Identification of *COL3A1* variants associated with sporadic thoracic aortic dissection: a case-control study

The supplemental data include six figures and four tables.

Supplemental Data

Figure S1: Workflow of our study design and analysis procedure

Figure S2: Distribution of P values in permutation test

Figure S3: Quality control of whole exome sequencing

Figure S4: Homology analysis of COL3A1 across different species.

Figure S5: Dendrogram showing the sample cluster relationships

Figure S6: GO and KEGG pathway analysis in GSE52093

Table S1: The consistent of 1000 Genomes populations

Table S2: PCR primers and conditions for Sanger sequencing for *COL3A1* variants.

Table S3: Summary of the sequencing alignment and mapping coverage

Table S4: Candidate genes associated syndrome

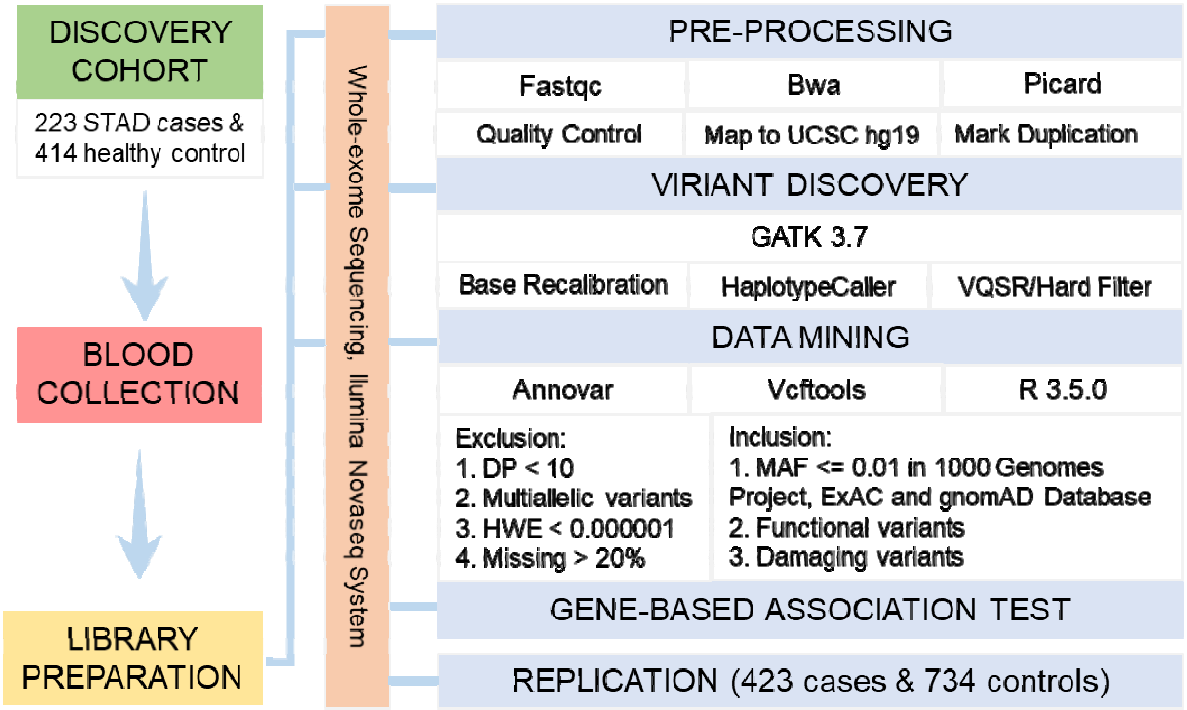

Figure S1: Workflow of our study design and analysis procedure. A total of 637 subjects were enrolled in our study (left). Exome sequencing procedure followed by GATK 3.7 best practice and variants control for gene-based associated test (right).

