Published in final edited form as:

Trends Cardiovasc Med. 2014 February ; 24(2): 53–60. doi:10.1016/j.tcm.2013.06.004.

Successes and Challenges of Using Whole Exome Sequencing to Identify Novel Genes Underlying an Inherited Predisposition for Thoracic Aortic Aneurysms and Acute Aortic Dissections

Dianna M. Milewicz, M.D., Ph.D.,

Division of Medical Genetics, Department of Internal Medicine at the University of Texas Health Science Center at Houston (UT Health) in Houston, TX 77030, USA

Ellen Regalado,

Division of Medical Genetics, Department of Internal Medicine at the University of Texas Health Science Center at Houston (UT Health) in Houston, TX 77030, USA

Jay Shendure,

Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA

Deborah A. Nickerson, and

Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA

Dongchuan Guo

Division of Medical Genetics, Department of Internal Medicine at the University of Texas Health Science Center at Houston (UT Health) in Houston, TX 77030, USA

Abstract

Thoracic aortic aneurysms involving the aortic root and/or ascending aorta can lead to acute aortic dissections. Approximately 20% of patients with thoracic aortic aneurysms and dissections (TAAD) have a family history of the disease, referred to as familial TAAD (FTAAD), which can be inherited in an autosomal dominant manner with variable expression with respect to disease presentation, age of onset and associated features. Whole exome sequencing (WES) has been used to identify causative mutations in novel genes for TAAD. The strategy used to reduce the large number of rare variants identified using WES is to sequence distant relatives with TAAD and filter for heterozygous rare variants that are shared between the relatives, predicted to disrupt protein function and segregate with the TAAD phenotype in other family members. Further validation of putative genes by sequencing of additional families with TAAD has successfully identified novel genes for FTAAD.

Address correspondence to: Dianna M. Milewicz, M.D., Ph.D., 6431 Fannin, Suite 6.100, Houston, Texas 77030, U.S.A.; (+1) 713 500 6715; fax: (+1) 713 500 0693; Dianna.M.Milewicz@uth.tmc.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Web Resources

NCBI dbSNP, www.ncbi.nlm.nih.gov/projects/SNP/

UCSC Genome Browser, <http://www.genome.ucsc.edu/>

Ensembl Genome Browser, <http://www.ensembl.org/index.html>

1000 Genome, <http://www.1000genomes.org/>

Exome Variant Server (NHLBI GO ESP), <http://evs.gs.washington.edu/EVS/>

PolyPhen-2 prediction of functional effects of nonsynonymous variants, <http://genetics.bwh.harvard.edu/pph2/>

Introduction

Thoracic aortic aneurysms leading to acute aortic dissections (TAAD) are a common cause of premature deaths, ranking as high as the 15th leading cause of death in the United States (Hoyert et al., 2001). If an individual is known to have a genetic predisposition to TAAD, clinical management can be initiated to prevent premature death due to aortic dissection. We and others have determined that up to 20% of TAAD patients without a genetic syndrome have a family history of TAAD, referred to as familial TAAD (FTAAD), indicating the disease has a significant genetic component (Biddinger et al., 1997; Coady et al., 1999). Analysis of FTAAD pedigrees showed TAAD is primarily inherited in families as an autosomal dominant condition with reduced penetrance, which indicates that a single gene mutation is the most likely cause of the disease (Albornoz et al., 2006; Milewicz et al., 1998). The disease presentation is variable, not only in the age of onset and aortic disease presentation, but also in the presence of other clinical features that segregate with TAAD, such as congenital defects (e.g., bicuspid aortic valve or BAV, patent ductus arteriosus or PDA) and vascular diseases elsewhere (e.g., intracranial aneurysms, coronary artery disease, occlusive cerebrovascular disease) (Guo et al., 2011; Loscalzo et al., 2007; Regalado et al., 2011a; Tran-Fadulu et al., 2009). Clinical heterogeneity of FTAAD suggests that there are multiple genes involved in the disease. Similar to other adult onset cardiovascular diseases due to single gene mutations, there is significant genetic heterogeneity for FTAAD.

