Journal of Cell Communication and Signaling (2021) 15:93–105
https://doi.org/10.1007/s12079-020-00598-9

RESEARCH ARTICLE

Differential intolerance to loss of function and missense mutations in genes that encode human matricellular proteins

Sukhbir Kaur[1] · **David D. Roberts**[1]

Received: 7 October 2020 / Accepted: 24 November 2020 / Published online: 7 January 2021

Abstract

Targeted gene disruption in mice has provided valuable insights into the functions of matricellular proteins. Apart from missense and loss of function mutations that have been associated with inherited diseases, however, their functions in humans remain unclear. The availability of deep exome sequencing data from over 140,000 individuals in the Genome Aggregation Database provided an opportunity to examine intolerance to loss of function and missense mutations in human matricellular genes. The probability of loss-of-function intolerance (pLI) differed widely within members of the thrombospondin, CYR61/CTGF/NOV (CCN), tenascin, small integrin-binding ligand N-linked glycoproteins (SIBLING), and secreted protein, acidic and rich in cysteine (SPARC) gene families. Notably, pLI values in humans had limited correlation with viability of the corresponding homozygous null mice. Among the thrombospondins, only *THBS1* was highly loss-intolerant (pLI = 1). In contrast, *Thbs1* is not essential for viability in mice. Several known thrombospondin-1 receptors were similarly loss-intolerant, although thrombospondin-1 is not the exclusive ligand for some of these receptors. The frequencies of missense mutations in *THBS1* and the gene encoding its signaling receptor CD47 indicated conservation of some residues implicated in specific receptor binding. Deficits in missense mutations were also observed for other thrombospondin genes and for *SPARC*, *SPOCK1*, *SPOCK2*, *TNR*, and *DSPP*. The intolerance of *THBS1* to loss of function in humans and elevated pLI values for *THBS2*, *SPARC*, *SPOCK1*, *TNR*, and *CCN1* support important functions for these matricellular protein genes in humans, some of which may relate to functions in reproduction or responding to environmental stresses.

Keywords Human genetic variation · Population genetics · Loss of function variants · Matricellular proteins · Gene families

Abbreviations

α2δ1	Precursor of the α2 and δ subunits of voltage-dependent calcium channel, encoded by *CACN2D1*
CCN	Cyr61/CTGF/NOV gene family
CD148	Membrane-bound tyrosine phosphatase, encoded by *PTPRJ*
COMP	Cartilage oligomeric matrix protein
DMP1	Dentin matrix protein 1
DSPP	Dentin sialophosphoglycoprotein
ExAC	Exome Aggregation Consortium
gnomAD	Genome Aggregation Database
IBSP	Bone sialoprotein
LoF	Loss of function
LRP1	Low density lipoprotein receptor-related protein 1
MEPE	Matrix extracellular phosphoglycoprotein
pLI	Probability of loss-of-function intolerance
SIBLING	Small integrin-binding ligand N-linked glycoproteins
SIRPα	Signal regulatory protein-α
SMOC	Secreted modular calcium-binding protein
SNV	Single nucleotide variation
SPARC	Secreted protein, acidic and rich in cysteine
SPOCK	Sparc/osteonectin, CWCV, and kazal-like domains proteoglycan (Testican)
SPP1	Osteopontin
STIM1	Stromal interaction molecule 1
THBS	Thrombospondin
TN	Tenascin

✉ Sukhbir Kaur
sukhbir.kaur@nih.gov

✉ David D. Roberts
droberts@mail.nih.gov

[1] Laboratory of Pathology, Center for Cancer Research, National Cancer Institute, NIH, Building 10 Room 2S235, 10 Center Drive MSC1500, Bethesda, MD 20892-1500, USA

Introduction

Matricellular proteins were defined by Paul Bornstein in 1995 as extracellular regulators of cell function that modulate cell behavior by interacting with structural components of the extracellular matrix, cytokines, or proteases and by binding to specific cell surface receptors (Bornstein 1995; Murphy-Ullrich and Sage 2014). With some exceptions, matricellular proteins do not serve structural roles in the extracellular matrix. Rather, they are transiently expressed at specific stages of development, during tissue remodeling, or in response to acute injuries or chronic disease. Original members included thrombospondin-1 and -2, tenascins, the secreted protein acidic and rich in cysteine (SPARC) family, and osteopontin, a member of the small integrin-binding ligand N-linked glycoproteins (SIBLING) family. Matricellular proteins currently include additional families including the Cyr61/CTGF/NOV (CCN) gene family, short fibulins, galectins, and R-spondins (Elola et al. 2007; Knight and Hankenson 2014; Leask 2020; Murphy-Ullrich and Sage 2014; Nakamura 2018).

Human genetics and transgenic mice have provided complementary insights into the functions of matricellular proteins. Missense, regulatory, or inactivating mutations in genes encoding specific matricellular proteins including *CCN6*, *SPARC*, *SMOC1*, *SPOCK1*, *TNXB*, *DMP1*, *DSPP*, *THBS1*, *THBS2*, and *COMP* have been linked to inherited genetic disorders or disease risk in humans (Abouzeid et al. 2011; Bristow et al. 2005; Burke et al. 2009; Dhamija et al. 2014; Hurvitz et al. 1999; Mendoza-Londono et al. 2015; Okada et al. 2011; Posey et al. 2018; Rainger et al. 2011; Staines et al. 2012; Stenina et al. 2007; Topol et al. 2001). However, genes that serve critical roles during human fetal development may escape detection. Conversely, disruption of matricellular genes in mice by homologous recombination identified *CCN1*, *CCN2*, and *SMOC1* to be essential for viability (Ivkovic et al. 2003; Mo and Lau 2006; Mo et al. 2002; Okada et al. 2011), but other homozygous null mice were viable, and some initially lacked an obvious phenotype (Bouleftour et al. 2016; Bradshaw 2009; Canalis et al. 2010; Hankenson et al. 2005a, b; Jones and Jones 2000; Kutz et al. 2005; Midwood and Orend 2009; Svensson et al. 2002). In some cases, important gene functions have been revealed when these mice were subjected to specific stresses (Calabro et al. 2014; Kim et al. 2018; Murphy-Ullrich and Sage 2014; Roberts et al. 2012; Soto-Pantoja et al. 2015; Stenina-Adognravi and Plow 2019).

Disruption of *Thbs1,* encoding thrombospondin-1 in mice, yielded viable mice that were fertile and appeared healthy except for lung inflammation (Lawler et al. 1998). The lung inflammation may relate to exposure to a specific pathogen because the lung phenotype was lost when the mice were rederived in a different vivarium (Isenberg et al. 2008a). Subsequent studies identified beneficial as well as detrimental effects of *Thbs1* gene disruption on the ability of mice to survive exposure to specific pathogens or respond to a variety of physiological stresses (Arun et al. 2020; Martin-Manso et al. 2012; McMaken et al. 2011; Qu et al. 2018; Soto-Pantoja et al. 2015; Zhao et al. 2015). These studies illustrate how gene functions can be influenced by the environmental context. Such environmental stresses that could reveal important adaptive functions of matricellular protein genes may be absent in the highly controlled environment of a laboratory vivarium.

Despite our ability to control our environment, the ability to survive numerous environmental stresses including acute injuries and ongoing exposures to endemic and novel pathogens has played an important role in human evolution. Genetic diversity is critical for the long-term survival of any species facing such unpredictable challenges, and identifying relevant variations in specific genes is one goal of population genetics. The Exome Aggregation Consortium (ExAC) assembled a data set containing variant calls across 60,706 human exomes to globally examine the prevalence of missense and predicted loss of function (LoF) mutations (Lek et al. 2016). Known and previously unrecognized essential genes were identified by having significantly fewer LoF mutations than expected. This genomic variation data was expanded to include 141,456 individuals in the Genome Aggregation Database (gnomAD), which currently includes 125,748 deep-sequenced exomes and 15,701 full genome sequences from unrelated individuals (Karczewski et al. 2020). Individuals with severe pediatric genetic diseases and their first-degree relatives were excluded to better reflect the incidence of recessive disease-causing alleles. Here we analyzed gnomAD v2.1.1 data to examine the rates of missense and predicted LoF mutations in several families of matricellular protein genes. Focusing on gene families that include several paralogs provided useful controls because the expected frequencies of LoF mutants depends in part on the length of their coding regions (Lek et al. 2016).