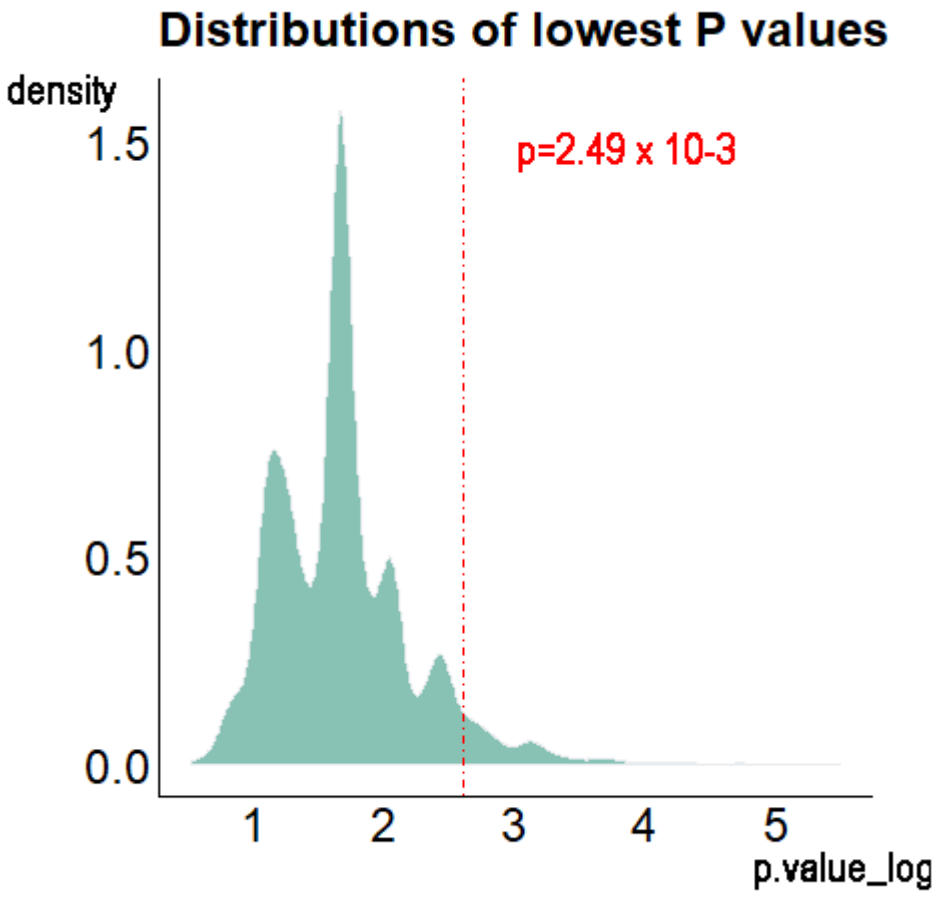

Figure S2: Distribution of P values in permutation test. Histograms show the distributions of the lowest P values across 10,000 permutations for the SKAT test. The red vertical line indicates the significance level of 0.05 for the most significant gene

