

Supplementary Tables

Supplementary Table S1. Homology modeling templates for individual filamin domains.

Supplementary Table S2. Short linkers between pairs of filamin A Ig domains.

Supplementary Table S3. Filamin sequences used in phylogenetic analysis.

Supplementary Table S4. Domains and number of isoform common ancestral residues.

Supplementary Table S5. Domains with low or high content of ancestral or distinctive residues.

Supplementary Figures

Supplementary Figure S1. Multiple sequence alignment used to obtain homology model templates of FLN domains. **A.** Alignment of homologous actin-binding domains (ABD) from plectin (PDB ID 1SH5), dystrophin (PDB ID 1DXX), α -actinin 1 (PDB ID 2EYI), α -actinin 3 (PDB ID 1WKU), human filamin A, human filamin B, and human filamin C. Muscle specific α -actinin 3 is used as a template to model filamin A's ABD. **B.** Alignment of target and template (asterisk (*)) filamin Ig-like domains. Excluded are predicted Ig-like domains that have close homologs in the PDB (e.g., we do not include the sequence IgFLNa12 along with IgFLNb12 (PDB ID 2DIC)). We show all isoforms of domain 16 for comparison.

Supplementary Figure S2. Isoform-specific divergence analysis. The analysis of isoform-specific divergence consists of ancestral reconstruction, aligning ancestral sequences, and an evolutionary trace of ancestral sequences. **A.** Segment alignment of a section of the extant filamin protein sequences. Lower alignment is of reconstructed ancestral sequences. X is indeterminate residue. **B.** Evolutionary trace results indicating sites are: “ancestral,” “distinctive,” or “other.” Ancestral implies the same biochemical type is maintained in each isoform. “Distinctive” implies change to a new biochemical type in one isoform and then is fixed, while at least one other isoform is ancestral. “Other” implies a variable site. **C.** The evolution of extant species are viewed as branching off the evolutionary tree that leads to humans. Ancestors at the branch points are Teleostei (blue), Amphibian (green), and Mammalian (yellow).

Supplementary Figure S3. Conservation within domains is bi-modal. The number of residues in each domain that are conserved to the common ancestor across filamin isoforms. N-terminal domains 1-16 alternate between the two conservation modes and have a mean of $62\pm9\%$. C-terminal domains 17-24, have a mean are more highly conserved overall (mean $71\pm8\%$). The histogram, right is bimodal, with a mode at $\sim70\%$ and a mode at $\sim55\%$. Individual highly conserved domains are CH 1 (88%), IgFLNa21 (84%) and IgFLNa22 (79%). Domains with the least conservation include 5, 8, 10, 11 ($\sim50\%$).

Supplementary Table S1.

Modeled domain (target)	Template(s)	PDB ID	Method	Verify-3d score
ABD	Human α -actinin 3 closed conformation	1wku	X-RAY	0.75
Ig-like				
1	hsIgFLNb9 hsIgFLNc14	2di9 2e9j	NMR NMR	0.72
2	hsIgFLNc14	2e9j	NMR	1.12
3	hsIgFLNb19	2di9	NMR	1.42
4	hsIgFLNa17	2bp3	X-RAY	0.87
5	hsIgFLNc14 ddFLN5 from ddFLN4-6	2e9j 1wlh	NMR X-RAY	0.88
6	hsIgFLNc14 ddFLN5 from ddFLN4-6	2d9j 1wlh	NMR X-RAY	0.88
7	hsIgFLNc14, ddFLN5 from ddFLN4-6	2d9j 1wlh	NMR X-RAY	0.97
8	hsIgFLNb10 hsIgFLNb19 ddFLN5 from ddFLN4-6	2dia 2di8 1wlh	NMR NMR X-RAY	0.73
9	hsIgFLNb9	2di9	NMR	1.14
10	hsIgFLNb10	2dia	NMR	1.01
11	hsIgFLNb13	2dj4	NMR	0.51
12	hsIgFLNc12	2dic	NMR	0.89
13	hsIgFLNb13	2dj4	NMR	0.58
14	hsIgFLNc14	2d9j	NMR	0.85
15	hsIgFLNb15	2dmb	NMR	0.72
16	hsIgFLNc10 hsIgFLNc16	2dia 2d7n	NMR NMR	0.64
17	hsIgFLNa17	2bp3	X-RAY	0.98
18	hsIgFLNb18	2dmc	NMR	0.11*
19	hsIgFLNb19	2di8	NMR	1.17
20	hsIgFLNb18	2dmc	NMR	0.65
21	hsIgFLNb19	2di8	NMR	1.13
22	hsIgFLNc22	2dlg	NMR	1.25
23	hsIgFLNc23	2d7q	NMR	1.45
24	hsIgFLNc24	1v05	X-RAY	1.04

Supplementary Table S2.

