Supporting Information

Ivarsson et al. 10.1073/pnas.1312296111

SI Methods

Protein Purifications. Overnight cultures of GST and GST-PDZ (protein-95/disks large/zonula occludens-1) fusion proteins were used to inoculate 50 mL autoinducing MagicMedia (Invitrogen) supplemented with 25 μg/mL kanamycin and were grown for 24 h at 37 °C with shaking. The bacteria were pelleted (8,000 rpm, 10 min), stored overnight at 20 °C, and purified using glutathione affinity resin (GE Healthcare). The coding regions were as described in Tonikian et al. (1).

Phage Selections. In brief, proteins were coated in 96-well Maxisorp microtiter plates (NUNC) overnight (15 μg/mL protein in 100 μL PBS per well). For the first two rounds of selection, three wells were used for each library, whereas a single well was used for the following rounds. Parallel plates were coated with GST alone to remove nonspecific binders by preselection. The next day, wells were blocked with BSA for 2 h with blocking buffer (PBS, 0.2% BSA). Phage pools representing the naïve peptide library were diluted 10-fold in PBS, precipitated with polyethylene glycol–NaCl [4% PEG-800 (wt/vol) and 0.5 M NaCl] and resuspended to a final concentration of 10^{12} cfu/mL in PBT. In the first selection round, 100 μL of the phage pool representing the naïve peptide library was added to each well of the preselection plate, incubated for 1 h, transferred to the target plate, and incubated for an additional 2 h. The plate was washed four times with cold wash buffer (PBS, 0.5% Tween-20) and bound phage was eluted by direct infection into bacteria by the addition of 100 μL of log-phase *Escherichia coli* SS230 (A_{600} = 0.8) in 2YT to each well and incubation for 30 min at 37 °C with shaking. M13K07 helper phage (New England Biolabs) was added to a final concentration of 10^{10} phage per milliliter to enable phage production, and the cultures were incubated for 45 min at 37 °C with shaking. The cultures were transferred to 20 mL of 2YT supplemented with kanamycin (25 μg/mL), carbenicillin (100 μg/mL), and isfopropyl-β-D-thiogalactopyranoside (IPTG; 0.4 mM), and shaken overnight at 37 °C. The bacteria was pelleted by centrifugation (10 min, 17,090 × *g*), the supernatant transferred to a new tube, and phage particles were precipitated by addition of one-fifth volume of polyethylene glycol·NaCl, incubated at 4 °C for 5 min, and centrifuged at 28,880 × *g* for 20 min. The supernatant was removed and the phage pellet was resuspended in 2 mL of PBT and then used for the next round of selection. The selections were carried out for five rounds and the progress followed by analyzing aliquots of phage supernatants in a phage ELISA (2).

The phage pools of rounds three to five and the naïve phage libraries were barcoded for Illumina sequencing as outlined by McLaughlin and Sidhu (3). Briefly, undiluted amplified phage pools (5 μL) were used as templates for 24 cycles of 50 μL PCR reactions using unique combinations of barcoded primers for each reaction (0.5 μM each; for sequences of amplicon and barcodes see ref. 3) and using Phusion High Fidelity DNA polymerase (New England Biolabs) using maximum polymerase and primer concentrations. The PCR products were confirmed by gel electrophoresis (2% agarose gel) of 1 μL of PCR products.

The amount of the DNA amplicons was normalized by PEG/NaCl precipitation in a 96-well plate using a limiting amount of Ampure XP magnetic beads (Beckman Coulter). The magnetic beads were diluted 16-fold in PEG/NaCl and 100 μL of this solution was mixed by pipetting with 40 μL PCR product, incubated at room temperature for 20 min and then on a magnetic plate for 5 min to collect the beads. The supernatant was removed and the beads were washed twice with 70% EtOH, dried for 20 min at room temperature, and eluted by addition of 20 μL TE buffer (10 mM Tris, pH 8.0, and 0.1 mM EDTA). The normalized PCR amplicons were pooled (15 μL per reaction) and concentrated using two columns of a QIAquick PCR purification kit. The pooled amplicons were run on 2% agarose gel (80 V for 30 min), excised, and purified on a column of a QIAquick gel extraction kit using a modified protocol that uses extended incubation at room temperature instead of heating in Buffer QG (4). The bound DNA was eluted with 30 μL TE buffer. The concentration of the DNA was estimated picogreen dye as previously described. The PCR amplicons (~3 mg) were sent to Cofactor Genomics (Saint Louis, MO) for deep sequencing (Illumina Miseq; paired end 150 base reads, 20% PhiX). The obtained sequencing reads were filtered by discarding reads with an average PHRED quality score <35 (99.95% sequencing accuracy) or having a minimal nucleotide position score lower than 26.