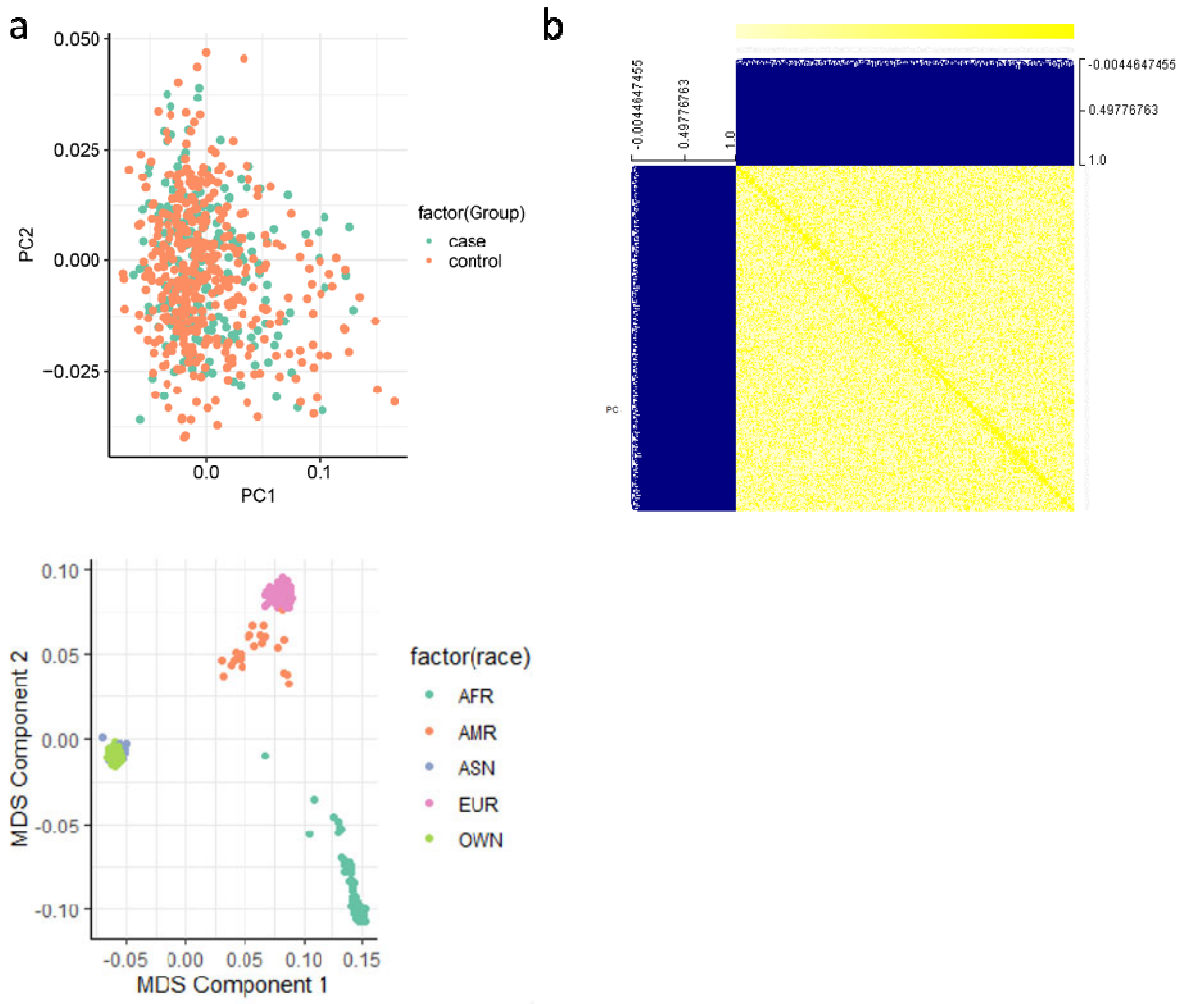

Figure S3: Quality control of whole exome sequencing. (a) Using the first three principal components to detect population stratification. Each point represented an individual, green and red dot represented case and control respectively. (b)Heatmap of the kinship matrix. Every dot presented the correlation of each pair. Color showed the degree of correlation. (c) MDS plot for ancestry analysis using 1000 Genomes populations. Different colors indicated various races.

