Supplementary Material

MYH9-related disease: description of a large Chinese pedigree and a survey of reported mutations

Shufang Zhang, Xueya Zhou, Shengnan Liu, Tingting Bai, Jing Wang, Xuegong Zhang & Binbin Wang

To compile a list of reported disease-associated mutations, we systematically reviewed the data from Human Gene Mutation Database (HGMD)(Stenson, et al., 2003). Related literature published before March 2013 was checked for functional data and co-segregation with phenotypes in pedigrees.

For each reported mutation, the cDNA position was mapped to the genomic coordinates based on the UCSC's RefSeq Gene annotation track. The reference alleles were checked to match the reference sequence. For small deletions within a stretch of repeated sequence, it was impossible to determine precisely which base(s) were deleted (as an example, both c.5818delG and c.5821delG were located in a stretch of G's); so we merged them to represent the same mutation, and place the deleted alleles at the leftmost of the repeated bases on the genomic coordinates. The missense SNVs were further annotated by four bioinformatics predictions on pathogenicity including SIFT, PolyPhen2, MutationTaster, and LRT queried from dbNSFP database (Liu, et al., 2011). 1000 Genomes Project (K1G) and NHLBI GO Exome Sequencing Project (ESP) were systematically searched for all exonic variants in the *MYH9* gene (assessed on March 15, 2013). If a disease-associated amino-acid change also appears in the public databases, its allele frequency based on the genotypes of 4300 Europeans in ESP will be shown. The results are summarized for SNVs and indels separately in the following two tables.

Several mutational hotspots accounting for the majority of reported mutations were well known (Kunishima and Saito, 2010). They included: six hotspots for point mutations (at residues S96, R702, R1165, D1424, E1841, and R1933), a hotspot for in-frame indels within exon 24(Miyazaki, et al., 2009b), and a hotspot for frame-shift indels in the last coding exon.

To summarize the patterns of all disease-associated mutations, we introduce following two definitions. We define an amino acid change as "*singleton*" mutation if the specific mutant allele was only reported in one family or a single patient; otherwise, it is referred to as "non-singleton" mutation. An amino acid mutation is referred to as "*recurrent*" if it met **either** one of the following criteria:

(1) Not a singleton mutation (i.e. reported in >=2 families or unrelated patients), which includes W33R, N93K, S96L, R702C, R702H, T115I, R1162T, R1162S, R1165C, D1424N, D1424H, D1424Y, V1516L, E1841K, R1933X, and Q1068_L1074del;
(2) Located within a position that mutated into >=2 different alleles found exclusively in patient (including mutations at hotspots and at residues W33, A95, T1155, R1162, D1447, V1516).

It should be noted all point mutations in the hotspots satisfy the definition of recurrence, so are the additional 16 mutations outside hotspots. The remaining mutations were "non-recurrent". They include ten SNVs and six indels shown in gray shadow in the tables and referred in the main text.

Following the above definitions, although a majority of point mutations (22) are singletons; many of them are recurrent. In total, 27 point mutations (including all 11 in hotspots) residing at 14 different residues are recurrent; and the remaining 10 are non-recurrent. For each singleton mutations, the original publication was checked and co-segregation patterns are indicated in the tables.

Table S1 All reported missense or nonsense single nucleotide variants (SNVs) associated with *MYH9*-RD (before March 2013). Mutational hotspots are shaded in light orange; non-recurrent mutations are shaded in light gray.