Ig domain Pairs	Linker length	Short linker
1-2	4	³⁷⁵ NKSQ ³⁷⁹
2-3	2	⁴⁷⁶ GA ⁴⁷⁷
3-4	2	⁵⁷² TE ⁵⁷³
4-5	6	⁶⁶⁴ RDAPQD ⁶⁶⁹
5-6	3	⁷⁶⁴ GAG ⁷⁶⁶
6-7	3	⁸⁶⁷ EPS ⁸⁶⁹
7-8	3	⁹⁶⁶ SPS ⁹⁶⁹
8-9	3	¹⁰⁶² VAP ¹⁰⁶⁴
9-10	3	¹¹⁵⁵ VPC ¹¹⁵⁷
10-11	3	¹²⁵⁰ EPA ¹²⁵³
11-12	3	¹³⁵⁰ TEG ¹³⁵²
12-13	3	¹⁴⁴³ HDV ¹⁴⁴⁵
13-14	3	¹⁵⁴⁰ LPT ¹⁵⁴²
14-15	3	¹⁶³⁷ VPT ¹⁶³⁸
16-17	2	¹⁸⁶¹ DY ¹⁸⁶²
18-19	4	²⁰³⁹ SQSE ²⁰⁴²
20-21	5	²²³¹ GPLGE ²²³⁵
21-22	5	²²²⁶ ASPSG ²³³⁰

Supplementary Table S3.

Protein name	ID
Human filamin A	NP_001447.2
<i>Canis familiaris</i> filamin A	ENSCAFT00000031079
<i>Mus musculus</i> filamin A	ENSMUST00000052678
<i>Rattus norvegicus</i> filamin A	ENSRNOT00000012796
<i>Xenopus tropicalis</i> filamin A	ENSXETT00000031243
<i>Danio rerio</i> filamin A	ENSDART00000046261
<i>Tetraodon nigroviridis</i> filamin A	GSTENT00008691001
<i>Takifugu rubripes</i> filamin A	NEWSINFRUT00000159596
Human filamin B	NP_001448.2
<i>Canis familiaris</i> filamin B	ENSCAFT00000012236
<i>Mus musculus</i> filamin B	ENSMUST00000052678
<i>Rattus norvegicus</i> filamin B	ENSRNOT00000012796
<i>Xenopus tropicalis</i> filamin B	ENSXETT00000018970
<i>Danio rerio</i> filamin B	ENSDART00000086028
<i>Tetraodon nigroviridis</i> filamin B	GSTENT00012499001
<i>Takifugu rubripes</i> filamin B	NEWSINFRUT00000147431
Human filamin C	NP_001449.3
<i>Canis familiaris</i> filamin C	ENSCAFT00000002504
<i>Mus musculus</i> filamin C	ENSMUST00000065090
<i>Rattus norvegicus</i> filamin C	XP_342654.2
<i>Xenopus tropicalis</i> filamin C	ENSXETT00000032738
<i>Danio rerio</i> filamin C	ENSDART00000073442
<i>Tetraodon nigroviridis</i> filamin C	GSTENT00029754001
<i>Takifugu rubripes</i> filamin C	NEWSINFRUT00000181322
<i>Ciona intestinalis</i> , <i>Ciona savignyi</i>	ENSCINT00000006732 (42,48)/ ENSCING00000003272 (gene), ENSCSAVT00000017772/ ENSCSAVG00000010354(gene)
<i>Strongylocentrotus purpuratus</i>	XP_784715.2
<i>Anopheles gambiae</i>	AGAP004335-PA

Supplementary Table S4.

FLN domain	Isoform	Range (Filamin A)	Residues
common ancestral residues			
ABD			
CH1	97	40-151	112
CH2	75	168-264	97
Ig-like			
1	62	279-374	96
2	52	379-475	97
3	67	478-71	94
4	90	574-663	90
5	45	670-763	94
6	67	767-866	100
7	64	870-965	96
8	46	969-1061	93
9	67	1065-1154	90
10	46	1158-1249	92
11	48	1253-1349	97
12	62	1352-1442	91
13	55	1446-1539	94
14	67	1543-1636	94
15	56	1640-1744	105
16	53	1771-1860	90
17	66	1863-1952	90
18	44	1968-2039	72
19	64	2044-2134	91
20	48	2159-2230	72
21	77	2236-2325	90
22	72	2331-2420	90
23	62	2427-2516	90
24	62	2553-2646	94

Supplementary Table S5.

Isoform	Domain	Ancestral Conservation ¹	Distinctive Modifications ² (teleostei period)
A	CH1	High <i>c</i>	Low
	CH2	High	High
	12	High	High
	14	High	No -
	17	High	High
	18	Low	High +
	21	High <i>i</i>	Low
B	CH1	High <i>c</i>	Low -
	CH2	High	High
	6	High	High
	12	High	High
	15	Low	High
	16	Low	High
	17	High	High
	18	Low	High +
	19	High	High
	20	High	Low -
	21	High <i>i</i>	Low
	22	High	Low
	23	High	Low
C	CH1	High <i>c</i>	Low
	CH2	High	Low
	6	High	Low
	9	High	No -
	11	High	High +
	12	High	Low
	16	High	High +
	17	High	No -
	19	High	Low
	21	High <i>i</i>	No -

¹ High/Low ancestral conservation implies greater than or less than 0.6 percent ancestral.