Materials and methods

Data for missense and LoF mutants in genes that encode human matricellular proteins was accessed and analyzed using the ExAC browser (http://exac.broadinstitute.org) and, subsequently, the Genome Aggregation Database (gnomAD, https://gnomad.broadinstitute.org) (Karczewski et al. 2020; Lek et al. 2016). Mouse knockout phenotypes were obtained from mousephenotype.org or http://www.informatics.jax.org/ (Bult et al. 2019) and published studies where indicated.

Data for clinical associations with variants in human genes was obtained from ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/).

Results

Variation in the frequencies of LoF mutations in the thrombospondin gene family

Previous studies demonstrated that mice with homozygous LoF mutants in any single *THBS* family gene are viable (Frolova et al. 2010; Hankenson et al. 2005a, b; Lawler et al. 1998; Svensson et al. 2002). Characterization of strains bearing multiple *THBS* gene knockouts indicated minimal functional cross-compensation between the five *THBS* genes (Posey et al. 2008). In contrast, the ExAC data from 60,706 individuals and the expanded human gnomAD dataset representing 141,456 individuals showed deficits in the observed versus expected numbers of individuals with LoF mutants for *THBS1* and *THBS2* (Table 1). The pLI for *THBS1* was 1.00 in both datasets, indicating this gene to be highly loss intolerant. None of the individuals with *THBS1* LoF mutant alleles were homozygotes. *THBS2* also showed substantial loss intolerance, whereas the numbers of observed LoF mutants in *THBS3*, *THBS4*, and *COMP* did not differ significantly from the expected numbers. None of the *THBS2*, *THBS3*, or *COMP* LoF mutants were homozygous, whereas the *THBS4* mutants in two individuals were homozygous frameshifts (p.Thr915GlnfsTer and p.Thr915GlnfsTer).

LoF mutations in the *SPARC* gene family

Among the *Sparc* gene family members in mice, only *Smoc1* is essential for viability based on studies of homozygous null mice (Okada et al. 2011). Homozygous nonsense and splice mutants in murine *Smoc1* recapitulated the features of the *SMOC1*-dependent human autosomal-recessive disorder microphthalmia with limb anomalies (Waardenburg Anophthalmia syndrome), and mice bearing these mutations died shortly after birth. Despite being essential for normal ocular and limb development in mice and humans (Okada et al. 2011; Rainger et al. 2011), *SMOC1* did not show a significant pLI in humans (Table 2). One caveat in interpreting this result is that the number of expected LoF mutants for *SPARC* family genes is smaller than for the *THBS* family because the

Table 1 LoF mutants in the thrombospondin gene family

Gene	ORF (kb)	Expected LoF mutants	Observed LoF mutants	Observed/expected (90% range)	pLI	Null mouse phenotype
THBS1	3.5	56	7	0.13 (0.07–0.23)	1.00	viable
THBS2	3.5	59.5	13	0.22 (0.14–0.35)	0.56	viable
THBS3	2.9	57.4	41	0.71 (0.56–0.93)	0.0	viable
THBS4	2.9	51.6	36	0.7 (0.53–0.92)	0.0	viable
COMP	2.3	39.5	21	0.53 (0.38–0.77)	0.0	viable

Observed numbers of LoF mutants identified based on exome sequencing from 141,456 individuals in the Genome Aggregation Database (v2.1.1) were used to calculate the probability the indicated genes are loss-intolerant (pLI) as described (Karczewski et al. 2020; Lek et al. 2016). Where multiple isoforms exist, ORF length is presented for isoform 1 except where noted. Null mouse viability phenotypes are from http://www.informatics.jax.org/ (Bult et al. 2019).

Table 2 LoF mutations in the *SPARC* gene family

Gene	ORF (kb)	Expected LoF mutants	Observed LoF mutants	Observed/expected (90% range)	pLI	Null mouse phenotype
SPARC	0.9	16	2	0.12 (0.05–0.39)	0.89	Viable
SPARCL1	2.0	28.8	19	0.66 (0.46–0.97)	0.0	Viable
SPOCK1	1.3	23.4	4	0.17 (0.08–0.39)	0.83	Viable
SPOCK2	1.3[b]	24.9	6	0.24 (0.13–0.47)	0.24	Pending[a]
SPOCK3	1.3[b]	25.4	8	0.32 (0.18–0.57)	0.01	Viable
SMOC1	1.3	24.4	8	0.33 (0.19–0.59)	0.01	Neonatal lethal
SMOC2	1.4	25.3	9	0.36 (0.21–0.62)	0.0	Viable

[a]Mouse registered but phenotype not currently available at www.mousephenotype.org

[b]Isoform 2

former have much shorter coding regions. Because the 90% confidence range for all of the *SPOCK* and *SMOC* paralogs extends to significant observed/expected LoF ratios, future availability of exome data for a larger human population will be required to confirm or exclude significant pLI values for any of these genes.

Despite being nonessential for development in mice (Gilmour et al. 1998; Roll et al. 2006), *SPARC* and *SPOCK1* had high pLI values in the human data (Table 2). Humans lacking *SPARC* have not been reported to date, but missense mutations in *SPARC* cause osteogenesis imperfecta, type XVII (Mendoza-Londono et al. 2015). Most studies of SPARC in humans have focused on its role in various cancers and their metastatic spread (Nagaraju et al. 2014). Abnormalities in *Sparc* null mice include cataract formation and rupture of the lens capsule in the eye, severe osteopenia, and accelerated closure of dermal wounds (Bradshaw 2009; Gilmour et al. 1998).

Humans lacking *SPOCK1* have not been reported to date, but a missense mutation in *SPOCK1* (p.D80V) was identified in a patient with developmental delay, agenesis of the corpus callosum, and microcephaly (Dhamija et al. 2014). *SPOCK1* encodes the proteoglycan testican-1, which is involved in neurogenesis and epithelial-to-mesenchymal transition (Roll et al. 2006; Sun et al. 2020). The lack of obvious abnormalities in *Spock1* null mice may be due to functional redundancies that were reported with testican-2 (*Spock2*) and testican-3 (*Spock3*) (Roll et al. 2006).

LoF mutations in the *CCN* gene family

CCN1 was the only member of the CCN gene family with an elevated pLI (Table 3). This is consistent with the embryonic and perinatal lethal phenotype of a homozygous *Ccn1* LoF mutant in mice (Mo and Lau 2006; Mo et al. 2002). As with the SPARC family, the relatively short coding sequences of *CCN* genes may require exome data from a larger population to reliably confirm or exclude intolerance to LoF, especially for *CCN3* and *CCN6* where the 90% range for observed/expected LoF mutants extends below the standard 0.3 pLI threshold. Loss of *Ccn2* limits viability in mice based on its roles in bone and lung development (Baguma-Nibasheka and Kablar 2008; Ivkovic et al. 2003; Kawaki et al. 2008). None of the coding variations for *CCN2* currently in ClinVar have a known clinical relevance, but analysis of additional exomes may determine whether human functions of CCN2 parallel those identified in the null mice. Three siblings diagnosed with early-onset parkinsonism were homozygous for a p.D82G mutation in *CCN3* (Bentley et al. 2020), which suggests a pathophysiological function that would not be detected by this LoF screen. The same mutation occurred once as a heterozygous variant in the gnomAD dataset (Karczewski et al. 2020; Lek et al. 2016).

Previous studies indicated that the phenotypes associated with LoF mutants in *CCN6* diverge between humans and mice (Kutz et al. 2005). Inactivating mutations in human *CCN6* cause an autosomal recessive skeletal disorder, progressive pseudorheumatoid dysplasia (Hurvitz et al. 1999), whereas comparable mutants of *Ccn6* in mice had no skeletal phenotype (Kutz et al. 2005). In humans with pseudorheumatoid dysplasia, a nonsense variant of *CCN6* was always in cis with a G83E missense allele (Supplemental Table 20 in (Lek et al. 2016)).

LoF mutations in the tenascin gene family

TNR was the only member of the tenascin gene family with an elevated pLI (Table 4). Because *Tnr* null mice are viable, and the associated null phenotypes involve altered cognitive functions (Weber et al. 1999), the rationale for a deficit in LoF mutants in humans is unclear. Tenascin-R is primarily expressed in the central nervous system, and homozygous deletion of *TNR* was found in a patient with intellectual disability (Dufresne et al. 2012). The SNV rs6686722 in *TNR* was associated with attention deficit hyperactivity disorder in a hypothesis-free genome-wide association study (Hawi et al. 2018).