Exon	AA change[i]	cDNA change[i]	Genomic position (build 37)[ii]	nsSNV Predic-tions[iii]	Singleton[iv]/ AlleleFreq[v]	References[vi]
1	W33R	97T>A	22:36745185_A/T	+ + + +		Jang, et al. (2012)
		97T>C	22:36745185_A/G			Sun, et al. (2013)
1	W33C	99G>T	22:36745183_C/A	+ + + +	D	Kahr, et al. (2009)
1	V34G	101T>G	22:36745181_A/C	+ + + +	F(3)	De Rocco, et al. (2013)
1	N93K	279C>G	22:36745003_G/C	+ + + +		Kunishima, et al. (2003); Seri, et al. (2000)
1	A95T	283G>A	22:36744999_C/T	+ + + +	F	Kunishima, et al. (2001b)
1	A95D	284C>A	22:36744998_G/T	+ + + +	F(8)	De Rocco, et al. (2009a)
1	S96L	287C>A	22:36744995_G/A	+ + + +		Arrondel, et al. (2002); Dong, et al. (2005); Kunishima, et al. (2003)
10	K373N[vii]	1119G>C	22:36714360_C/G	+ + + +	F	Heath, et al. (2001)
16	R702C	2104C>T	22:36702031_G/A	+ + + +		Heath, et al. (2001); Sekine, et al. (2010); Seri, et al. (2000)
16	R702H	2105G>A	22:36702030_C/T	+ + + +		Dong, et al. (2005); Heath, et al. (2001); Seri, et al. (2002)
16	R702S	2104C>A	22:36702031_G/T	+ + + +	D	De Rocco, et al. (2013)
16	Q706E	2116C>G	22:36702019_G/C	+ + + +	S	Otsubo, et al. (2006)
16	R718W	2152C>T	22:36701983_G/A	+ + + +	D	Pecci, et al. (2008)
21	K910Q	2728A>C	22:36697007_T/G	+ + − +	F*(2)	Seri, et al. (2003)
25	S1114P	3340T>C	22:36691696_A/G	− + − −	S/1.2e-4	Heath, et al. (2001)
25	T1155A	3463A>G	22:36691573_T/C	+ + + +	D	Pecci, et al. (2008)
25	T1155I	3464C>T	22:36691572_G/A	+ + + +		Kelley, et al. (2000); Seri, et al. (2003)
25	R1162T	3485G>C	22:36691551_C/G	+ + + +		Savoia, et al. (2010a); Vettore, et al. (2010)
26	R1162S	3486G>T[viii]	22:36691122_C/A	+ + + +	S	Kunishima, et al. (2012)
26	R1165C	3493C>T	22:36691115_G/A	+ + + +		Dong, et al. (2005); Seri, et al. (2000)
26	R1165L	3494G>T	22:36691114_C/A	+ + + +	D	Kunishima, et al. (2001b); Pecci, et al. (2008)
30	R1400W	4198C>T	22:36688178_G/A	+ + + +	F*(2)/ 2.2e-3	Arrondel, et al. (2002)

30	D1424N	4207G>A	22:36688106_C/T	+ + + +		Dong, et al. (2005); Kunishima, et al. (2001a); Seri, et al. (2003)
30	D1424H	4207G>C	22:36688106_C/G	+ + + +		Pecci, et al. (2008); Seri, et al. (2000); Seri, et al. (2003)
30	D1424Y	4207G>T	22:36688106_C/A	+ + + +		Dong, et al. (2005); Kunishima, et al. (2001b)
30	D1447H	4339G>C	22:36688037_C/G	+ + − +	S	Burt, et al. (2008)
30	D1447Y	4339G>T	22:36688037_C/A	+ + − +	S	De Rocco, et al. (2013)
30	D1447G	4340A>G	22:36688036_T/C	+ + − +	S	Schleinitz, et al. (2006)
30	D1447V	4340A>T	22:36688036_T/A	+ + − +	F(4)	Pecci, et al. (2008)
31	V1516L	4546G>T	22:36685142_C/A	+ + + +		Ma, et al. (2006)
31		4546G>C	22:36685142_C/G	+ + + +		This study
31	V1516M	4546G>A	22:36685142_C/T	+ + + +	F(2)	Pecci, et al. (2010)
32	R1557L	4670G>C	22:36684873_C/A	+ + + +	F(2)/0[ix]	Pecci, et al. (2010)
34	M1651T	4952T>G	22:36682873_A/G	+ − − −	F(2)/ 1.4e-3	Provaznikova, et al. (2009)
37	I1816V	5446A>G	22:36681204_T/C	− − + −	F	Kunishima, et al. (2003)
38	E1841K	5521G>A	22:36680520_C/T	+ + + +		Kunishima, et al. (2001a); Pecci, et al. (2008); Seri, et al. (2000)
40	R1933X	5797C>T	22:36678800_G/A	n.a.		Kelley, et al. (2000); Kunishima, et al. (2001a); Seri, et al. (2000)
40	E1945X	5833G>T	22:36678764_C/A	n.a.		Seri, et al. (2003); Sun, et al. (2013)