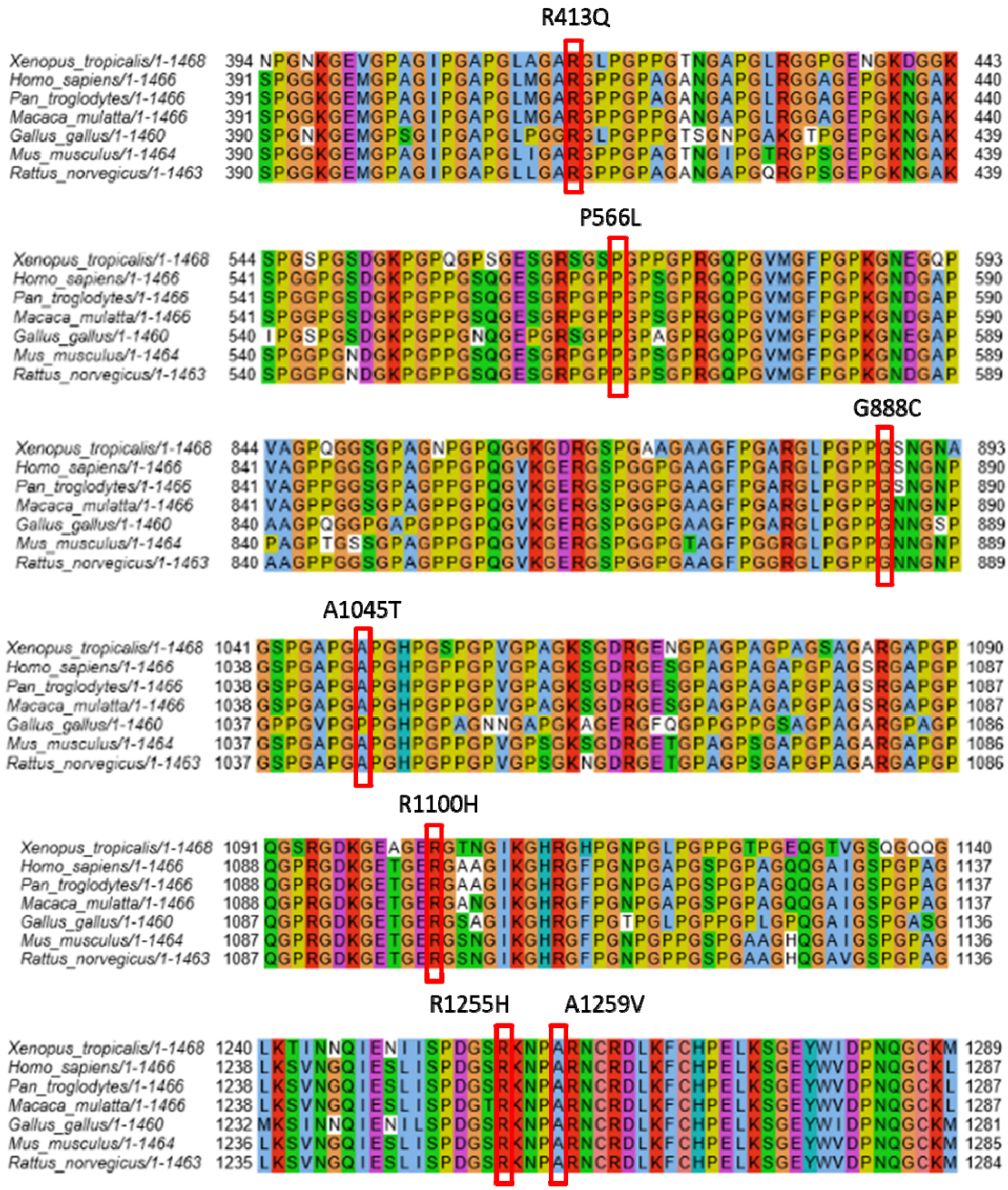

Figure S4: Homology analysis of COL3A1 rare deleterious variants in WES cohort across different species.