² High/Low distinctive residue implies per domain (A: > 6 / < 2 B: > 8 / < 4 C: > 6 / < 3).

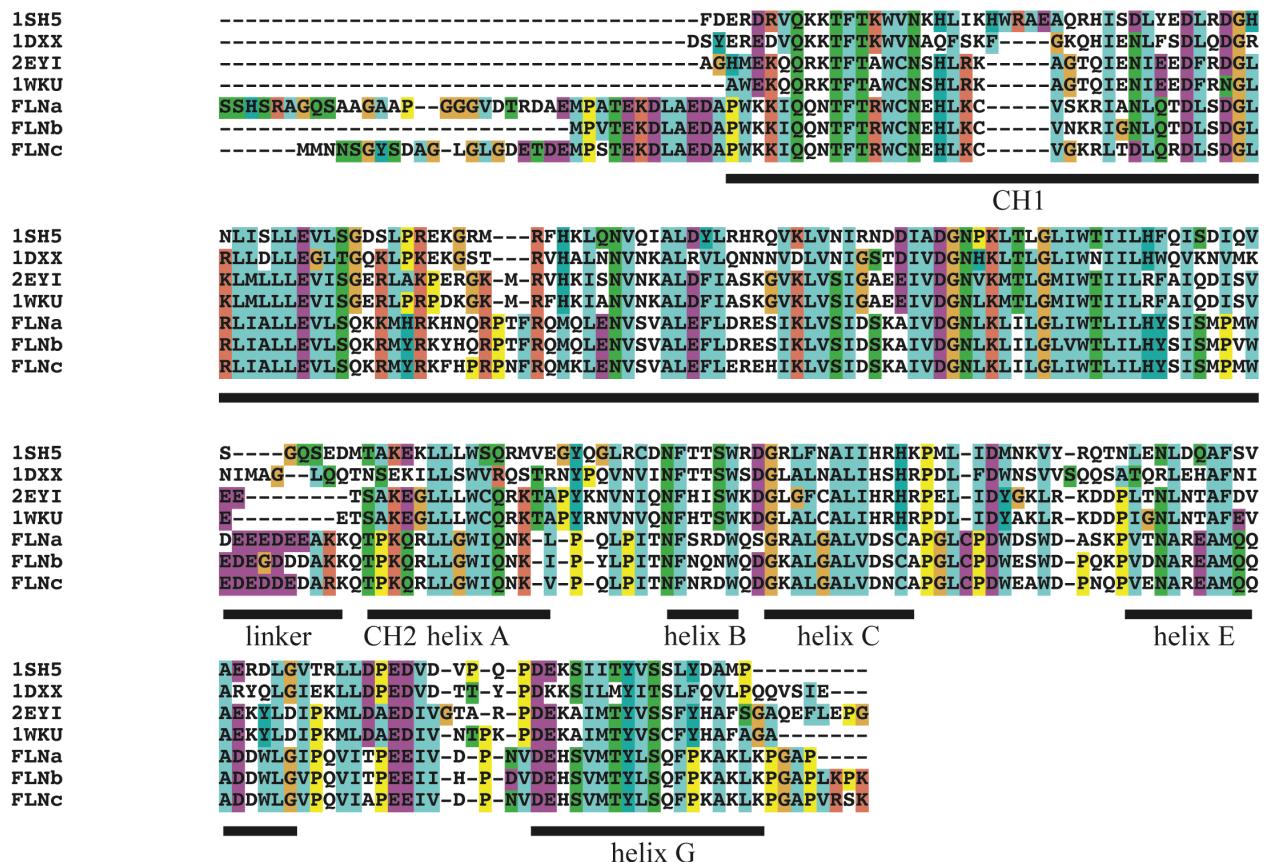
c, i Most Ancestral Residues for Isoform (*c* = CH domain, *I* = Ig-like domain)