[i] The amino acid positions are based on the translated protein sequence from NCBI RefSeq NM_002473.4. The DNA positions start with the translation initiation site.
[ii] The genomic position (NCBI Build 37) and DNA changes on the forward strand (chrom:pos_ref/alt allele).
[iii] Results from three non-synonymous SNV effect prediction algorithm, from left to right: PolyPhen2, SIFT, Mutation Taster. +: deleterious or damaging, −: benign.
[iv] For mutations only reported in one family or single patient (singletons), F: reported in one family, co-segregate with the phenotype, followed by the number of tested patients in the family if known; F*, reported in one family, but not co-segregate with the phenotype; D: reported in one sporadic patient, confirmed as de novo; S: reported in a single patients, no parental DNA available.
[v] Allele frequencies for the variants were calculated based on the genotypes of 4300 Europeans in ESP.
[vi] If reported more than three references, only the first three will be listed.
[vii] Originally reported as K371N, which was likely a typo.
[viii] This missense variant was reported to cause exon 26 skipping; but our bioinformatics prediction does not support the splicing effect.
[ix] R1557Q was observed once in ESP exomes caused by C>T change in the same nucleotide position.

Table S2 All reported insertion/deletions/block substitutions associated with *MYH9*-RD. Mutational hotspots are shaded in light orange; non-recurrent mutations are shaded in light gray.

Exon	AA Change	cDNA Change	Genomic Position (Build 37)	Singleton[x]	References
1	W33_P35delinsCVA	99_103delinsTGTGG[xi]	22:36745179_ 36745183delins CCACA	D	Miyajima and Kunishima (2009a)
1	N76_S81del	228_245del	22:36745037_ 36745054del	S	Miyazaki, et al. (2009b)
20	M847_E853dup	2539_2559dup	22:36697652_ 36697672dup	D	De Rocco, et al. (2013)
24	K1048_E1054del	3142_3162del	22:36692999_ 36693019del	D	De Rocco, et al. (2013)
24	G1055_Q1068del	3164_3205del	22:36692956_ 36692997del	F	Miyazaki, et al. (2009b)
24	Q1068_L1074dup	3195_3215dup	22:36692946_ 36692966dup	F(3)	De Rocco, et al. (2009b)
24	Q1068_L1074del	3195_3215del	22:36692946_ 36692966del		Miyazaki, et al. (2009b); Seri, et al. (2003)
24	E1084del	3250_3252del	22:36692909_ 36692911del	F	Miyazaki, et al. (2009b)
25	V1092_R1162del	n.a.[xii]	22:36691354_ 36692573del	S	Kunishima, et al. (2008b)
26	L1205_Q1207del	3613_3621del	22:36690987_ 36690995del	F	Kunishima, et al. (2001b)
30	Q1443_K1445dup	4327_4335dup	22:36688041_ 36688049dup	S	Sun, et al. (2013)
40	D1925TfsX23	5773del	22:36678824del		Kunishima, et al. (2008a); Pecci, et al. (2008)
40	D1925AfsX23	5774del	22:36678823del	F	Kunishima, et al. (2001b)
40	G1924RfsX21	5770_5779del	22:36678818_ 36678827del	F(3)	Pecci, et al. (2010)
40	P1927RfsX21	5780del	22:36678817del	F	Kunishima, et al. (2001a)
40	V1930CfsX18	5788del	22:36678809del	S	Pecci, et al. (2010)
40	R1933EfsX15	5797del	22:36678800del	F(2)	Pecci, et al. (2008)
40	M1934WfsX14	5800del	22:36678797del	F(3)	Savoia, et al. (2010b)
40	D1941MfsX7	5821del[xiii]	22:36678776del		Kunishima, et al. (2003); Pecci, et al. (2008)

[x] F: reported in one family, co-segregate with the phenotype, followed by the number of tested patients in the family if reported; D: reported in one sporadic patient, confirmed as de novo; S: reported in a single patients, no parental DNA available.