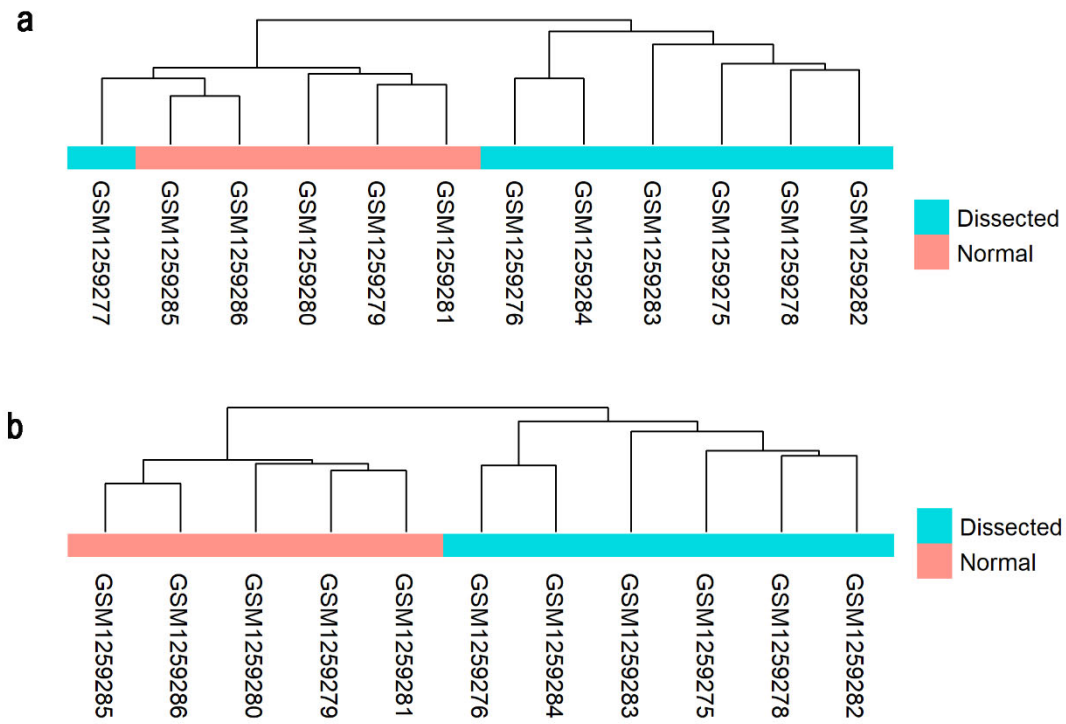

Figure S5: Dendrogram showing the sample cluster relationships (a) Dendrogram showing all sample clustering in GSE52093. Dissected refers dissected aortas from TAD patients, normal refers normal aortas from non-TAD patients. (b) Dendrogram showing sample clustering after excluding GSM1259277.