Multiple genes for FTAAD have been identified using different genetic techniques. Positional cloning successfully identified mutations in three genes causing FTAAD: *TGFBR2* encoding transforming growth factor- β receptor type II and *ACTA2* and *MYH11* encoding the smooth muscle specific isoforms of α -actin and β -myosin, respectively (Guo et al., 2007; Pannu et al., 2005; Pannu et al., 2007; Zhu et al., 2006). A candidate gene approach was used to identify FTAAD genes: *TGFBR1* encoding transforming growth factor- β receptor type I, *FBN1* encoding fibrillin, and *MYLK* encoding myosin light chain kinase (Francke et al., 1995; Milewicz et al., 1996; Tran-Fadulu et al., 2009; Wang et al., 2010). Thus far, the genes implicated for FTAAD highlight disruption of two distinct molecular pathways involved in the etiology of the thoracic aortic disease: TGF- β signaling and smooth muscle cell (SMC) contractile function (Milewicz et al., 2008). Furthermore, specific clinical manifestations associated with mutations in these genes helped explain the clinical heterogeneity observed for this condition (Guo et al., 2009; Loeys et al., 2005; van de Laar et al., 2011; Zhu et al., 2006).

Despite identifying six genes for FTAAD using these genetic strategies, only about 20% of TAAD families have mutations in these genes. This fact suggests that the search for novel genes containing rare variants leading to disease will be complicated by the possibility that each gene will contribute to disease in only a small proportion of families. We have recently turned to whole exome sequencing (WES) to identify rare variants in novel genes for TAAD and increase the pace of novel gene identification. As recently reviewed (Marian, 2012), WES provides an unbiased assessment of genetic variants in the protein coding regions of the genome. Since approximately 85% of causative mutations for Mendelian diseases are found in the coding regions or in canonical splice sites, WES is an efficient approach to identify causative mutations for these diseases (Ng et al., 2009; Turner et al., 2009). This review will outline the strategy we have used to successfully identify rare gene variants causing familial TAAD using WES, as well as discuss the challenges going forward in using this approach (Boileau et al., 2012; Guo DC et al., 2013; Regalado et al., 2011b).

Approach

The commercial human exome capture arrays and next generation sequencing technologies are designed to capture and sequence approximately 97% - 99% of the exons and flanking introns that are in current genomic databases (Reinhardt et al., 1996), although in practice performance can fall modestly short of this (Clark et al., 2011). Additionally, computer programs have been developed to map sequence reads to the reference human genome, identify genetic variants, annotate amino acid changes, predict the impact of an amino acid substitution on the structure and function of a human protein, and search minor allele frequencies in different populations using multiple publicly available databases, such as the National Heart Lung and Blood Institute GO Exome Sequencing Project (NHLBI GO ESP), 1000 Genome Project, and NCBI dbSNP databases. Thus, WES has evolved into a robust technology for identifying coding variation at the genome-wide level and enabling researchers to identify causative rare variants predisposing to diseases rapidly and efficiently (Marneros et al., 2013).

The strategy that we have used to identify causative mutations for familial TAAD is to perform WES using DNA from two distantly related affected individuals in a family. The initial families selected for WES had TAAD inherited in an autosomal dominant manner, but each family had a different aortic disease presentation or evidence that the defective gene was causing additional vascular diseases in the family (Fig. 1). Family TAA549 shows dominant inheritance of TAAD with reduced penetrance, but also has family members with intracranial, iliac artery and abdominal aortic aneurysms. Family members of TAA288 has autosomal dominant inheritance of thoracic aortic root aneurysms, along with intracranial aneurysms and mild systemic features of Marfan syndrome (Regalado et al., 2011a). Family TAA292 demonstrates aggressive aortic disease with early onset aortic dissections in both men and women (Tran-Fadulu et al., 2006). Affected individuals of family TAA059 have aortic root aneurysms associated with bicuspid aortic valves. Distantly related family members who were clearly affected with TAAD were selected for WES (circled symbols in the pedigrees).