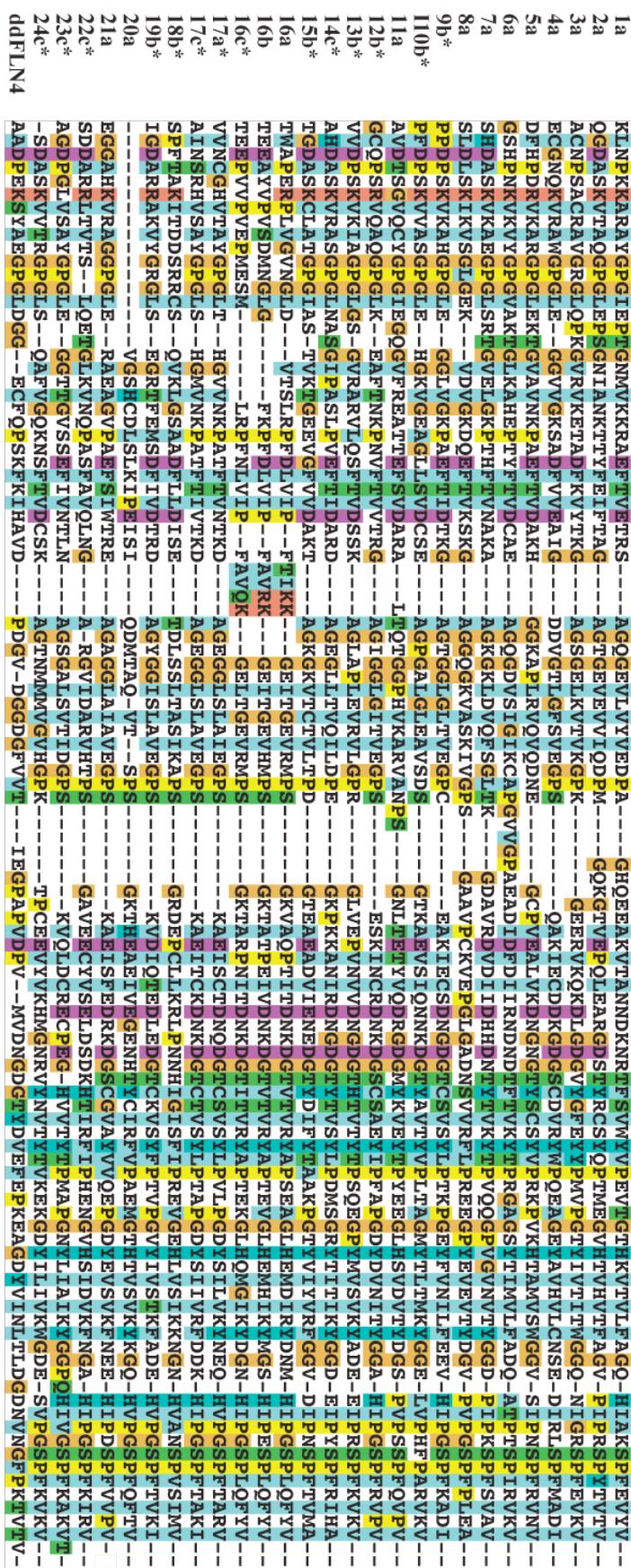
+- Most/Least Distinctive Residues.

consecutive **distinctive/ancestral**

Supplementary Figure S1.

A



B

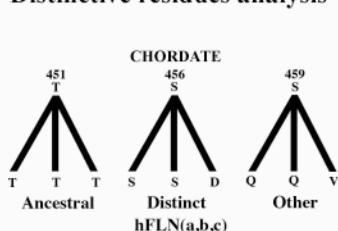
Supplementary Figure S2.

A Ancestral reconstruction

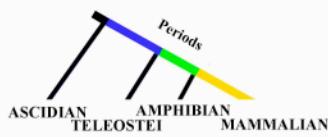
A. GAMBIA TVETFSAGKGNVDVAV
 S. PURPURATUS TVETFGAGSGNVEVL
 CIONA TVETVSAGSGEVLVYL
 H. SAPIEN TVETRSAGQGEVLVYY
 X. TROPICALIS TVETISAGMGEVLVYY
 D. RERIO TVETISAGQGEVLVYY
 H. SAPIEN TVDTISAGQGDVMVFV
 X. TROPICALIS YIVQIISIWSSSVLNLV
 D. RERIO TVDTFSAGQGQVMVYY
 H. SAPIEN TVQTVDAVGGEVLVYYI
 X. TROPICALIS TVETINAGLGEVLVFV
 D. RERIO TVETLEAGLGEVIVYYV

FLN_a FLN_b FLN_c
 CHORDATE TVETXSAGXGEVLVYY
 MAMMALAIN TVETRSAGQGEVLVYY
 AMPHIBIAN TVETXSAGXGEVLVYY
 TELEOSTEI TVETISAGXGEVLVYY
 MAMMALIAN TVDTISAGQGDVMVFV
 AMPHIBIAN TVXTISAGSGXVLVLV
 TELEOSTEI TVDTFSAGQGXVMVYY
 MAMMALIAN TVQTVDAVGGEVLVYYI
 AMPHIBIAN TVETVDAGVGEVLVYYI
 TELEOSTEI TVETVEAGLGEVLVYYV