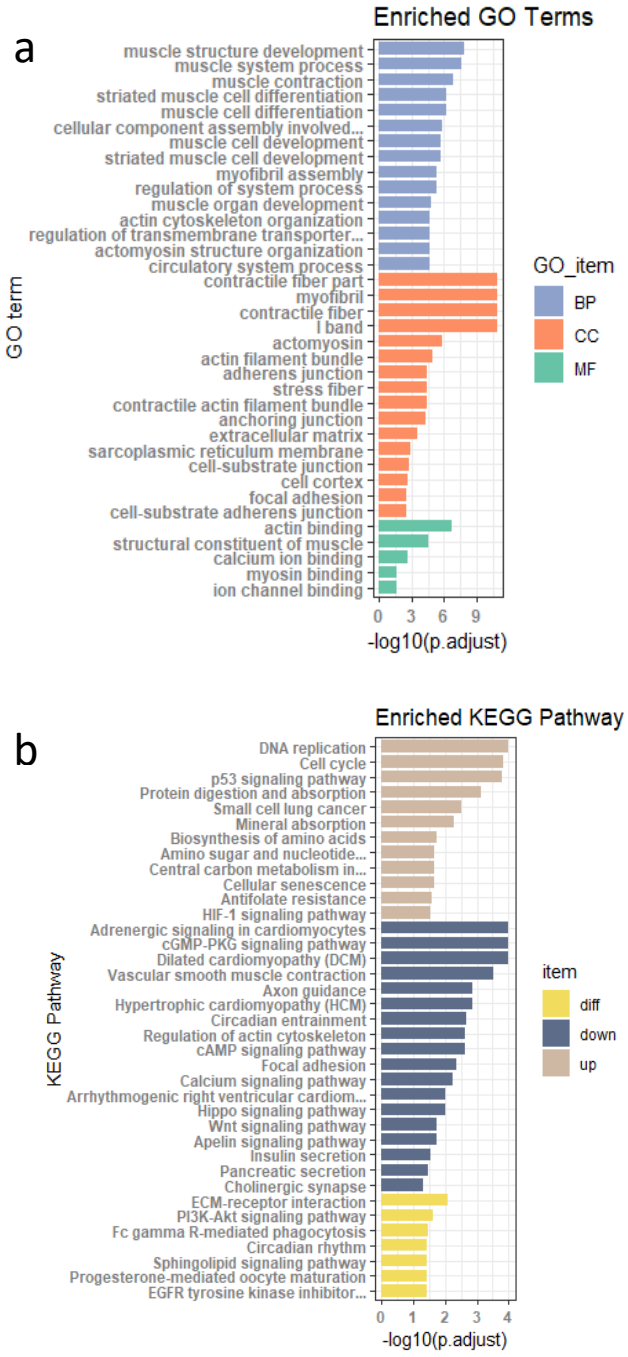

Figure S6: (a) bar plot showed the result of GO enrichment analysis using differentially expressed mRNAs (b) the result of KEGG pathway enrichment, we used differentially expressed mRNAs (diff), up-regulated mRNAs (up) in TAD

patients and down-regulated mRNAs (down) in TAD patients to perform KEGG pathway analysis.

Table S1: The consist of 1000 Genomes populations

region		1000 Genomes populations
African	ASW	HapMap African ancestry individuals from SW US
	YRI	Yoruba individuals
	LWK	Luhya individuals
American	MXL	HapMap Mexican individuals from LA California
	PUR	Puerto Rican in Puerto Rico
Asian	JPT	Japanese individuals
	CHB	Han Chinese in Beijing
	CHD	Chinese in metropolitan Denver
	CHS	Han Chinese South
European	CEU	CEPH individuals
	TSI	Toscan individuals
	GBR	British individuals from England and Scotland
	FIN	HapMap Finnish individual from Finland

Table S2: PCR primers and conditions for Sanger sequencing for *COL3A1* variants.

target		PCR primers(5'-3')	Product Size (bp)	PCR condition
R413Q	forward:	AGGCGAAATGGTAAGCTGTCC (sequencing primer)	491	
	reverse:	GCACCATTCTTACCAGGCTCA		
P566L	forward:	ATGTGCAAATCTGAGGCTTCAC	634	
	reverse:	CTCCCTTCCTACATGCAATTTG (sequencing primer)		
G666S	forward:	GAGAAAATGGAAAACCTGGGG (sequencing primer)	628	
	reverse:	GAATGTCAGAGAATTGGACCTTCC		
G888C	forward:	GTACAATGCAGATCATGCCACA (sequencing primer)	776	
	reverse:	TGAAGGCTACCAAAGGAAGGA		
R992H	forward:	TTCAGCCCTTTGCCATCC (sequencing primer)	445	
	reverse:	AGATAGAAGAACTTTGCTATGTGGC		
A1045T	forward:	GTCTGGACATTGTGACCCTTTG (sequencing primer)	720	
	reverse:	ACAGAAAAGTTGGGCAAGAGC		
R1100H R1100C	forward:	AAAGTTCTTTGGAGCCAAGGATAG (sequencing primer)	760	
	reverse:	TATCATGACATCACAATGCCAGAA		
R1255H A1259V	forward:	AGCAATGTATTCTTAGAGTGGCGA (sequencing primer)	569	
	reverse:	TGTATTTGGCATGATCTGAAGTTATG		
H1325N	forward:	TGCAGACACATTAGCAGTCAACA (sequencing primer)	605	
	reverse:	GAAGTTCAGGATTGCCGTAGC		