LoF mutations in the SIBLING gene family

None of the SIBLING family genes are essential for viability in mice, and reported disease-associated mutations in

Table 3 LoF mutations in the *CCN* gene family

Gene	ORF (kb)	Expected LoF mutants	Observed LoF mutants	Observed/expected (90% range)	pLI	Null mouse phenotype
CCN1 (Cyr61)	1.1	17.2	3	0.17 (0.08–0.45)	0.71	Embryonic or perinatal lethal
CCN2 (CTGF)	1.0	12.4	7	0.56 (0.32–1.06)	0.0	Perinatal lethal
CCN3 (NOV)	1.1	14.7	6	0.41 (0.22–0.81)	0.01	Viable
CCN4 (WISP1)	1.1	17.3	12	0.69 (0.44–1.13)	0.0	Viable
CCN5 (WISP2)	0.8	9.8	8	0.82 (0.48–1.46)	0.0	Viable
CCN6 (WISP3)	1.2	17.1	8	0.47 (0.27–0.84)	0.0	Viable

Table 4 LoF mutations in tenascins

Gene	ORF (kb)	Expected LoF mutants	Observed LoF mutants	Observed/expected (90% range)	pLI	Null mouse phenotype
TNC	6.6	91.4	30	0.33 (0.24–0.45)	0.00	Viable
TNXB	2.0[b]	21.8	7	0.32 (0.18–0.6)	0.02	Viable
TNN	3.9	59.6	49	0.82 (0.65–1.04)	0.00	Pending*
TNR	4.1	72.4	16	0.22 (0.15–0.34)	0.52	Viable

*Mouse registered but phenotype not currently available at www.mousephenotype.org

[b]Isoform 2

humans generally have postnatal effects (Bouleftour et al. 2016; Staines et al. 2012). Correspondingly, elevated pLI values were not found for these genes in the gnomAD data (Table 5). Mutations in *DMP1* cause autosomal recessive hypophosphatemic rickets, type 1 and osteomalacia (Feng et al. 2006; Lorenz-Depiereux et al. 2006). Mutations in *DSPP* are associated with dentinogenesis imperfecta and dentin dysplasia (Song et al. 2008; Zhang et al. 2001).

Distribution and frequencies of missense mutations in *THBS* family genes

In the ExAC data, deficits in LoF mutants positively correlated with deficits in missense SNVs (Lek et al. 2016). Consistent with this global correlation and the pLI data in Table 1, *THBS1* had the highest Z-score in the thrombospondin gene family for observed/expected missense mutants in the gnomAD data, and *THBS2* was the second highest (Table 6). Notably, all members of the *THBS* family had deficits in observed versus expected missense SNVs that exceeded the 90% range.

The frequency of missense variants across the *THBS1* coding sequence is presented in Fig. 1. One of the most frequent variants is at N700. The N700S variant was associated with increased risk for early myocardial infarction (Topol et al. 2001), and biochemical studies established that this variant decreases the affinity for calcium binding (Stenina et al. 2005) and destabilizes the protein (Carlson et al. 2008). In addition to altering the secretion or stability of thrombospondin-1, missense mutations could interfere with its interactions with other proteins that mediate its functions, including multiple cell surface receptors (Resovi et al. 2014). Except for more frequent variations at G454 and R517 in sequences identified to be recognized by CD36 (Dawson et al. 1997; Tolsma et al. 1993), all variations in sequences previously implicated in interactions of thrombospondin-1 with its integrin and non-integrin receptors (Calzada and Roberts 2005) occurred at frequencies $< 10^{-4}$. Furthermore, the frequencies of rare variants in these sequences were similar to those for SNVs occurring throughout the coding sequence. Therefore, the frequency of missense SNVs in these putative functional sequences is insufficient to infer protection of specific receptor binding sites.

Missense SNV frequencies in other matricellular gene families

The loss-intolerant SPARC family members *SPARC* and *SPOCK1* also had elevated Z-scores for deficits in missense SNVs (Table 6). As noted previously, missense mutations in *SPARC* (E263K, R166H) cause osteogenesis imperfecta, type XVII (Mendoza-Londono et al. 2015), and a p.D80V missense mutation in *SPOCK1* was associated with developmental delay, agenesis of the corpus callosum, and microcephaly (Dhamija et al. 2014). Only one SPARC p.Arg166His allele (frequency 3.98×10^{-6}) was found in gnomAD. SPARC p.E263K and SPOCK1 p.D80V variants were not found in the gnomAD data. Consistent with *Smoc1*

Table 5 LoF mutations in SIBLING gene family members

Gene	ORF (kb)	Expected LoF mutants	Observed LoF mutants	Observed/expected (90% range)	pLI	Null mouse phenotype
SPP1	0.9	9.7	9	0.93 (0.56–1.59)	0.0	Viable
DMP1	1.5	18.1	13	0.72 (0.47–1.14)	0.0	Viable
DSPP	3.9	17.6	13	0.74 (0.48–1.17)	0.0	Viable
MEPE	1.6	7	4	0.57 (0.28–1.29)	0.01	Viable
IBSP	1.0	17.5	17	0.97 (0.67–1.46)	0.0	Viable

Osteopontin (SPP1), Dentin matrix protein 1 (DMP1), Dentin sialophosphoglycoprotein (DSPP), Matrix extracellular phosphoglycoprotein (MEPE), and Bone sialoprotein (IBSP).

Table 6 Missense mutation frequencies in matricellular protein genes

Gene	Expected missense SNVs	Observed missense SNVs	Observed/expected (90% range)	Z score
THBS1	721.4	516	0.72 (0.67–0.77)	2.72
THBS2	758.7	587	0.77 (0.72–0.83)	2.21
THBS3	574.9	455	0.79 (0.73–0.85)	1.78
THBS4	553.9	474	0.86 (0.79–0.92)	1.21
COMP	454.8	348	0.77 (0.7–0.84)	1.78
SPARC	180.5	139	0.77 (0.67–0.89)	1.10
SPARCL1	340.3	337	0.99 (0.91–1.08)	0.06
SPOCK1	241.7	190	0.79 (0.7–0.89)	1.18
SPOCK2	249	199	0.8 (0.71–0.9)	1.13
SPOCK3	235.9	219	0.93 (0.83–1.04)	0.39
SMOC1	246.8	214	0.87 (0.78–0.97)	0.74
SMOC2	284.5	254	0.89 (0.81–0.99)	0.64
CCN1	216.2	201	0.93 (0.83–1.04)	0.37
CCN2	182.8	163	0.89 (0.78–1.01)	0.52
CCN3	201.3	192	0.95 (0.85–1.07)	0.23
CCN4	245.5	236	0.96 (0.86–1.07)	0.22
CCN5	156.8	144	0.92 (0.8–1.05)	0.36
CCN6	190.2	184	0.97 (0.86–1.09)	0.16
TNC	1287.7	1296	1.01 (0.96–1.05)	-0.08
TNXB	246.9	245	0.99 (0.89–1.1)	0.04
TNN	781.6	824	1.05 (0.99–1.12)	-0.54
TNR	813	685	0.84 (0.79–0.9)	1.60
SPP1	176.5	166	0.94 (0.83–1.07)	0.28
DMP1	267.6	255	0.95 (0.86–1.06)	0.27
DSPP	676.8	596	0.88 (0.82–0.94)	1.10
MEPE	276.8	268	0.97 (0.88–1.07)	0.19
IBSP	173.2	156	0.9 (0.79–1.03)	0.46

Higher positive Z-scores indicate increased selective pressure to limit missense mutations

being essential for viability in mice and potential cross compensation between *SMOC1* and *SMOC2* (DeGroot et al. 2019), *SMOC1* and *SMOC2* had similar moderate deficits in missense SNVs (Z = 0.74 and 0.64, respectively).

Among the tenascins, only TNR had a deficit in observed versus expected missense SNVs that exceeded the 90% range (Z = 1.60, Table 6). Several TNR missense variants (C155S, T166A, N180H, T592A) were previously linked to familial Parkinson disease, but their pathologic significance remained uncertain (Farlow et al. 2016), The variant p.Cys155Ser occurred in 13 alleles in gnomAD with a frequency of 4.63×10^{-5}, p.Thr166Ala in 1175 alleles (4.2×10^{-3}), p.Asn180His in 1164 alleles (4.12×10^{-3}), and p.Thr592Ala in 53 alleles (1.98×10^{-4}). The high frequencies of these variants raise caution regarding their disease relevance. Missense variants in TNR (p.Arg1192Trp, p.Ala397Thr) were also linked to a nonprogressive neurodevelopmental disorder with spasticity and transient opisthotonos, with the former variant being clinically significant (Wagner et al. 2020). The R1192W variant was not found in the gnomAD data, which supports its role in disease. The A397T variant occurred in 4 alleles with a frequency of 1.59×10^{-5}, suggesting need for further investigation.