Approximately 15,000-20,000 single nucleotide variants (SNVs) and short indel variants are identified by WES in each individual. Candidate disease-causing variants are identified from this large number of variants using a straightforward filtering strategy (Fig. 2). First, non-synonymous, nonsense, loss-of-stop codon variants, insertion/deletions, and splice site variants are included for further study since we expect the disease causing variant to disrupt protein function; all other variants are excluded. Second, since we expect the disease variant to be rare, SNVs with a minor allele frequency (MAF) of more than 0.2% in 1500 locally generated exome sequencing controls and SNVs with a MAF of more than 0.05% in the NHLBI GO ESP database are excluded. Third, homozygous SNVs are excluded since autosomal dominant inheritance of TAAD predicts a heterozygous rare variant is causing the disease. Based on the first three filters, approximately 250-400 candidate SNVs are identified in each individual. Fourth, only heterozygous variants shared by the two affected relatives are considered, which further narrows down the number of possible disease-causing variants for each family. Additionally, because the initial filtering step is to identify heterozygous SNVs shared between affected relatives, the number of false positive SNVs is significantly reduced. The false positive rates vary depending on the stringency of filtering criteria and WES data quality; the number of false positive calls can be as low as 5% but is often much higher (Marian, 2012). We have performed Sanger sequencing assay to validate rare SNVs identified by WES, and determined that the false positive rate was approximately 3%. The number of candidate variants left depends on the degree of relationship between the two affected individuals. Using the filtering strategy described above, the average number of variants left are 15, 27, and 38 for families wherein the relatives sequenced shared 1/16, 1/8,

and 1/4 of their genomes, respectively (Fig. 3). To further reduce the number of candidate variants, DNA from additional family members are sequenced and only the candidate gene variants that segregated with TAAD are studied further. In families with available genome-wide linkage data, candidate gene variants that are located within putative disease loci are also prioritized to further reduce the number of candidate SNVs (Fig. 4).

The remaining candidate rare variants are prioritized for validation based on the following criteria: a) genes encoding proteins that are involved in known cellular pathways of TAAD, specifically the TGF- β pathway and SMC contraction; b) missense SNVs that are predicted to significantly alter protein function based on PolyPhen-2 or SIFT programs and variants that cause frameshift, nonsense, and loss of stop codon; and c) candidate SNVs in genes that are highly expressed in SMCs. After prioritizing the candidate genes, Sanger sequencing assays are performed in 200 affected probands from unrelated FTAAD families to validate a putative gene by identifying additional families with rare variants in the candidate gene that segregate with TAAD. Finally, biochemical, cell biology, pathology, and/or mouse model studies are done to link the putative mutant gene to the clinical phenotype.

The above strategy is based on the premise that there is a single, rare gene variant within or disrupting exomes or flanking introns that is responsible for the familial disease. Therefore, the critical elements for significantly reducing the number of candidate SNVs and successfully identifying the causative mutation are the following: a) truly affected individuals with early age of onset of aortic dissection or aneurysmal disease are selected for WES to confirm that these family members have the disease-causing variant; b) pedigree analysis and correct assignment of mode of inheritance; c) selection of most distantly related affected individuals for WES or using families for which linkage mapping data is available so that the rare variant causing disease can be rapidly identified.

Identification of novel FTAAD genes encoding proteins in the TGF- β pathway

The “low hanging fruits” as far as using WES and large pedigrees to identify novel genes for FTAAD were rare variants in genes encoding proteins in the canonical TGF- β signaling pathways, specifically *SMAD3* and *TGFB2*, encoding Smad3 and TGF- β 2, respectively (Boileau et al., 2012; Regalado et al., 2011b). Sequencing first cousins once removed in family TAA549 rapidly narrowed down the number of potentially disease-causing rare variants to 11 variants, one of which was a novel frame shift mutation, p. N218fs (c. 652delA), in the *SMAD3* gene (Fig. 1). This variant segregated with the vascular diseases in this family with a LOD score of 2.52. The gene was rapidly validated after sequencing of 181 probands with FTAAD identified three additional *SMAD3* mutations in 4 families, resulting in a combined LOD score of 5.21. *SMAD3* mutations are responsible for 2% of familial TAAD and were identified in families with TAAD alone, but also in families with TAAD, intracranial aneurysms, aortic and bilateral iliac aneurysms and mild skeletal features of Marfan syndrome (Regalado et al., 2011a). In addition, other investigators have also identified *SMAD3* mutations using a positional cloning approach causing early onset osteoarthritis and aneurysms (van de Laar et al., 2011). A subset of *SMAD3* mutations are frameshift or nonsense mutations which are predicted to lead to nonsense mediated decay and *SMAD3* haploinsufficiency. Other disease-causing mutations are located in the MH2 protein domain of *SMAD3* and are predicted to affect hydrogen-bond formation, thus potentially disrupting protein-protein interactions.