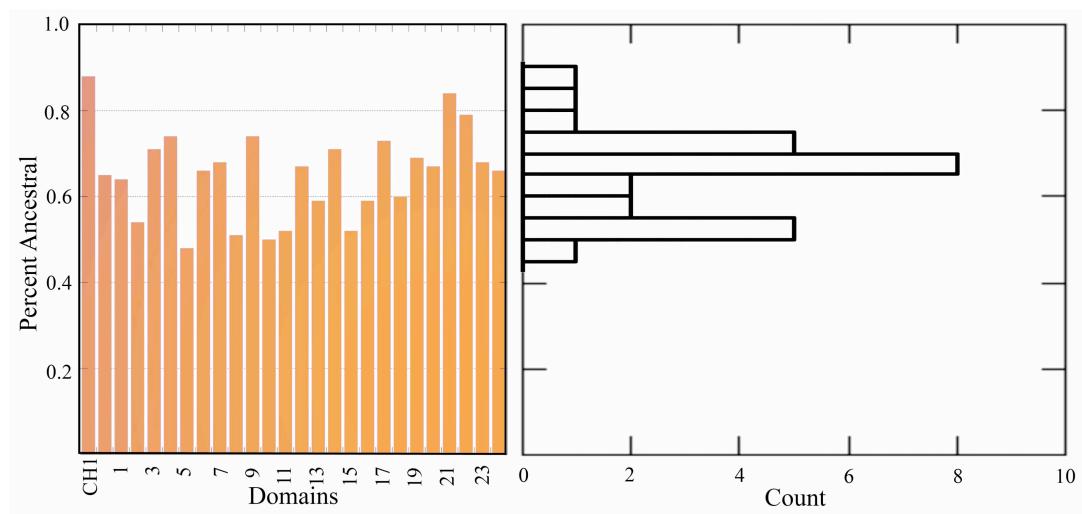
B Distinctive residues analysis



C Period of divergence



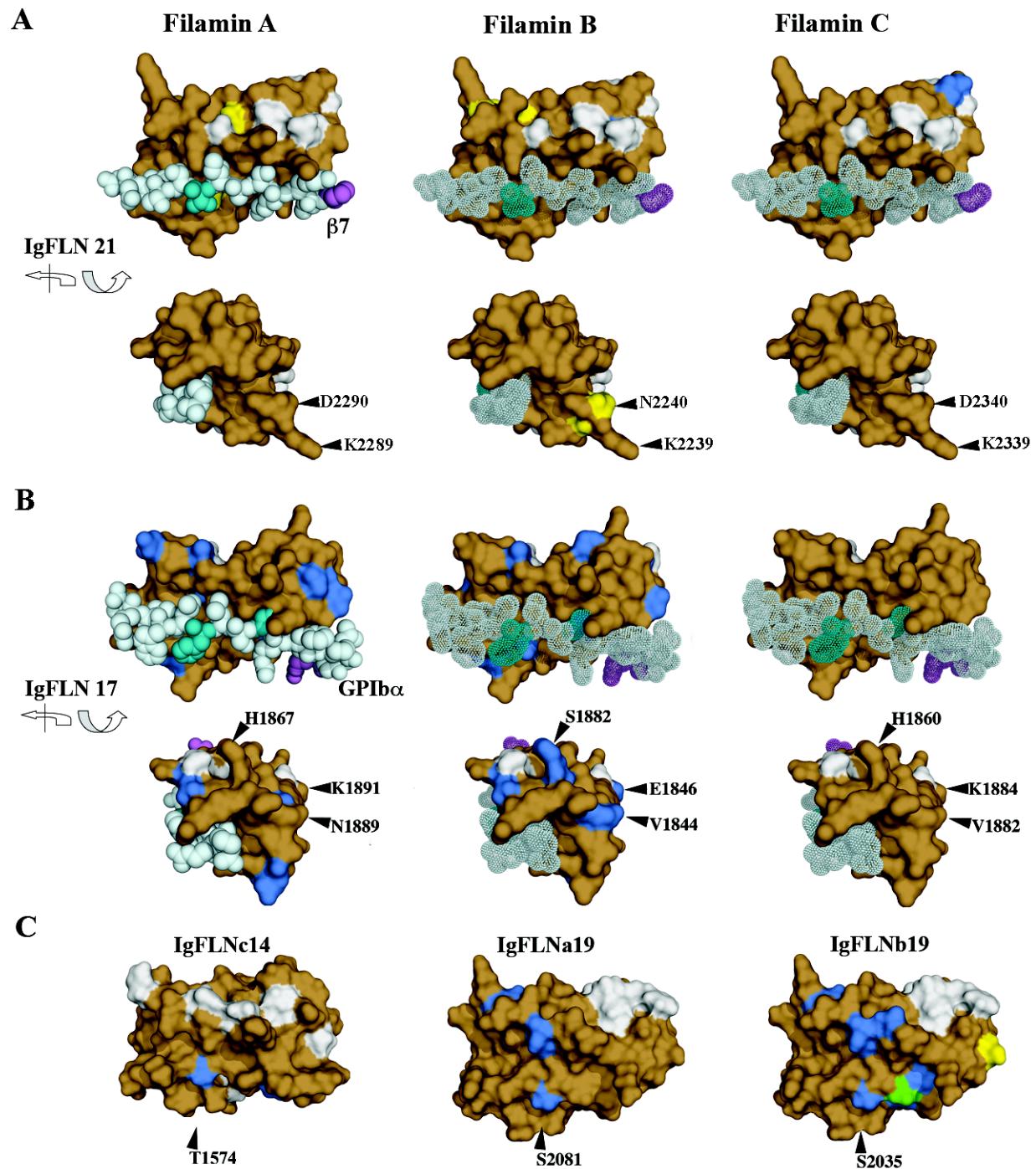
Supplementary Figure S3.