In the SIBLING family, only DSPP had a deficit in observed versus expected missense SNVs that exceeded the 90% range (Z = 1.10, Table 6). Multiple deletion and missense mutations in *DSPP* including p.Ala15Val, p.Pro17Thr, p.Val18Phe, and p.Pro19Leu variants result in dentinogenesis imperfecta, deafness, and autosomal dominant nonsyndromic sensorineural 39 (de La Dure-Molla et al. 2015; Liang et al. 2019). Among these residues only p.Pro17Ser had a single variant in gnomAD. The occurrence of multiple pathogenesis-associated missense mutations in DSPP may account for the overall deficit in mutations in this gene among healthy individuals.

Intolerance to LoF mutations in thrombospondin-1 receptors

The pLI values for known thrombospondin-1 receptors were examined to identify genetic evidence for which receptors mediate critical functions of thrombospondin-1 (Table 7). CD36 was the first receptor linked to the anti-angiogenic activity of thrombospondin-1 (Dawson et al. 1997), but LoF mutations in *CD36* occurred at 2.8 times the expected rate for this gene (pLI 0.0, Table 7). The elevated frequency of LoF mutations is consistent with prior reports that genetic deficiencies associated with loss of CD36 expression on red blood cells (Nak^a-negative phenotype) are common in Asian and African populations (Curtis and Aster 1996; Hirano et al. 2003). Type 1 LoF *CD36* mutations in these populations may be related to resistance to malaria (Chilongola et al. 2009; Liu et al. 2020), which would provide a rationale for the high number of observed LoF mutants. Conversely, deletion of *CD36* has been linked with cardiovascular disease and to insulin resistance (Miyaoka et al. 2001; Yuasa-Kawase et al. 2012). These functions of CD36 may be independent of its role as a thrombospondin-1 receptor, thereby accounting for the divergence of the pLI values for these genes.

CD47 mediates signaling functions of CD47 in several cell types (Soto-Pantoja et al. 2015). In contrast to *CD36*, *CD47* had a low frequency of LoF mutations and high loss intolerance (pLI = 0.92, Table 7). CD47 is not essential for viability in mice (Lindberg et al. 1996; Soto-Pantoja et al. 2015), and loss of CD47 expression has only been reported in human red blood cells in the context of mutations in protein 4.2 (*EPB42*) that cause hereditary spherocytosis (Bruce et al. 2002). CD47 also has an independent function as the counter-receptor of SIRPα (Barclay and Van den Berg 2014), and *SIRPA* also had elevated loss-intolerance

Fig. 1 Missense mutations of putative functional sequences in human thrombospondin-1 (P07996). Yellow highlighted regions indicate peptide sequences reported to engage the indicated thrombospondin-1 receptors or ligands. Residues are colored based on variant frequency: no variants (black), > 10^{-6} (blue), > 10^{-5} (green), > 10^{-4} (orange), and 10^{-3} to 10^{-1} (red)

(pLI = 0.67, o/e = 0.19 (0.09–0.43). Therefore, the basis for loss intolerance in human *CD47* remains unclear, but it is a candidate for playing a significant role in the loss intolerance of *THBS1*. Analysis of the frequency of coding variants in CD47 revealed only rare variations in residues involved in its interaction with the counter-receptor SIRPα (Hatherley et al. 2008) and those subject to post-translational modifications including the glycosylation required for THBS1-dependent signaling (Kaur et al. 2011) (Fig. 2).

Thrombospondin-1 binding increases the enzymatic activity of the membrane-bound tyrosine phosphatase CD148, encoded by *PTPRJ* (Takahashi et al. 2012). *PTPRJ* is a tumor suppressor gene, and allelic loss or loss of heterozygosity (LOH) occurs in human sporadic colorectal, lung and breast carcinomas and non-Hodgkin's lymphomas (Aya-Bonilla et al. 2013; Ruivenkamp et al. 2002). Two LoF variants that result in frameshift and insertion of a premature stop codon in *PTPRJ* (g.48131608A > G (c.97-2A > G and g.48158556delG (c.1875delG) are associated with an inherited autosomal recessive thrombocytopenia (Marconi et al. 2019). Consistent with *Ptprj*-deficient mice being viable, fertile, and without any anatomical abnormalities (Trapasso et al. 2006), *PTPRJ* did not exhibit

Table 7 LoF mutants in genes encoding thrombospondin-1 receptors

Gene	Expected LoF mutants	Observed LoF mutants	Observed/expected (90% range)	pLI	Null mouse phenotype
CD36	23.4	66	2.82 (1.77–2)	0.0	Viable
CD47	16.9	2	0.12 (0.05–0.37)	0.92	Viable
PTPRJ (CD148)	62.4	29	0.46 (0.34–0.63)	0.0	Viable/lethal
CACNA2D1 (α2δ1)	74.4	11	0.15 (0.09–0.24)	1.00	Viable
STIM1	31.3	6	0.19 (0.1–0.38)	0.78	Peri- and postnatal lethal
LRP1	246.9	8	0.03 (0.02–0.06)	1.00	Embryonic lethal
ITGA3	58.9	22	0.37 (0.27–0.53)	0.0	Perinatal lethal
ITGA4	62.9	19	0.3 (0.21–0.44)	0.0	Embryonic lethal
ITGA6	61.6	23	0.37 (0.27–0.53)	0.0	Perinatal lethal
ITGB1	38.7	6	0.15 (0.08–0.31)	0.98	Embryonic lethal
ITGAV	62.9	19	0.3 (0.21–0.44)	0.0	Viable
ITGB3	37.4	12	0.32 (0.2–0.52)	0.0	Reduced viability

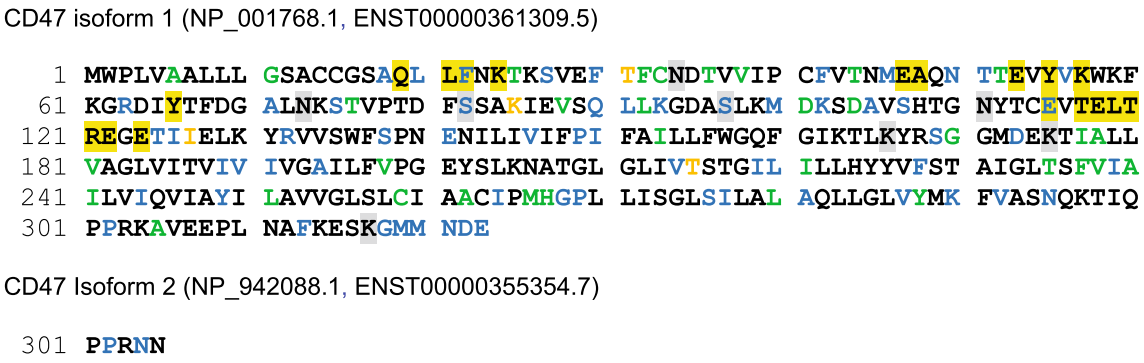

Fig. 2 Missense mutation frequencies in human CD47. Residues involved in SIRPα binding (yellow highlight) generally lack mutations. Residues subject to posttranslational modifications are highlighted in gray. Thrombospondin-1 signaling requires posttranslational modification of S^{82} (Kaur et al. 2011)

an elevated pLI (Table 7). Therefore, this interaction is unlikely to account for the loss intolerance of *THBS1*.

The auxiliary subunit of voltage-gated calcium channels α2δ1 interacts with several members of the *THBS* gene family that regulate channel function including thrombospondin-1 (Eroglu et al. 2009; Taylor and Harris 2020). The low frequency of LoF for *CACNA2D1* resulted in a pLI of 1.0 (Table 7), which is consistent with clinical pathologies associated with *CACNA2D1* mutations. Missense mutations in the extracellular region of α2δ1 at c.2867C > A p.S956T, c.2126 G > A p.S709N, and in the Cache domain (c.1648 G > T p.D550Y) have been associated with early repolarization syndrome (ERS) and inherited Brugada syndrome /J-wave syndromes that cause sudden cardiac death (Burashnikov et al. 2010). Loss of function mutants of *CACNA2D1* were also reported in a patient with Short QT syndrome (Templin et al. 2011) and associated with epilepsy and intellectual disability (Vergult et al. 2015). Because THBS4 also interacts with and regulates α2δ1 but was not loss intolerant, the relevance of this receptor to the elevated pLI for *THBS1* is unclear.