In a similar manner, *TGFB2* was identified as the defective gene causing thoracic aortic disease, intracranial aneurysms and minimal systemic features of the Marfan syndrome in two large unrelated families using genome-wide linkage analysis, followed by whole exome

sequencing of affected relatives (Fig. 1) (Boileau et al., 2012). In these families, a frameshift mutation in exon 6 and a nonsense mutation in exon 4 were identified that segregated with disease with a combined LOD score of 7.7. Sanger sequencing of 276 probands from families with inherited thoracic aortic disease identified two additional *TGFB2* mutations, a nonsense and frameshift mutation. *TGFB2* encodes the transforming growth factor beta-2 (TGF- β 2) and three of the identified mutations are predicted to lead to nonsense mediated RNA decay resulting in haploinsufficiency of TGF- β 2 and predisposing to thoracic aortic disease.

Mutations in these novel genes indicate that haploinsufficiency of *SMAD3* or *TGFB2* is the initiating step in the pathogenic events leading to thoracic aortic disease. Furthermore, the missense mutations in *TGFB1* and *TGFB2* either have been shown or are predicted to disrupt kinase function, which is critical for propagating TGF- β signaling, and thus can be classified as loss of function mutations (Inamoto et al., 2010; Mizuguchi et al., 2004; Regalado et al., 2011b; Stheneur et al., 2008; van de Laar et al., 2011). Despite the fact that many of the mutations in these genes are predicted to decrease TGF- β signaling, these defective genes trigger poorly understood signaling events that lead to a secondary increase of TGF- β signaling in SMCs in aortic tissue, predominantly assayed by increased nuclear phosphorylated (activated) SMAD2 immunostaining in diseased tissue. Although the data supports the conclusion that the initiating event leading to vascular diseases is diminished TGF- β signaling, further work is needed to unravel the pathways that are secondarily altered in the pathogenesis of thoracic aortic disease and cause paradoxical increase in *TGFB2* expression and TGF- β signaling in diseased aortic tissue.

Identification of novel FTAAD genes encoding proteins involved in SMC contraction

Similar to the identification of rare variants in genes encoding proteins in the canonical TGF- β pathway, we also rapidly identified mutations in another gene regulating SMC contraction as a cause of FTAAD. We and others had previously identified mutations in genes encoding the major proteins forming the thin and thick filaments of smooth muscle: α -actin (*ACTA2*) and myosin heavy chain (*MYH11*) (Guo et al., 2007; Zhu et al., 2006). In addition, we identified mutations in *MYLK* encoding the myosin light chain kinase (MLCK) which is a ubiquitously expressed kinase whose only known target of phosphorylation is the 20kDa regulatory light chain (RLC) of smooth and nonmuscle myosin II (Ding et al., 2009; Isotani et al., 2004; Wang et al., 2010). MLCK is highly expressed in SMCs, where phosphorylation of RLC by MLCK initiates the physiologic contraction of smooth muscle within hollow organs. Exome sequencing of distant relatives with TAAD, followed by sequencing of additional probands with FTAAD (Fig. 1) identified the same rare variant, *PRKG1* c.530G>A, p.R177Q, in four unrelated families (Guo DC et al., 2013). This mutation segregated with TAAD in these families with a combined two-point LOD score of 7.9. This *PRKG1* R177Q mutation demonstrated complete penetrance for TAAD with early onset aortic dissections in the majority of family members as young as 17 years of age.

PRKG1 encodes a type I cGMP-dependent protein kinase (PKG I) that is activated upon binding of cGMP and controls SMC relaxation (Murthy, 2006; Nakamura et al., 2007). Although the R177Q disrupts binding to the high affinity cGMP binding site within the regulatory domain, the mutant PKG I is constitutively active even in the absence of cGMP. The increased PKG I activity leads to decreased phosphorylation of the myosin regulatory light chain in fibroblasts and is predicted to cause decreased contraction of vascular SMCs.

Although rare variants in *PRKG1* were found in other FTAAD families and are present in the population, only the *PRKG1* mutation resulting in R177Q was confirmed to be disease-

causing, most likely due the unique gain-of-function effect of this mutation on PKG I α activity. These data illustrate the complexity of identifying rare variants with complex, adult onset diseases, even when the variant is highly penetrant. Association studies based on a burden test of rare variants in *PRKG1* with thoracic aortic disease would not demonstrate an association, and studies with the specific *PRKG1* p.R177Q variant would only show an association in a large cohort of patients or if family-based studies were pursued.