Supplementary Figure S4. Distinctive divergence outside the binding interface of Ig-like domains 17 and 21. Coloring (sand = ancestral, class-distinctive period of divergence (blue = Teleostei, green = Amphibian, yellow = Mammalian), white = other) **A.** IgFLN 21 in complex with β 7 integrin (PDB ID 2BRQ). Top row has the same orientation as Figure 4. Row 2, 180° rotation, views the CD face of the domain. β strand D is opposite ligand. All isoforms have class-distinctive residues (yellow) surrounded by highly ancestral regions outside the binding pocket. **B.** IgFLN 17 in complex with GPIba ligand residues (PDB ID 2BP3). Similarly, domain 17 has distinctive residues outside the binding pocket surrounded by ancestral residues (Filamin A and B only) **C.** Domains 14 and 19 have a profile of distinctive residues along β strand C that is similar to both IgFLN 17 and IgFLN 21.

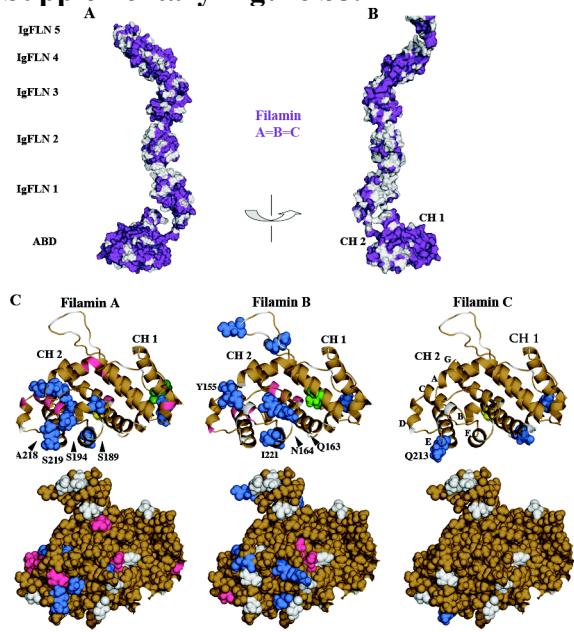
Supplementary Figure S5. Spatial and temporal distribution of divergence in the actin-binding domain. Coloring is the same as in Supplementary Figure S4. In addition (top row) violet is ancestral and similar among isoforms; bottom row (pink) shows mutations leading to disease. **A.** Surface representation of the actin-binding domain (ABD) and filamin Ig domains 1-5. Surface is highly ancestral. **B.** 180° rotation of A about the vertical axis. Domains 1 and 2 have an asymmetric conservation pattern. **C.** The ABD subdomains, calponin homology (CH) domains 1 and 2. CH domains contain seven helices labeled A-G on filamin C. Top is cartoon, bottom is space-filled model. Most of the distinctive residues in the ABD are localized to one region of CH 2, composed of helices B, C, E and F. Class-distinctive residues are in the vicinity of disease-causing mutations. Filamin A class-distinctive substitutions are mostly serine residues, while filamin B favors large bulky residues. Filamin C has fewer class-distinctive residues and is more ancestral than filamin A and B.

Supplementary Figure S4.



KEY: Ancestral, Distinctive (Teleostei, Amphibian, Mammalian)

Supplementary Figure S5.

