DA, Cox DR, Hovingh GK, Kastelein JJ. Comprehensive whole-genome and candidate gene analysis for response to statin therapy in the Treating to New Targets (TNT) cohort. Circ Cardiovasc Genet. 2009 Apr;2(2):173-81.
25. Roden DM, Viswanathan PC. Genetics of acquired long QT syndrome. J Clin Invest. 2005 Aug;115(8):2025-32. Review.