Surprisingly, we have not identified mutations in the genes for other proteins in the contractile complex as a cause for FTAAD. Hypertrophic cardiomyopathy is a common form of inherited heart disease characterized by left ventricular wall thickening and contractile dysfunction. Several genes have been implicated in this condition mostly genes encoding sarcomeric proteins such as myosin-7 or cardiac muscle beta myosin heavy chain (*MYH7*), cardiac myosin-binding protein C (*MYBPC3*), cardiac muscle troponin T (*TNNT2*) and tropomyosin (*TPM1*). Altogether mutations in these genes explain more than 80% of the disease (Frey et al., 2012). Other genes encoding sarcomeric proteins such as actin (*ACTC1*), alpha-actinin-2 (*ACTN2*) and titin (*TTN*) have also been identified, but are individually rare causes of HCM. Although we identified mutations in the genes for the two major proteins in the contractile complex (*MYH11* and *ACTA2*) and kinases controlling phosphorylation of the RLC (*MYLK* and *PRKG1*), mutations in additional genes coding smooth muscle contractile proteins, such as calmodulin (*CALM1*), caldesmon (*CALD1*), calponin (*CNN1*) and RLC (*MYL2*), have not been identified as causes for FTAAD.

Identification of Further Genes for FTAAD: Where are we?

Whole exome sequencing of affected relative pairs, followed by validation of putative genes by sequencing additional families, has successfully identified novel genes for FTAAD. The genes reported to date have disrupted known pathways to thoracic aortic disease and were easily confirmed with minimal additional molecular, cellular or animal studies. Although we have genetic evidence to strongly support some additional genes as disease-causing for FTAAD, these genes are predicted to disrupt new pathways to the disease or have an unknown function. The molecular, cellular and animal studies required to link the gene variant to the disease phenotype are extensive and often require development of novel techniques and establishing new collaborations. As an example, WES and segregation of putative disease-causing rare variants have been completed in a large family with FTAAD, and only one rare gene variant remains as a possible cause of TAAD based on segregation of the variant with disease in the family. The variant disrupts an amino acid critical for the function of the corresponding protein. The gene encodes a protein that is highly expressed in SMCs and is responsible for controlling many SMC functions. But we have not been able to validate this gene since no other families in our cohort have a mutation in this gene. Is the genetic heterogeneity underlying FTAAD extensive enough that some families will not only have their own private mutation but also a “private gene” causing their familial thoracic aortic disease?

Future Directions

Our research group has successfully used an approach involving WES and large families to identify defective genes resulting in disruption of known disease pathways associated with FTAAD. The identification of further genes for FTAAD highlights the genetic heterogeneity and complexity of identifying causative genes for this disease. Going forward, there are some hurdles to overcome to identify causative mutations for FTAAD as discussed below. First, there is significant genetic heterogeneity for FTAAD and the vast majority of the defective genes identified to date are responsible for disease in only a very small proportion of families. To overcome such limitation, large cohorts of FTAAD families with DNA

samples and clinical data collected from all affected family members, including those with other vascular diseases, are needed to validate putative genes, confirm segregation of rare gene variants with the disease, and define the vascular disease phenotype associated with novel gene mutations. Furthermore, collaboration between groups with large cohorts of patients is essential to rapidly and efficiently validate candidate genes. Extending collaborative efforts to other researchers with WES data on cohorts of patients with other vascular disease, such as intracranial aneurysms, may help to define the spectrum of vascular diseases associated with novel FTAAD genes. Second, biological studies to establish links between the defective gene and the disease are currently the rate limiting step to confirm a gene as causative. The lack of readily available model system, such as the zebrafish model system, to aid in the confirmation that disruption of a gene alters aortic function hinders the identification and characterization of novel genes for FTAAD. Animal models can provide biological evidences for confirming mutations (Lindsay et al., 2012; Wang et al., 2010)); however, mouse models are costly and time-consuming. Third, the filtering strategy used to reduce the number of putative disease-causing rare variants may inadvertently exclude disease-causing rare variants. If we fail to identify the disease-causing variant in a large number of FTAAD families, we plan to relax our filtering strategies and reassess the putative variants identified. Lastly, WES has limitations as far as the extent of the genome sequenced, types of variants detected, and coverage. We know that causative mutations can be located a distance from coding regions or can be missed when WES is used. Genomic copy number variants (CNV) and rare variants that are located at introns and disrupt mRNA splicing can be the cause of FTAAD and will not readily be identified by WES (Guo et al., 2008; Kuang et al., 2011). Several software programs, such as XHMM and CoNIFER, have been recently developed to estimate copy number variants in genes based on the read depth of WES (Fromer et al., 2012; Krumm et al., 2012). These software programs use principal components analysis to reduce background noise and are able to reliably predict more than 90% of both *de novo* and inherited CNVs involving three or more consecutive exons. (Krumm et al., 2012). In addition, approximately 2 -3% of exomes are not captured efficiently by exome capture arrays for WES, especially in regions with high GC content. These limitations will be overcome by whole genome sequencing in the future.