Stromal interaction molecule (STIM1) is an essential regulator of store-operated Ca2 + entry (SOCE) and Ca^{2+} release activated Ca^{2+} (CRAC) channels by binding to *ORAI1* (Feske 2010). In addition to its intracellular roles in calcium signaling, some STIM1 is on the cell surface and interacts with thrombospondin-1 (Ambily et al. 2014; Duquette et al. 2014). *STIM1* had a deficit in LoF mutations that indicates loss intolerance (pLI = 0.78, Table 7), which is consistent with the perinatal lethality of the null mouse (Varga-Szabo et al. 2008). LoF mutations in human *STIM1* and *ORAI1* abolished CRAC and SOCE channel currents and are associated with severe combined immunodeficiency, congenital myopathy, and anhydrotic ectodermal dysplasia (Feske 2010; Lacruz and Feske 2015). Therefore, *STIM1* is also a candidate to play a role in the loss intolerance of *THBS1*.

Low density lipoprotein receptor-related protein 1 (LRP1) is a scavenger receptor that mediates endocytosis of multiple ligands including lipoproteins and Thrombospondin-1 (Gonias et al. 2004). LRP1 modulates downstream intracellular signaling controlling cell survival associated with tissue remodeling in response to injury via a thrombospondin-1-calreticulin complex (Pallero et al. 2008). LoF mutations in *LRP1* were rare, indicating a high degree of loss intolerance (pLI = 1.0, Table 7), consistent with the embryonic lethal phenotype of *Lrp1* null mice at embryonic implantation (Herz et al. 1992). The promiscuity of this receptor precludes assessment of its relevance to the loss intolerance of *THBS1*.

Several integrin heterodimers act as thrombospondin-1 receptors including α3β1, α4β1, α6β1, and αvβ3 (Calzada and Roberts 2005; Resovi et al. 2014). Of the genes encoding their respective subunits, only *ITGB1* showed an elevated (pLI = 0.98, Table 7). Although *Itgb1* is essential for embryonic development in mice, the role of β1-integrins as receptors for multiple ECM proteins including members of the THBS, CCN, and tenascin families precludes assigning a specific role for any individual integrin in the loss intolerance for *THBS1* (Humphries et al. 2006).

Discussion

The elevated pLI values observed for several matricellular protein genes infer that LoF mutations in those genes confers a significant survival or reproductive disadvantage in humans (Lek et al. 2016). When homozygous knockout of the murine ortholog indicates an essential role in development, as reported for *CCN1*, loss-intolerance for the human gene is consistent with a similar role in human embryonic development. On the other hand, loss-intolerant matricellular genes such as *THBS1*, *THBS2*, *SPARC*, *SPOCK1*, and *TNR* that are not essential for murine embryonic development may have critical roles in human postnatal survival or in adult reproductive function. In the case of *THBS1*, the selective pressure to maintain this gene may have more than one origin. Studies in mice and primates indicated specific roles for *Thbs1* in reproduction (Bender et al. 2019; Greenaway et al. 2007) and in the ability of adult mice, rats and pigs to repair dermal wounds and survive exposure to ischemic injuries or genotoxic stress (Isenberg et al. 2007, 2008a, b; Soto-Pantoja et al. 2015). Loss of *Thbs1* in mice also alters their survival following exposure to several pathogens (Arun et al. 2020; Binsker et al. 2019; Lawler et al. 1998; Martin-Manso et al. 2012; Qu et al. 2018). Loss of *Thbs1* impairs survival for some of these stresses while improving survival or recovery from other stresses. These animal studies suggest that LoF mutants in human *THBS1* could alter the probability of surviving acute injuries and infections, some of which could lead to a decrease in longevity or success in reproduction. The multiplicity of functions for thrombospondin-1 and other matricellular proteins suggests that no single function will account for the selection against individuals carrying LoF or missense mutants.

Although the expectation–maximization algorithm used in calculating the pLI values compensates for the influence of coding sequence length on the expected number of LoF mutants, genes with longer coding sequences remain more likely to achieve a significant pLI at a given sample number (Lek et al. 2016). Therefore, the predictions of loss-intolerance are more reliable for the thrombospondin and tenascin family members with longer ORFs than for CCN or SIBLING family genes. Despite this limitation, the *CCN1* data demonstrated loss-intolerance in humans, as expected based on its essential role in mouse embryonic development (Mo and Lau 2006; Mo et al. 2002). Additional factors including the breadth of tissue expression also correlate broadly with obtaining an elevated pLI (Lek et al. 2016), suggesting that matricellular genes with more restricted tissue expression or organ-specific essential functions are less likely to be detected in the gnomAD data. Temporal differences in the expression of matricellular proteins may also be a factor, as was documented for THBS1 versus THBS2 during dermal wound repair (Agah et al. 2002).

Significant loss intolerance was not demonstrated for some of the genes with shorter ORFs such as SMOC1, but the 90% confidence range extended beyond the cutoff for significant loss-intolerance, consistent with its essential role in murine embryonic development and other evidence that *SMOC1* plays an important role in human embryonic development (Okada et al. 2011; Rainger et al. 2011). As the number of available human genomes in gnomAD increases, additional matricellular genes may be identified to be significantly LoF-intolerant.

The overall deficit in missense mutations in *THBS1* and *THBS2* are consistent with selection pressures to maintain the functional integrity of these proteins. However, the distribution of missense mutations in *THBS1* did not clearly identify specific residues or regions of the protein that mediate these functions beyond those residues previously identified through genome-wide association studies linking one polymorphism in THBS1 with cardiovascular disease. In contrast, multiple residues in *DSPP* that are subject to disease-causing missense mutations were invariant in the gnomAD data. Thus, another use for this broad population data is to validate the absence of previously identified disease-causing mutations in a large healthy population. Further analyses of missense mutations in other matricellular genes may also provide insights into specific interactions that mediate functional roles of these proteins in human development and disease.

Recent advances in understanding the complex effects of LoF mutations on gene regulation may help explain the observed divergence between pLI values for human matricellular protein genes such as *THBS1*, *SMOC1* and *CCN2* and the viability of mice bearing LoF mutants in the corresponding murine orthologs. One advantage of the gnomAD LoF analysis over the knockout mouse studies is that pLI values are derived from multiple independent LoF mutants in each human gene, whereas the mouse phenotypes are typically based on a single gene knockout strategy. Gene knockout strategies can have unanticipated effects that extend beyond the targeted gene. As was reported for *Thbs3* in mice, matricellular gene regulation may involve elements located within an adjacent essential gene (Collins et al. 1998). Using different methods to inactivate a gene can also result in contradictory phenotypes by triggering compensatory responses including transcriptional adaptation (Kontarakis and Stainier 2020). A subset of mutations causing premature transcript termination can result in gain of function rather than LoF alleles (Coban-Akdemir et al. 2018), An analysis of the ExAC data using an algorithm to predict such dominant gain of function alleles did not identify potential gain of function alleles for *THBS1*, *SMOC1* and *CCN2* (Coban-Akdemir et al. 2018). However, potential gain of function alleles were identified for *DMP1* and *DSPP* in the ExAC data (Supplemental Table 4 in Coban-Akdemir et al. 2018). Based on these data, clinically relevant gain of function mutants in SIBLING family genes should be further investigated.

Funding This work was supported by the Intramural Research Program of the NIH/NCI (ZIA SC009172).

Availability of data and materials All data is contained in the manuscript or the indicated public databases.

Compliance with ethical standards

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

Abouzeid H et al (2011) Mutations in the SPARC-related modular calcium-binding protein 1 gene, SMOC1, cause waardenburg anophthalmia syndrome. Am J Hum Genet 88:92–98. https://doi.org/10.1016/j.ajhg.2010.12.002

Agah A, Kyriakides TR, Lawler J, Bornstein P (2002) The lack of thrombospondin-1 (TSP1) dictates the course of wound healing in double-TSP1/TSP2-null mice. Am J Pathol 161:831–839. https://doi.org/10.1016/s0002-9440(10)64243-5

Ambily A et al (2014) The role of plasma membrane STIM1 and Ca(2+)entry in platelet aggregation. STIM1 binds to novel proteins in human platelets. Cell Signal 26:502–511. https://doi.org/10.1016/j.cellsig.2013.11.025

Arun A et al (2020) Thrombospondin-1 plays an essential role in yes-associated protein nuclear translocation during the early phase of trypanosoma cruzi infection in heart endothelial cells. Int J Mol Sci. https://doi.org/10.3390/ijms21144912

Aya-Bonilla C et al (2013) High-resolution loss of heterozygosity screening implicates PTPRJ as a potential tumor suppressor gene that affects susceptibility to Non-Hodgkin's lymphoma. Genes Chromosomes Cancer 52:467–479. https://doi.org/10.1002/gcc.22044