[xi] This mutations was initially reported as two point de novo mutations (c.99G>T and c.103C>G) in *cis*. However, we found it could be most parsimoniously explained as one single block substitution.

[xii] This was a whole exon deletion whose break point boundaries were located in flanking introns.

[xiii] The deletion was first reported as 5828del in Kunishima S et al. (2001b) later corrected in Kunishima S et al. (2003).

References

Arrondel, C., *et al.* (2002) Expression of the nonmuscle myosin heavy chain IIA in the human kidney and screening for MYH9 mutations in Epstein and Fechtner syndromes, *Journal of the American Society of Nephrology : JASN*, **13**, 65-74.

Burt, R.A., *et al.* (2008) Description of a novel mutation leading to MYH9-related disease, *Thrombosis research*, **122**, 861-863.

De Rocco, D., *et al.* (2009a) MYH9 related disease: a novel missense Ala95Asp mutation of the MYH9 gene, *Platelets*, **20**, 598-602.

De Rocco, D., *et al.* (2009b) Identification of the first duplication in MYH9-related disease: a hot spot for unequal crossing-over within exon 24 of the MYH9 gene, *European journal of medical genetics*, **52**, 191-194.

De Rocco, D., *et al.* (2013) MYH9-related disease: five novel mutations expanding the spectrum of causative mutations and confirming genotype/phenotype correlations, *European journal of medical genetics*, **56**, 7-12.

Dong, F., *et al.* (2005) Genotype-phenotype correlation in MYH9-related thrombocytopenia, *British journal of haematology*, **130**, 620-627.

Heath, K.E., *et al.* (2001) Nonmuscle myosin heavy chain IIA mutations define a spectrum of autosomal dominant macrothrombocytopenias: May-Hegglin anomaly and Fechtner, Sebastian, Epstein, and Alport-like syndromes, *American journal of human genetics*, **69**, 1033-1045.

Jang, M.J., *et al.* (2012) A Trp33Arg mutation at exon 1 of the MYH9 gene in a Korean patient with May-Hegglin anomaly, *Yonsei medical journal*, **53**, 662-666.

Kahr, W.H., *et al.* (2009) Megakaryocyte and platelet abnormalities in a patient with a W33C mutation in the conserved SH3-like domain of myosin heavy chain IIA, *Thrombosis and haemostasis*, **102**, 1241-1250.

Kelley, M.J., *et al.* (2000) Mutation of MYH9, encoding non-muscle myosin heavy chain A, in May-Hegglin anomaly, *Nature genetics*, **26**, 106-108.

Kunishima, S., Hamaguchi, M. and Saito, H. (2008a) Differential expression of wild-type and mutant NMMHC-IIA polypeptides in blood cells suggests cell-specific regulation mechanisms in MYH9 disorders, *Blood*, **111**, 3015-3023.

Kunishima, S., *et al.* (2001a) Mutations in the NMMHC-A gene cause autosomal dominant macrothrombocytopenia with leukocyte inclusions (May-Hegglin anomaly/Sebastian syndrome), *Blood*, **97**, 1147-1149.

Kunishima, S., *et al.* (2008b) Identification and characterization of the first large deletion of the MYH9 gene associated with MYH9 disorders, *European journal of haematology*, **80**, 540-544.

Kunishima, S., *et al.* (2001b) Identification of six novel MYH9 mutations and genotype-phenotype relationships in autosomal dominant macrothrombocytopenia with leukocyte inclusions, *Journal of human genetics*, **46**, 722-729.

Kunishima, S., *et al.* (2003) Immunofluorescence analysis of neutrophil nonmuscle myosin heavy chain-A in MYH9 disorders: association of subcellular localization with MYH9 mutations, *Laboratory investigation; a journal of technical methods and pathology*, **83**, 115-122.