Acknowledgments

We are extremely grateful to the families who participated in these studies. We would like to acknowledge the following sources that provided support for these studies: RO1 HL62594 (D.M.M.), P50HL083794-01 (D.M.M.), UL1 RR024148 (CTSA), Vivian L. Smith Foundation (D.M.M.), TexGen Foundation (D.M.M.), the Richard T. Pisani Funds (D.M.M.), the Lung Cohorts Sequencing Project (HL-102923), the WHI Sequencing Project (HL-102924), the Heart Cohorts Sequencing Project (HL-103010), the Broad Institute Sequencing Project (HL-102925), the Northwest Genomics Center Sequencing Project (HL-102926, J. S and D.A.N).

Reference List

- Albornoz G, Coady MA, Roberts M, et al. Familial thoracic aortic aneurysms and dissections--incidence, modes of inheritance, and phenotypic patterns. *Ann Thorac Surg.* 2006; 82:1400–1405. [PubMed: 16996941]
- Biddinger A, Rocklin M, Coselli J, et al. Familial thoracic aortic dilatations and dissections: a case control study. *J Vasc Surg.* 1997; 25:506–511. [PubMed: 9081132]
- Boileau C, Guo DC, Hanna N, et al. TGFB2 mutations cause familial thoracic aortic aneurysms and dissections associated with mild systemic features of Marfan syndrome. *Nat Genet.* 2012; 44:916–921. [PubMed: 22772371]
- Clark MJ, Chen R, Lam HY, et al. Performance comparison of exome DNA sequencing technologies. *Nat Biotechnol.* 2011; 29:908–914. [PubMed: 21947028]
- Coady MA, Davies RR, Roberts M, et al. Familial patterns of thoracic aortic aneurysms. *Arch Surg.* 1999; 134:361–367. [PubMed: 10199307]