Baguma-Nibasheka M, Kablar B (2008) Pulmonary hypoplasia in the connective tissue growth factor (Ctgf) null mouse. DevDyn 237:485–493. https://doi.org/10.1002/dvdy.21433

Barclay AN, Van den Berg TK (2014) The interaction between signal regulatory protein alpha (SIRPalpha) and CD47: structure, function, and therapeutic target. Annu Rev Immunol 32:25–50. https://doi.org/10.1146/annurev-immunol-032713-120142

Bender HR, Campbell GE, Aytoda P, Mathiesen AH, Duffy DM (2019) Thrombospondin 1 (THBS1) promotes follicular angiogenesis, luteinization, and ovulation in primates. Front Endocrinol (Lausanne) 10:727. https://doi.org/10.3389/fendo.2019.00727

Bentley SR et al (2020) Evidence of a recessively inherited CCN3 mutation as a rare cause of early-onset parkinsonism. Front Neurol 11:331. https://doi.org/10.3389/fneur.2020.00331

Binsker U, Kohler TP, Hammerschmidt S (2019) Contribution of human thrombospondin-1 to the pathogenesis of gram-positive bacteria. J Innate Immun 11:303–315. https://doi.org/10.1159/000496033

Bornstein P (1995) Diversity of function is inherent in matricellular proteins: an appraisal of thrombospondin 1. J Cell Biol 130:503–506. https://doi.org/10.1083/jcb.130.3.503

Bouleftour W et al (2016) The role of the SIBLING, Bone Sialoprotein in skeletal biology - Contribution of mouse experimental genetics. Matrix Biol 52–54:60–77. https://doi.org/10.1016/j.matbio.2015.12.011

Bradshaw AD (2009) The role of SPARC in extracellular matrix assembly. J Cell Commun Signal 3:239–246. https://doi.org/10.1007/s12079-009-0062-6

Bristow J, Carey W, Egging D, Schalkwijk J (2005) Tenascin-X, collagen, elastin, and the Ehlers–Danlos syndrome. Am J Med Genet C Semin Med Genet 139C:24–30. https://doi.org/10.1002/ajmg.c.30071

Bruce LJ et al (2002) Absence of CD47 in protein 4.2-deficient hereditary spherocytosis in man: an interaction between the Rh complex and the band 3 complex. Blood 100:1878–1885. https://doi.org/10.1182/blood-2002-03-0706

Bult CJ, Blake JA, Smith CL, Kadin JA, Richardson JE, Mouse Genome Database G (2019) Mouse Genome Database (MGD) 2019. Nucleic Acids Res 47:D801–D806. https://doi.org/10.1093/nar/gky1056

Burashnikov E et al (2010) Mutations in the cardiac L-type calcium channel associated with inherited J-wave syndromes and sudden cardiac death. Heart Rhythm 7:1872–1882. https://doi.org/10.1016/j.hrthm.2010.08.026

Burke A, Creighton W, Tavora F, Li L, Fowler D (2009) Decreased frequency of the 3'UTR T>G single nucleotide polymorphism of thrombospondin-2 gene in sudden death due to plaque erosion. Cardiovasc Pathol. https://doi.org/10.1016/j.carpath.2008.12.013

Calabro NE, Kristofik NJ, Kyriakides TR (2014) Thrombospondin-2 and extracellular matrix assembly. Biochim Biophys Acta 1840:2396–2402. https://doi.org/10.1016/j.bbagen.2014.01.013

Calzada MJ, Roberts DD (2005) Novel integrin antagonists derived from thrombospondins. Curr Pharm Des 11:849–866

Canalis E, Smerdel-Ramoya A, Durant D, Economides AN, Beamer WG, Zanotti S (2010) Nephroblastoma overexpressed (Nov) inactivation sensitizes osteoblasts to bone morphogenetic protein-2, but Nov is dispensable for skeletal homeostasis. Endocrinology 151:221–233. https://doi.org/10.1210/en.2009-0574

Carlson CB, Liu Y, Keck JL, Mosher DF (2008) Influences of the N700S thrombospondin-1 polymorphism on protein structure and stability. J Biol Chem 283:20069–20076. https://doi.org/10.1074/jbc.m800223200

Chilongola J, Balthazary S, Mpina M, Mhando M, Mbugi E (2009) CD36 deficiency protects against malarial anaemia in children by reducing *Plasmodium falciparum*-infected red blood cell adherence to vascular endothelium. Trop Med Int Health 14:810–816. https://doi.org/10.1111/j.1365-3156.2009.02298.x

Coban-Akdemir Z et al (2018) Identifying genes whose mutant transcripts cause dominant disease traits by potential gain-of-function alleles. Am J Hum Genet 103:171–187. https://doi.org/10.1016/j.ajhg.2018.06.009

Collins M, Rojnuckarin P, Zhu YH, Bornstein P (1998) A far upstream, cell type-specific enhancer of the mouse thrombospondin 3 gene is located within intron 6 of the adjacent metaxin gene. J Biol Chem 273:21816–21824. https://doi.org/10.1074/jbc.273.34.21816

Curtis BR, Aster RH (1996) Incidence of the Nak(a)-negative platelet phenotype in African Americans is similar to that of Asians. Transfusion 36:331–334. https://doi.org/10.1046/j.1537-2995.1996.36496226147.x

Dawson DW, Pearce SF, Zhong R, Silverstein RL, Frazier WA, Bouck NP (1997) CD36 mediates the In vitro inhibitory effects of thrombospondin-1 on endothelial cells. J Cell Biol 138:707–717. https://doi.org/10.1083/jcb.138.3.707

de La Dure-Molla M, Philippe Fournier B, Berdal A (2015) Isolated dentinogenesis imperfecta and dentin dysplasia: revision of the classification. Eur J Hum Genet 23:445–451. https://doi.org/10.1038/ejhg.2014.159

DeGroot MS, Shi H, Eastman A, McKillop AN, Liu J (2019) The *Caenorhabditis elegans* SMOC-1 protein acts cell nonautonomously to promote bone morphogenetic protein signaling. Genetics 211:683–702. https://doi.org/10.1534/genetics.118.301805

Dhamija R, Graham JM Jr, Smaoui N, Thorland E, Kirmani S (2014) Novel de novo SPOCK1 mutation in a proband with developmental delay, microcephaly and agenesis of corpus callosum. Eur J Med Genet 57:181–184. https://doi.org/10.1016/j.ejmg.2014.02.009

Dufresne D, Hamdan FF, Rosenfeld JA, Torchia B, Rosenblatt B, Michaud JL, Srour M (2012) Homozygous deletion of Tenascin-R in a patient with intellectual disability. J Med Genet 49:451–454. https://doi.org/10.1136/jmedgenet-2012-100831

Duquette M, Nadler M, Okuhara D, Thompson J, Shuttleworth T, Lawler J (2014) Members of the thrombospondin gene family bind stromal interaction molecule 1 and regulate calcium channel activity. Matrix Biol 37:15–24. https://doi.org/10.1016/j.matbio.2014.05.004

Elola MT, Wolfenstein-Todel C, Troncoso MF, Vasta GR, Rabinovich GA (2007) Galectins: matricellular glycan-binding proteins linking cell adhesion, migration, and survival. Cell Mol Life Sci 64:1679–1700. https://doi.org/10.1007/s00018-007-7044-8

Eroglu C et al (2009) Gabapentin receptor alpha2delta-1 is a neuronal thrombospondin receptor responsible for excitatory CNS synaptogenesis. Cell 139:380–392. https://doi.org/10.1016/j.cell.2009.09.025

Farlow JL et al (2016) Whole-exome sequencing in familial parkinson disease. JAMA Neurol 73:68–75. https://doi.org/10.1001/jamaneurol.2015.3266

Feng JQ et al (2006) Loss of DMP1 causes rickets and osteomalacia and identifies a role for osteocytes in mineral metabolism. Nat Genet 38:1310–1315. https://doi.org/10.1038/ng1905

Feske S (2010) CRAC channelopathies. Pflugers Arch 460:417–435. https://doi.org/10.1007/s00424-009-0777-5

Frolova EG et al (2010) Thrombospondin-4 regulates vascular inflammation and atherogenesis. Circ Res 107:1313–1325. https://doi.org/10.1161/CIRCRESAHA.110.232371

Gilmour DT et al (1998) Mice deficient for the secreted glycoprotein SPARC/osteonectin/BM40 develop normally but show severe age-onset cataract formation and disruption of the lens. EMBO J 17:1860–1870. https://doi.org/10.1093/emboj/17.7.1860