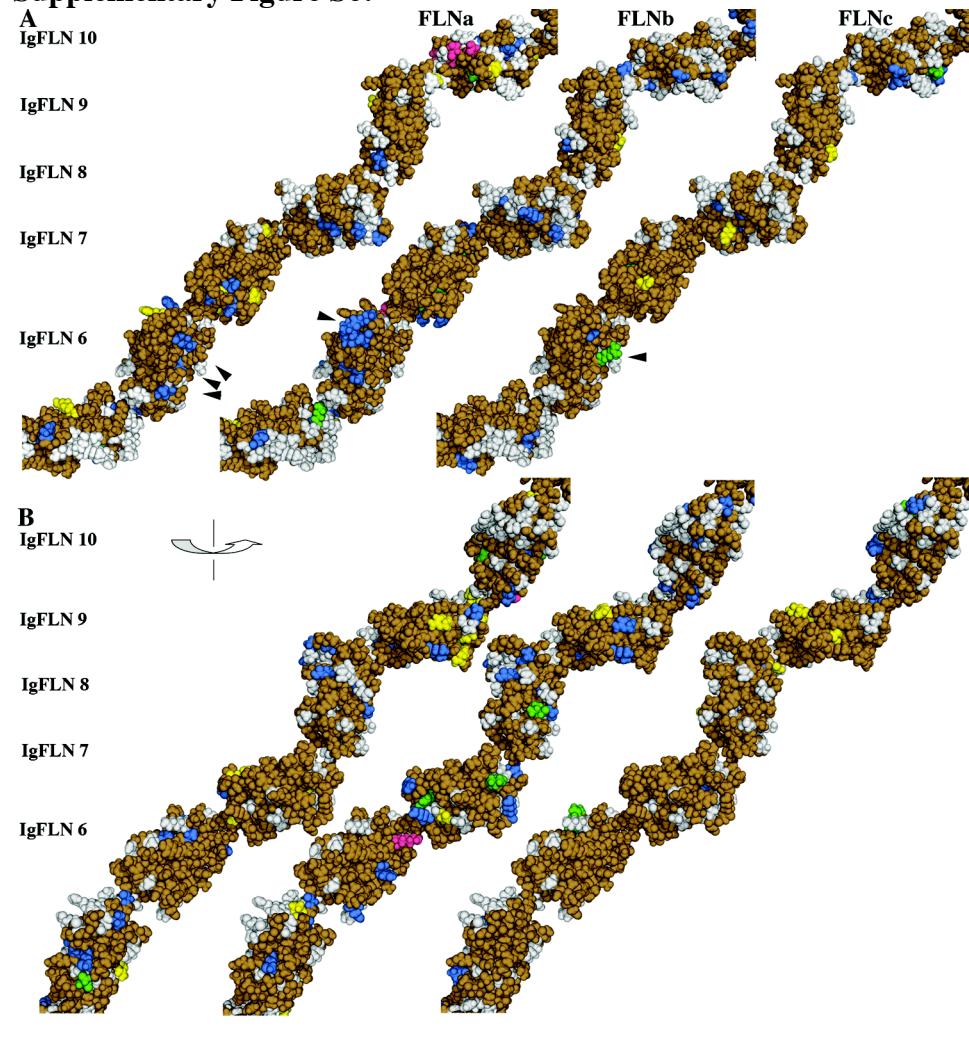


KEY: Ancestral (brown), Distinctive (Teleostei, Amphibian, Mammalian), Disease (pink)

Supplementary Figure S6. Ancestral and class-distinctive residues of Ig-like filamin

domains 5-10. Space-filling model of Ig-like domains 5-10 with coloring the same as Supplementary Figure S5. **A.** Some domains are highly conserved (e.g., IgFLN 7), and some are highly divergent in each isoform (e.g., IgFLN 6). Arrows pointing to IgFLN 6 point to the clustering of class-distinctive residues, including to an entire surface containing the CD loop of filamin B. In general, filamin C is more ancestral but nearly as variable as the other two isoforms. Unitary Mammalian period class-distinctive residues appear on many domains, although clusters of Mammalian period class-distinctive residues exist (e.g., IgFLNa9). **B.** A 180° rotation of A. This side of these domains appears to have more class-distinctive residues than the other side, especially in filamin B. Missense mutations leading to skeletal disorders are surface-exposed and close to the ends of domains. Often these localize to the vicinity of distinctive residues, even across domains (e.g., IgFLNb6 and IgFLNb7) and isoforms (e.g., IgFLNa10 and IgFLNb10).

Supplementary Figure S6.