- Ding HL, Ryder JW, Stull JT, et al. Signaling processes for initiating smooth muscle contraction upon neural stimulation. *J Biol Chem*. 2009; 284:15541–15548. [PubMed: 19349274]
- Francke U, Berg MA, Tynan K, et al. A Gly1127Ser Mutation in An Egf-Like Domain of the Fibrillin-1 Gene Is A Risk Factor for Ascending Aortic-Aneurysm and Dissection. *American Journal of Human Genetics*. 1995; 56:1287–1296. [PubMed: 7762551]
- Frey N, Luedde M, Katus HA. Mechanisms of disease: hypertrophic cardiomyopathy. *Nat Rev Cardiol*. 2012; 9:91–100. [PubMed: 22027658]
- Fromer M, Moran JL, Chambert K, et al. Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth. *Am J Hum Genet*. 2012; 91:597–607. [PubMed: 23040492]
- Guo DC, Regalado E, Casteel DE, et al. Recurrent gain of function mutation in PRKG1 causes thoracic aortic aneurysms and acute aortic dissections. *Am J Hum Genet*. 2013 (in press).
- Guo DC, Gupta P, Tran-Fadulu V, et al. An FBN1 pseudoexon mutation in a patient with Marfan syndrome: confirmation of cryptic mutations leading to disease. *J Hum Genet*. 2008; 53:1007–1011. [PubMed: 18795226]
- Guo DC, Pannu H, Papke CL, et al. Mutations in smooth muscle alpha-actin (ACTA2) lead to thoracic aortic aneurysms and dissections. *Nat Genet*. 2007; 39:1488–1493. [PubMed: 17994018]
- Guo DC, Papke CL, Tran-Fadulu V, et al. Mutations in smooth muscle alpha-actin (ACTA2) cause coronary artery disease, stroke, and moyamoya disease, along with thoracic aortic disease. *Am J Hum Genet*. 2009; 84:617–627. [PubMed: 19409525]
- Guo DC, Regalado ES, Minn C, et al. Familial thoracic aortic aneurysms and dissections: identification of a novel locus for stable aneurysms with a low risk for progression to aortic dissection. *Circ Cardiovasc Genet*. 2011; 4:36–42. [PubMed: 21163914]
- Hoyert DL, Arias E, Smith BL, et al. Deaths: final data for 1999. *Natl Vital Stat Rep* 2001. 49:1–113.
- Inamoto S, Kwartler CS, Lafont AL, et al. TGFB2 Mutations Alter Smooth Muscle Cell Phenotype and Predispose to Thoracic Aortic Aneurysms and Dissections. *Cardiovasc Res*. 2010; 88:520–529. [PubMed: 20628007]
- Isotani E, Zhi G, Lau KS, et al. Real-time evaluation of myosin light chain kinase activation in smooth muscle tissues from a transgenic calmodulin-biosensor mouse. *Proc Natl Acad Sci U S A*. 2004; 101:6279–6284. [PubMed: 15071183]
- Krumm N, Sudmant PH, Ko A, et al. Copy number variation detection and genotyping from exome sequence data. *Genome Res*. 2012; 22:1525–1532. [PubMed: 22585873]
- Kuang SQ, Guo DC, Prakash SK, et al. Recurrent chromosome 16p13.1 duplications are a risk factor for aortic dissections. *PLoS Genet*. 2011; 7:e1002118. [PubMed: 21698135]
- Lindsay ME, Schepers D, Bolar NA, et al. Loss-of-function mutations in TGFB2 cause a syndromic presentation of thoracic aortic aneurysm. *Nat Genet*. 2012; 44:922–927. [PubMed: 22772368]
- Loeys BL, Chen J, Neptune ER, et al. A syndrome of altered cardiovascular, craniofacial, neurocognitive and skeletal development caused by mutations in TGFB1 or TGFB2. *Nat Genet*. 2005; 37:275–281. [PubMed: 15731757]
- Loscalzo ML, Goh DL, Loeys B, et al. Familial thoracic aortic dilation and bicommissural aortic valve: a prospective analysis of natural history and inheritance. *Am J Med Genet A*. 2007; 143:1960–1967. [PubMed: 17676603]
- Marian AJ. Challenges in medical applications of whole exome/genome sequencing discoveries. *Trends Cardiovasc Med*. 2012; 22:219–223. [PubMed: 22921985]
- Marneros AG, Beck AE, Turner EH, et al. Mutations in KCTD1 Cause Scalp-Ear-Nipple Syndrome. *Am J Hum Genet*. 2013; 92:621–626. [PubMed: 23541344]
- Milewicz DM, Chen H, Park ES, et al. Reduced penetrance and variable expressivity of familial thoracic aortic aneurysms/dissections. *Am J Cardiol*. 1998; 82:474–479. [PubMed: 9723636]
- Milewicz DM, Guo D, Tran-Fadulu V, et al. Genetic Basis of Thoracic Aortic Aneurysms and Dissections: Focus on Smooth Muscle Cell Contractile Dysfunction. *Annu Rev Genomics Hum Genet*. 2008; 9:283–302. [PubMed: 18544034]
- Milewicz DM, Michael K, Fisher N, et al. Fibrillin-1 (FBN1) mutations in patients with thoracic aortic aneurysms. *Circulation*. 1996; 94:2708–2711. [PubMed: 8941093]