Gonias SL, Wu L, Salicioni AM (2004) Low density lipoprotein receptor-related protein: regulation of the plasma membrane proteome. Thromb Haemost 91:1056–1064. https://doi.org/10.1160/TH04-01-0023

Greenaway J, Lawler J, Moorehead R, Bornstein P, Lamarre J, Petrik J (2007) Thrombospondin-1 inhibits VEGF levels in the ovary directly by binding and internalization via the low density lipoprotein receptor-related protein-1 (LRP-1). J Cell Physiol 210:807–818. https://doi.org/10.1002/jcp.20904

Hankenson KD, Hormuzdi SG, Meganck JA, Bornstein P (2005a) Mice with a disruption of the thrombospondin 3 gene differ in geometric and biomechanical properties of bone and have accelerated development of the femoral head. Mol Cell Biol 25:5599–5606. https://doi.org/10.1128/mcb.25.13.5599-5606.2005

Hankenson KD et al (2005b) Increased osteoblastogenesis and decreased bone resorption protect against ovariectomy-induced bone loss in thrombospondin-2-null mice. Matrix Biol 24:362–370. https://doi.org/10.1016/j.matbio.2005.05.008

Hatherley D, Graham SC, Turner J, Harlos K, Stuart DI, Barclay AN (2008) Paired receptor specificity explained by structures of signal regulatory proteins alone and complexed with CD47. Mol Cell 31:266–277. https://doi.org/10.1016/j.molcel.2008.05.026

Hawi Z et al (2018) A case-control genome-wide association study of ADHD discovers a novel association with the tenascin R (TNR) gene. Transl Psychiatry 8:284. https://doi.org/10.1038/s41398-018-0329-x

Herz J, Clouthier DE, Hammer RE (1992) LDL receptor-related protein internalizes and degrades uPA-PAI-1 complexes and is essential for embryo implantation. Cell 71:411–421. https://doi.org/10.1016/0092-8674(92)90511-a

Hirano K, Kuwasako T, Nakagawa-Toyama Y, Janabi M, Yamashita S, Matsuzawa Y (2003) Pathophysiology of human genetic CD36 deficiency. Trends Cardiovasc Med 13:136–141. https://doi.org/10.1016/s1050-1738(03)00026-4

Humphries JD, Byron A, Humphries MJ (2006) Integrin ligands at a glance. J Cell Sci 119:3901–3903. https://doi.org/10.1242/jcs.03098

Hurvitz JR et al (1999) Mutations in the CCN gene family member WISP3 cause progressive pseudorheumatoid dysplasia. Nat Genet 23:94–98. https://doi.org/10.1038/12699

Isenberg JS et al (2007) Thrombospondin-1 limits ischemic tissue survival by inhibiting nitric oxide-mediated vascular smooth muscle relaxation. Blood 109:1945–1952. https://doi.org/10.1182/blood-2006-08-041368

Isenberg JS et al (2008a) Thrombospondin-1 and CD47 limit cell and tissue survival of radiation injury. Am J Pathol 173:1100–1112. https://doi.org/10.2353/ajpath.2008.080237

Isenberg JS, Romeo MJ, Maxhimer JB, Smedley J, Frazier WA, Roberts DD (2008b) Gene silencing of CD47 and antibody ligation of thrombospondin-1 enhance ischemic tissue survival in a porcine model: implications for human disease. Ann Surg 247:860–868. https://doi.org/10.1097/SLA.0b013e31816c4006

Ivkovic S et al (2003) Connective tissue growth factor coordinates chondrogenesis and angiogenesis during skeletal development. Development 130:2779–2791. https://doi.org/10.1242/dev.00505

Jones FS, Jones PL (2000) The tenascin family of ECM glycoproteins: structure, function, and regulation during embryonic development and tissue remodeling. Dev Dyn 218:235–259

Karczewski KJ et al (2020) The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 581:434–443. https://doi.org/10.1038/s41586-020-2308-7

Kaur S et al (2011) Heparan sulfate modification of the transmembrane receptor CD47 is necessary for inhibition of T cell receptor signaling by thrombospondin-1. J Biol Chem 286:14991–15002. https://doi.org/10.1074/jbc.M110.179663

Kawaki H et al (2008) Functional requirement of CCN2 for intramembranous bone formation in embryonic mice. Biochem Biophys Res Commun 366:450–456. https://doi.org/10.1016/j.bbrc.2007.11.155

Kim KH, Won JH, Cheng N, Lau LF (2018) The matricellular protein CCN1 in tissue injury repair. J Cell Commun Signal 12:273–279. https://doi.org/10.1007/s12079-018-0450-x

Knight MN, Hankenson KD (2014) R-spondins: novel matricellular regulators of the skeleton. Matrix Biol 37:157–161. https://doi.org/10.1016/j.matbio.2014.06.003

Kontarakis Z, Stainier DYR (2020) Genetics in light of transcriptional adaptation. Trends Genet 36:926–935. https://doi.org/10.1016/j.tig.2020.08.008

Kutz WE, Gong Y, Warman ML (2005) WISP3, the gene responsible for the human skeletal disease progressive pseudorheumatoid dysplasia, is not essential for skeletal function in mice. Mol Cell Biol 25:414–421. https://doi.org/10.1128/mcb.25.1.414-421.2005

Lacruz RS, Feske S (2015) Diseases caused by mutations in ORAI1 and STIM1. Ann N Y Acad Sci 1356:45–79. https://doi.org/10.1111/nyas.12938

Lawler J, Sunday M, Thibert V, Duquette M, George EL, Rayburn H, Hynes RO (1998) Thrombospondin-1 is required for normal murine pulmonary homeostasis and its absence causes pneumonia. J Clin Invest 101:982–992. https://doi.org/10.1172/JCI1684

Leask A (2020) Conjunction junction, what's the function? CCN proteins as targets in fibrosis and cancers. Am J Physiol Cell Physiol 318:C1046–C1054. https://doi.org/10.1152/ajpcell.00028.2020

Lek M et al (2016) Analysis of protein-coding genetic variation in 60,706 humans. Nature 536:285–291. https://doi.org/10.1038/nature19057

Liang T, Zhang H, Xu Q, Wang S, Qin C, Lu Y (2019) Mutant dentin sialophosphoprotein causes dentinogenesis imperfecta. J Dent Res 98:912–919. https://doi.org/10.1177/0022034519854029

Lindberg FP, Bullard DC, Caver TE, Gresham HD, Beaudet AL, Brown EJ (1996) Decreased resistance to bacterial infection and granulocyte defects in IAP-deficient mice. Science 274:795–798. https://doi.org/10.1126/science.274.5288.795

Liu J et al (2020) Distribution of CD36 deficiency in different Chinese ethnic groups. Hum Immunol 81:366–371. https://doi.org/10.1016/j.humimm.2020.05.004

Lorenz-Depiereux B et al (2006) DMP1 mutations in autosomal recessive hypophosphatemia implicate a bone matrix protein in the regulation of phosphate homeostasis. Nat Genet 38:1248–1250. https://doi.org/10.1038/ng1868

Marconi C et al (2019) Loss-of-function mutations in PTPRJ cause a new form of inherited thrombocytopenia. Blood 133:1346–1357. https://doi.org/10.1182/blood-2018-07-859496

Martin-Manso G, Navarathna DH, Galli S, Soto-Pantoja DR, Kuznetsova SA, Tsokos M, Roberts DD (2012a) Endogenous thrombospondin-1 regulates leukocyte recruitment and activation and accelerates death from systemic candidiasis. PLoS ONE 7:e48775. https://doi.org/10.1371/journal.pone.0048775

McMaken S et al (2011) Thrombospondin-1 contributes to mortality in murine sepsis through effects on innate immunity. PLoS ONE 6:e19654. https://doi.org/10.1371/journal.pone.0019654

Mendoza-Londono R et al (2015) Recessive osteogenesis imperfecta caused by missense mutations in SPARC. Am J Hum Genet 96:979–985. https://doi.org/10.1016/j.ajhg.2015.04.021

Midwood KS, Orend G (2009) The role of tenascin-C in tissue injury and tumorigenesis. J Cell Commun Signal. https://doi.org/10.1007/s12079-009-0075-1

Miyaoka K, Kuwasako T, Hirano K, Nozaki S, Yamashita S, Matsuzawa Y (2001) CD36 deficiency associated with insulin resistance. Lancet 357:686–687. https://doi.org/10.1016/s0140-6736(00)04138-6