Table S3: Summary of the sequencing alignment and mapping coverage

	All(N=637)	Case(N=223)	Con(N=414)
mean coverage	110.14	100.63	115.24
coverage 1X %	97.97%	98.11%	97.89%
coverage 10X %	97.09%	97.17%	97.05%
coverage 20X %	95.58%	95.27%	95.75%
coverage 50X %	81.05%	77.05%	83.20%

Table S4 Candidate genes associated syndrome

Associated Syndrome	Gene Symble	Inheritance	GenBank	OMIM	Key Function
FTAAD	ACTA2	AD	NM_001141945	611788	SMC Contraction
	BGN	X-linked	NM_001711	300989	Extracellular Matrix
	FOXE3	AD	NM_012186	617349	SMC Contraction
	LOX	AD	NM_001178102	617168	Extracellular Matrix
	MFAP5	AD	NM_001297712	616166	Extracellular Matrix
	MYH11	AD	NM_002474	132900	SMC Contraction
	MYLK	AD	NM_001321309	613780	SMC Contraction
	PRKG1	AD	NM_001098512	615436	SMC Contraction
LDS	SMAD3	AD	NM_001145104	613795	TGF-β Pathway
	TGFB2	AD	NM_001135599	614816	TGF-β Pathway
	TGFB3	AD	NM_001329938	615582	TGF-β Pathway
	TGFBR1	AD	NM_001306210	609192	TGF-β Pathway
	TGFBR2	AD	NM_003242	610168	TGF-β Pathway
EDS	COL1A1	AD	NM_000088	130000	Extracellular Matrix
	COL1A2	AD	NM_000089	130060	Extracellular Matrix
	COL3A1	AD	NM_000090	130050	Extracellular Matrix
	PLOD1	AR	NM_000302	225400	Extracellular Matrix
BAV	NOTCH1	AD	NM_017617	109730	Neural Crest Migration
	SMAD6	AD	NM_005585	614823	TGF-β Pathway
CLS	EFEMP2	AR	NM_016938	614437	Extracellular Matrix
	ELN	AD, AR	NM_000501	123700	Extracellular Matrix

AS	COL4A5	X-linked	NM_000495	301050	Extracellular Matrix
ATS	SLC2A10	AR	NM_030777	208050	TGF-β Pathway
CCA	FBN2	AD	NM_001999	121050	Extracellular Matrix
CVD	FLNA	X-linked	NM_001456	300049	SMC Contraction
HANAC	COL4A1	AD	NM_001845	611773	Extracellular Matrix
JP/HHT	SMAD4	AD	NM_005359	175050	TGF-β Pathway
MFS	FBN1	AD	NM_000138	154700	Extracellular Matrix
SGS	SKI	AD	NM_003036	182212	TGF-β Pathway

AD, autosomal dominant. AOS, aneurysm osteoarthritis syndrome. AR, autosomal recessive. AS: Alport syndrome. ATS, arterial tortuosity syndrome. BAV: Bicuspid aortic valve. CCA: Congenital contractural arachnodactyly. CLS: Cutis laxa syndrome. CVD: Cardiac valvular dysplasia. EDS: Ehlers-Danlos syndrome. FTAAD: Familial Thoracic Aortic Aneurysm and Dissection. HANACS: Hereditary Angiopathy with Neuropathy, Aneurysms, and Muscle Cramps. JP/HHT: Juvenile polyposis/hereditary hemorrhagic telangiectasia syndrome. LDS: Loeys-Dietz syndrome. MFS: Marfan syndrome. OMIM, Online Mendelian Inheritance in Man. SGS: Shprintzen-Goldberg syndrome; SMC: smooth muscle cell; TGF-β: transforming growth factor beta