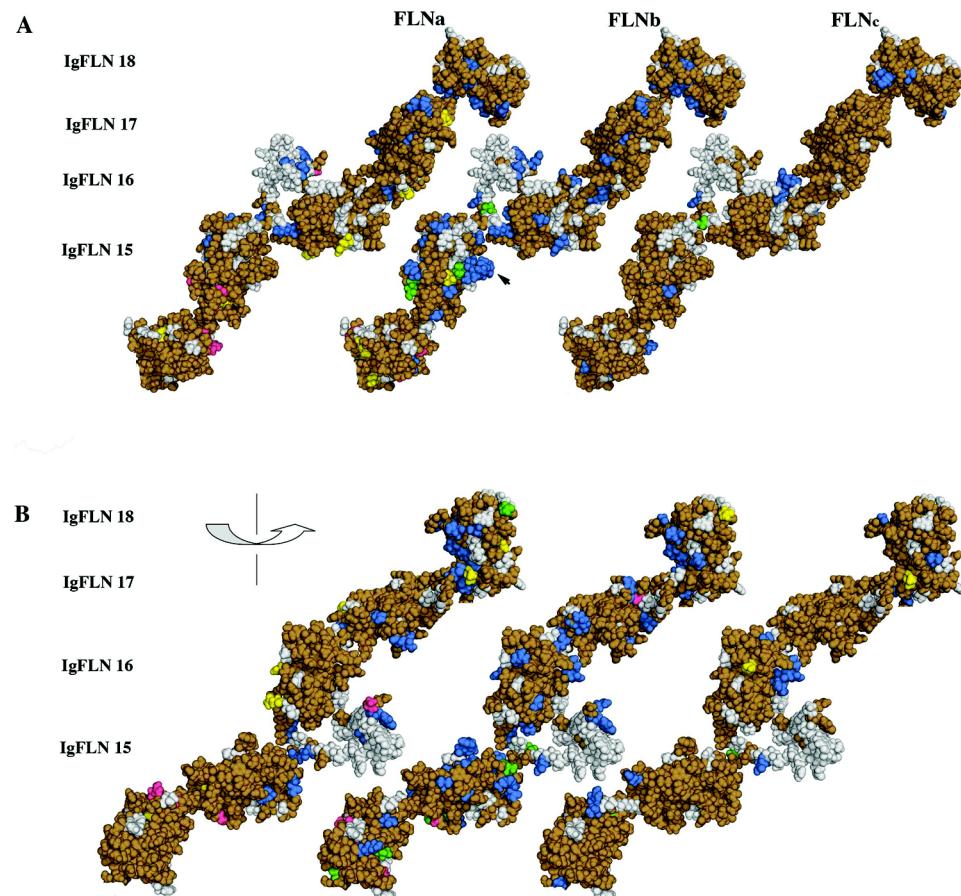


KEY: Ancestral, Distinctive (Teleostei, Amphibian, Mammalian), Disease

Supplementary Figure S7. Ancestral and distinctive residues of Ig-like filamin domains 14-

18. A space filling model of Ig-like domains 14-18 of filamin A, B and C with coloring same as Supplementary Figure S5. **A.** Filamin A is similarly conserved as filamin C, except for domain 17 and 18 where filamin A and B have similarly localized clusters of class-distinctive residues. The distinctive region of domain 18 includes the residues that would make up the first β strand of domain 18, however in our homology model these residues form a loop. Filamin B domain 15 is highly divergent (arrow). The loop between β -strand A and B on IgFLN 15 contains a distinctive deletion in filamin B of 8 amino acids that contain a putative GSK-3 phosphorylation site in both filamin A and C. Other non-deletion class-distinctive sites surround the distinctive deletion. **B.** A 180° rotation of A. Filamin B on this surface of appears to have more class-distinctive residues than that of filamin A and C, although there are more recent distinctive residues in both A and C. For example, filamin A domain 16 has Mammalian period distinctive residues that are homologous with filamin B Teleostei period class-distinctive residues.

Supplementary Figure S7.



KEY:Ancestral, Distinctive (Teleostei, Amphibian, Mammalian), Disease

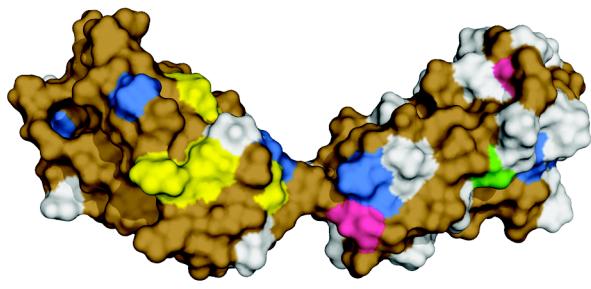
KEY:Ancestral, Distinctive (Teleostei, Amphibian, Mammalian), Disease

Supplementary Figure S8. Distinctive residues associated with mutations causing disease.

Coloring same as Supplementary Figure S5. **A.** IgFLNa9 and IgFLNa10 are surface models showing localization of distinctive and ancestral residues and residues with disease-causing mutations. **B.** IgFLNb14 also shows similar co-localization

Supplementary Figure S8.

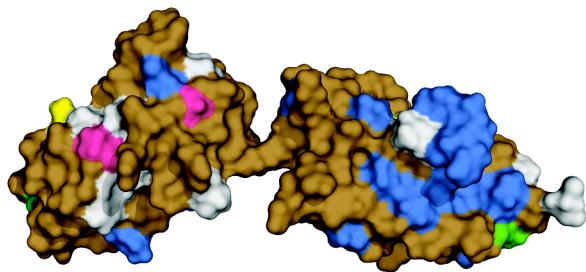
A



IgFLNa9

IgFLNa10

B



IgFLNb14

IgFLNb15

KEY: Ancestral, Distinctive (Teleostei, Amphibian, Mammalian), Disease