Mo FE, Lau LF (2006) The matricellular protein CCN1 is essential for cardiac development. Circ Res 99:961–969. https://doi.org/10.1161/01.RES.0000248426.35019.89

Mo FE, Muntean AG, Chen CC, Stolz DB, Watkins SC, Lau LF (2002) CYR61 (CCN1) is essential for placental development and vascular integrity. Mol Cell Biol 22:8709–8720. https://doi.org/10.1128/mcb.22.24.8709-8720.2002

Murphy-Ullrich JE, Sage EH (2014) Revisiting the matricellular concept. Matrix Biol 37:1–14. https://doi.org/10.1016/j.matbio.2014.07.005

Nagaraju GP, Dontula R, El-Rayes BF, Lakka SS (2014) Molecular mechanisms underlying the divergent roles of SPARC in human carcinogenesis. Carcinogenesis 35:967–973. https://doi.org/10.1093/carcin/bgu072

Nakamura T (2018) Roles of short fibulins, a family of matricellular proteins, in lung matrix assembly and disease. Matrix Biol 73:21–33. https://doi.org/10.1016/j.matbio.2018.02.003

Okada I et al (2011) SMOC1 is essential for ocular and limb development in humans and mice. Am J Hum Genet 88:30–41. https://doi.org/10.1016/j.ajhg.2010.11.012

Pallero MA, Elzie CA, Chen J, Mosher DF, Murphy-Ullrich JE (2008) Thrombospondin 1 binding to calreticulin-LRP1 signals resistance to anoikis. Faseb J 22:3968–3979. https://doi.org/10.1096/fj.07-104802

Posey KL, Coustry F, Hecht JT (2018) Cartilage oligomeric matrix protein: COMPopathies and beyond. Matrix Biol 71–72:161–173. https://doi.org/10.1016/j.matbio.2018.02.023

Posey KL, Hankenson K, Veerisetty AC, Bornstein P, Lawler J, Hecht JT (2008) Skeletal abnormalities in mice lacking extracellular matrix proteins, thrombospondin-1, thrombospondin-3, thrombospondin-5, and type IX collagen. Am J Pathol 172:1664–1674. https://doi.org/10.2353/ajpath.2008.071094

Qu Y et al (2018) Thrombospondin-1 protects against pathogen-induced lung injury by limiting extracellular matrix proteolysis. JCI Insight. https://doi.org/10.1172/jci.insight.96914

Rainger J et al (2011) Loss of the BMP antagonist, SMOC-1, causes Ophthalmo-acromelic (Waardenburg Anophthalmia) syndrome in humans and mice. PLoS Genet 7:e1002114. https://doi.org/10.1371/journal.pgen.1002114

Resovi A, Pinessi D, Chiorino G, Taraboletti G (2014) Current understanding of the thrombospondin-1 interactome. Matrix Biol 37:83–91. https://doi.org/10.1016/j.matbio.2014.01.012

Roberts DD, Miller TW, Rogers NM, Yao M, Isenberg JS (2012) The matricellular protein thrombospondin-1 globally regulates cardiovascular function and responses to stress. Matrix Biol 31:162–169. https://doi.org/10.1016/j.matbio.2012.01.005

Roll S, Seul J, Paulsson M, Hartmann U (2006) Testican-1 is dispensable for mouse development. Matrix Biol 25:373–381. https://doi.org/10.1016/j.matbio.2006.05.004

Ruivenkamp CA et al (2002) Ptprj is a candidate for the mouse colon-cancer susceptibility locus Scc1 and is frequently deleted in

human cancers. Nat Genet 31:295–300. https://doi.org/10.1038/ng903

Song YL, Wang CN, Fan MW, Su B, Bian Z (2008) Dentin phosphoprotein frameshift mutations in hereditary dentin disorders and their variation patterns in normal human population. J Med Genet 45:457–464. https://doi.org/10.1136/jmg.2007.056911

Soto-Pantoja DR, Kaur S, Roberts DD (2015) CD47 signaling pathways controlling cellular differentiation and responses to stress. Crit Rev Biochem Mol Biol 50:212–230. https://doi.org/10.3109/10409238.2015.1014024

Staines KA, MacRae VE, Farquharson C (2012) The importance of the SIBLING family of proteins on skeletal mineralisation and bone remodelling. J Endocrinol 214:241–255. https://doi.org/10.1530/JOE-12-0143

Stenina-Adognravi O, Plow EF (2019) Thrombospondin-4 in tissue remodeling. Matrix Biol 75–76:300–313. https://doi.org/10.1016/j.matbio.2017.11.006

Stenina OI, Topol EJ, Plow EF (2007) Thrombospondins, their polymorphisms, and cardiovascular disease. Arterioscler Thromb Vasc Biol 27:1886–1894

Stenina OI, Ustinov V, Krukovets I, Marinic T, Topol EJ, Plow EF (2005) Polymorphisms A387P in thrombospondin-4 and N700S in thrombospondin-1 perturb calcium binding sites. Faseb J 19:1893–1895. https://doi.org/10.1096/fj.05-3712fje

Sun LR, Li SY, Guo QS, Zhou W, Zhang HM (2020) SPOCK1 involvement in epithelial-to-mesenchymal transition: a new target in cancer therapy? Cancer Manag Res 12:3561–3569. https://doi.org/10.2147/CMAR.S249754

Svensson L, Aszodi A, Heinegard D, Hunziker EB, Reinholt FP, Fassler R, Oldberg A (2002) Cartilage oligomeric matrix protein-deficient mice have normal skeletal development. Mol Cell Biol 22:4366–4371. https://doi.org/10.1128/mcb.22.12.4366-4371.2002

Takahashi K et al (2012) Thrombospondin-1 acts as a ligand for CD148 tyrosine phosphatase. Proc Natl Acad Sci U S A 109:1985–1990. https://doi.org/10.1073/pnas.1106171109

Taylor CP, Harris EW (2020) Analgesia with gabapentin and pregabalin may involve *N*-methyl-d-aspartate receptors neurexins, and thrombospondins. J Pharmacol Exp Ther 374:161–174. https://doi.org/10.1124/jpet.120.266056

Templin C et al (2011) Identification of a novel loss-of-function calcium channel gene mutation in short QT syndrome (SQTS6). Eur Heart J 32:1077–1088. https://doi.org/10.1093/eurheartj/ehr076

Tolsma SS, Volpert OV, Good DJ, Frazier WA, Polverini PJ, Bouck N (1993) Peptides derived from two separate domains of the matrix protein thrombospondin-1 have anti-angiogenic activity. J Cell Biol 122:497–511. https://doi.org/10.1083/jcb.122.2.497

Topol EJ et al (2001) Single nucleotide polymorphisms in multiple novel thrombospondin genes may be associated with familial premature myocardial infarction. Circulation 104:2641–2644. https://doi.org/10.1161/hc4701.100910

Trapasso F et al (2006) Genetic ablation of Ptprj, a mouse cancer susceptibility gene, results in normal growth and development and does not predispose to spontaneous tumorigenesis DNA. Cell Biol 25:376–382. https://doi.org/10.1089/dna.2006.25.376

Varga-Szabo D et al (2008) The calcium sensor STIM1 is an essential mediator of arterial thrombosis and ischemic brain infarction. J Exp Med 205:1583–1591. https://doi.org/10.1084/jem.20080302

Vergult S et al (2015) Genomic aberrations of the CACNA2D1 gene in three patients with epilepsy and intellectual disability. Eur J Hum Genet 23:628–632. https://doi.org/10.1038/ejhg.2014.141

Wagner M et al (2020) Loss of TNR causes a nonprogressive neurodevelopmental disorder with spasticity and transient opisthotonus. Genet Med 22:1061–1068. https://doi.org/10.1038/s41436-020-0768-7

Weber P et al (1999) Mice deficient for tenascin-R display alterations of the extracellular matrix and decreased axonal conduction velocities in the CNS. J Neurosci 19:4245–4262. https://doi.org/10.1523/jneurosci.19-11-04245.1999

Yuasa-Kawase M et al (2012) Patients with CD36 deficiency are associated with enhanced atherosclerotic cardiovascular diseases. J Atheroscler Thromb 19:263–275. https://doi.org/10.5551/jat.10603

Zhang X et al (2001) DSPP mutation in dentinogenesis imperfecta Shields type II. Nat Genet 27:151–152. https://doi.org/10.1038/84765

Zhao Y et al (2015) Thrombospondin-1 restrains neutrophil granule serine protease function and regulates the innate immune response during *Klebsiella pneumoniae* infection. Mucosal Immunol 8:896–905. https://doi.org/10.1038/mi.2014.120