- Mizuguchi T, Collod-Beroud G, Akiyama T, et al. Heterozygous TGFBR2 mutations in Marfan syndrome. *Nat Genet.* 2004; 36:855–860. [PubMed: 15235604]
- Murthy KS. Signaling for contraction and relaxation in smooth muscle of the gut. *Annu Rev Physiol.* 2006; 68:345–374. [PubMed: 16460276]
- Nakamura K, Koga Y, Sakai H, et al. cGMP-dependent relaxation of smooth muscle is coupled with the change in the phosphorylation of myosin phosphatase. *Circ Res.* 2007; 101:712–722. [PubMed: 17673671]
- Ng SB, Turner EH, Robertson PD, et al. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature.* 2009; 461:272–276. [PubMed: 19684571]
- Pannu H, Fadulu V, Chang J, et al. Mutations in transforming growth factor-beta receptor type II cause familial thoracic aortic aneurysms and dissections. *Circulation.* 2005; 112:513–520. [PubMed: 16027248]
- Pannu H, Tran-Fadulu V, Papke CL, et al. MYH11 mutations result in a distinct vascular pathology driven by insulin-like growth factor I and angiotensin II. *Hum Mol Genet.* 2007; 16:3453–3462.
- Regalado E, Medrek S, Tran-Fadulu V, et al. Autosomal dominant inheritance of a predisposition to thoracic aortic aneurysms and dissections and intracranial saccular aneurysms. *Am J Med Genet A.* 2011; 155A:2125–2130. [PubMed: 21815248]
- Regalado ES, Guo DC, Villamizar C, et al. Exome Sequencing Identifies SMAD3 Mutations as a Cause of Familial Thoracic Aortic Aneurysm and Dissection With Intracranial and Other Arterial Aneurysms. *Circ Res.* 2011; 109:680–686. [PubMed: 21778426]
- Reinhardt DP, Sasaki T, Dzamba BJ, et al. Fibrillin-1 and fibulin-2 interact and are colocalized in some tissues. *J Biol Chem.* 1996; 271:19489–19496. [PubMed: 8702639]
- Steneur C, Collod-Beroud G, Faivre L, et al. Identification of 23 TGFBR2 and 6 TGFBR1 gene mutations and genotype-phenotype investigations in 457 patients with Marfan syndrome type I and II, Loeys-Dietz syndrome and related disorders. *Hum Mutat.* 2008; 29:E284–E295. [PubMed: 18781618]
- Tran-Fadulu V, Chen JH, Lemuth D, et al. Familial thoracic aortic aneurysms and dissections: three families with early-onset ascending and descending aortic dissections in women. *Am J Med Genet A.* 2006; 140:1196–1202. [PubMed: 16646045]
- Tran-Fadulu V, Pannu H, Kim DH, et al. Analysis of multigenerational families with thoracic aortic aneurysms and dissections due to TGFBR1 or TGFBR2 mutations. *J Med Genet.* 2009; 46:607–613. [PubMed: 19542084]
- Turner EH, Lee C, Ng SB, et al. Massively parallel exon capture and library-free resequencing across 16 genomes. *Nat Methods.* 2009; 6:315–316. [PubMed: 19349981]
- van de Laar IM, Oldenburg RA, Pals G, et al. Mutations in SMAD3 cause a syndromic form of aortic aneurysms and dissections with early-onset osteoarthritis. *Nat Genet.* 2011; 43:121–126. [PubMed: 21217753]
- Wang L, Guo DC, Cao J, et al. Mutations in Myosin light chain kinase cause familial aortic dissections. *Am J Hum Genet.* 2010; 87:701–707. [PubMed: 21055718]
- Zhu L, Vranckx R, Khau Van KP, et al. Mutations in myosin heavy chain 11 cause a syndrome associating thoracic aortic aneurysm/aortic dissection and patent ductus arteriosus. *Nat Genet.* 2006; 38:343–349. [PubMed: 16444274]

Fig. 1. Pedigrees of families TAA288 (*TGFB2* mutation), TAA216 (*PRKG1* mutation), TAA549 (*SMAD3* mutation), and TAA059 with TAAD and additional clinical phenotypes, such as abdominal aortic aneurysm, other arterial aneurysms, and bicuspid aortic valve. Squares represent males and circles represent females. The symbols represent different diagnoses as shown in the legend. The age at diagnosis (“dx”) is shown in years and “d” indicates age at death. Circled symbols indicate individuals whose DNA was used in WES assay.

Fig. 2.

Strategy of filtering WES data to decrease the number of candidate gene variants in a family. MAF- minor allele frequency; NHLBI GO ESP- National Heart Lung and Blood Institute GO Exome Sequencing Project.

Fig. 3.

The number of candidate rare variants remaining after filtering is dependent on the degree of relatedness between affected family members selected for WES. The x-axis indicates the degree of relatedness among affected individuals who were selected for WES and the y-axis indicates the number of candidate gene variants remaining in each family after the filtering strategy is completed.

Fig. 4.

Whole genome linkage analysis using DNA from large FTAAD families reduces the number of candidate gene variants identified. The x-axis indicates the families used for WES and the y-axis indicates the number of candidate gene variants identified. Open bars indicate the number of candidate variants identified by WES; and filled bars indicate the number of candidate variants identified by WES and located within putative disease loci identified by linkage analysis.