Kunishima, S. and Saito, H. (2010) Advances in the understanding of MYH9 disorders, *Current opinion in hematology*, **17**, 405-410.

Kunishima, S., *et al.* (2012) G to T transversion at the first nucleotide of exon 26 of the MYH9 gene results in a novel missense mutation and abnormal splicing in platelets: comment on "A G to C transversion at the last nucleotide of exon 25 of the MYH9 gene results in a missense mutation rather than in a splicing defecton" by Vettore et al, *European journal of medical genetics*, **55**, 763-765.

Liu, X., Jian, X. and Boerwinkle, E. (2011) dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions, *Human mutation*, **32**, 894-899.

Ma, E.S., *et al.* (2006) Hematologic and genetic characterization of an MYH9-related disorder in a Chinese family, *Haematologica*, **91**, 1002-1003.

Miyajima, Y. and Kunishima, S. (2009a) Identification of the first in cis mutations in MYH9 disorder, *European journal of haematology*, **82**, 288-291.

Miyazaki, K., *et al.* (2009b) Identification of three in-frame deletion mutations in MYH9 disorders suggesting an important hot spot for small rearrangements in MYH9 exon 24, *European journal of haematology*, **83**, 230-234.

Otsubo, K., *et al.* (2006) Identification of a novel MYH9 mutation in a patient with May-Hegglin anomaly, *Pediatric blood & cancer*, **47**, 968-969.

Pecci, A., *et al.* (2010) MYH9 related disease: four novel mutations of the tail domain of myosin-9 correlating with a mild clinical phenotype, *European journal of haematology*, **84**, 291-297.
Pecci, A., *et al.* (2008) Position of nonmuscle myosin heavy chain IIA (NMMHC-IIA) mutations predicts the natural history of MYH9-related disease, *Human mutation*, **29**, 409-417.
Provaznikova, D., *et al.* (2009) Clinical manifestation and molecular genetic characterization of MYH9 disorders, *Platelets*, **20**, 289-296.
Savoia, A., *et al.* (2010a) Heavy chain myosin 9-related disease (MYH9 -RD): neutrophil inclusions of myosin-9 as a pathognomonic sign of the disorder, *Thrombosis and haemostasis*, **103**, 826-832.
Savoia, A., *et al.* (2010b) MYH9-related disease: Report on five German families and description of a novel mutation, *Annals of hematology*, **89**, 1057-1059.
Schleinitz, N., *et al.* (2006) [The MYH9 syndrome: report of a new case with a new mutation of the MYH9 gene], *La Revue de medecine interne / fondee ... par la Societe nationale francaise de medecine interne*, **27**, 783-786.
Sekine, T., *et al.* (2010) Patients with Epstein-Fechtner syndromes owing to MYH9 R702 mutations develop progressive proteinuric renal disease, *Kidney international*, **78**, 207-214.
Seri, M., *et al.* (2000) Mutations in MYH9 result in the May-Hegglin anomaly, and Fechtner and Sebastian syndromes. The May-Heggllin/Fechtner Syndrome Consortium, *Nature genetics*, **26**, 103-105.
Seri, M., *et al.* (2003) MYH9-related disease: May-Hegglin anomaly, Sebastian syndrome, Fechtner syndrome, and Epstein syndrome are not distinct entities but represent a variable expression of a single illness, *Medicine*, **82**, 203-215.
Seri, M., *et al.* (2002) Epstein syndrome: another renal disorder with mutations in the nonmuscle myosin heavy chain 9 gene, *Human genetics*, **110**, 182-186.
Stenson, P.D., *et al.* (2003) Human Gene Mutation Database (HGMD): 2003 update, *Human mutation*, **21**, 577-581.
Sun, X.H., *et al.* (2013) Clinical, pathological, and genetic analysis of ten patients with MYH9-related disease, *Acta haematologica*, **129**, 106-113.
Vettore, S., *et al.* (2010) A G to C transversion at the last nucleotide of exon 25 of the MYH9 gene results in a missense mutation rather than in a splicing defect, *European journal of medical genetics*, **53**, 256-260.