Analysis of the Naïve Libraries. The quality of the proteomic peptide-phage display (ProP-PD) libraries were assessed from the deep-sequencing data by estimating the percent of starting templates, point mutations, and frame-shift mutations. The frequency of point mutations was estimated by assigning for each mutated sequence the most similar peptide sequence in the library design (denoted as parental sequence) and counting the amount of mutations as differences between the parental and mutated sequences on the DNA level. Frame-shifts were detected by aligning the DNA sequence of each mutated sequence to all sequences in the library design on DNA level using the Smith Waterman implementation provided by JAligner (parameters: identity matrix, gap opening penalty −5, gap extension −1).

Processing of Data from the Selections. The sequencing data contains selected wild-type parental peptides as well as mutant versions thereof (Fig. S6). To retrieve relevant peptides, we filtered the data for peptides occurring in the original library designs. To remove the noise we plotted histograms of the peptide frequencies (after matching to the actual library design) and manually assigned cut-off values after the prominent peak representing spurious binders after visual inspection. To focus on relevant peptides from the human ProP-PD, we subdivided the library entries into three groups based on the data available in April 2013 into a "high interest" set of true C termini comprising sequences that are in addition to either RefSeq from 2010 or Ensembl62 [also contained in one of either RefSeq or Uniprot in their 2013 versions (excluding sequences annotated as fragmentary)], a "proteolytic set" with an experimental support for a cleavage event listed in the TopFind database, and a "low interest" set with Ensembl62 entries not matching the two other sets. We filtered for peptides found in the high interest set (Table S1) and list identified targets from the low and medium interest sets in Table S3. To obtain viral targets of interest from the deep-sequencing data (Table S2) we assigned cut-off values to remove nonspecific peptides and filtered the data by removing three hits that did not originate from viruses targeting higher eukaryotes.

Comparison with Conventional Phage Display. Position weight matrices (PWMs) were generated using the MUSI software (5) with standard settings and without realignment of the C termini. For comparison between human targets predicted using conventional

phage display, a set of 7mer and of 10mer PWMs were calculated from the Tonikian et al. (1) data using MUSI. To compare the hydrophobicity of the retrieved ligands we calculated for the heptamer PWMs (from ProP-PD and randomized phage display, respectively), an accumulated hydrophobicity value as the sum of each amino acid hydrophobicity weight multiplied by each amino acid normalized frequency in the PWM matrix over each position (6) (Fig. S2). To compare if ProP-PD ligands would have been predicted by conventional methods, we used the 10mer PWMs based on Tonikian's data to scan a human library equivalent to the high interest set of our design using MOTIPs (7) and ranked the target peptides from 1 and up. Sequences with identical scores were ranked equally.

Peptide Synthesis. Peptides (Table 1) were synthesized using a Multipep synthesizer (Intavis AG Bioanalytical Instruments) on Wang resins (*p*-benzyloxybenzyl alcohol resin; AnaSpec) using 9-Fluorenyl methoxycarbonyl chemistry, with longer incubation or multiple cycles to conjugate the first C-terminal amino acid in the presence of 4-Dimethylaminopyridine (Sigma Aldrich). *N*-hydroxysuccinimide fluorescein (Pierce) was used to tag the N termini of the peptides with a fluorescent label. A 6-aminohexanoic acid moiety (AnaSpec) was used as a linker to separate the peptide from the fluorescein label to mitigate potential steric hinderance of protein–peptide interactions.

Fluorescence Polarization Assays. Binding affinities of PDZ domains for fluorescein-labeled peptides were determined using a 2103 Multilabel Reader (PerkinElmer). Briefly, fluorescein-labeled peptides were diluted to a final concentration of 2–5 nM and incubated with increasing concentrations of hexaHis-tagged-PDZ domains (0–100 μM; 12 datapoints), using duplicate protein titrations in 384-well Corning plates. After mixing on a shaking platform for 2 min at 500 rpm and centrifuging for 2 min at 1,000 × *g*, the fluorescence polarization signals from the wells were measured. The data were analyzed using the Graphpad Prism software and K_D values were determined by curve fitting the data to a single binding-site model.

Cloning. Full-length Scribble, mitogen-activated kinase 12 (MK12), guanylate cyclase soluble subunit α-2 (GCAY2) constructs were generated by Gateway cloning (Invitrogen) from entry clones in pDONR223 and shuttled into pcDNA5 FRT/TO with either an N terminus GFP or 3xFlag tag. CTNB1 was PCR-amplified and cloned into pCMV2B (Stratagene) that contains a Flag-tag sequence at the N terminus. PKP4 was PCR-amplified and cloned into the Creator vector 3xFlag N terminus expression vector using the Creator recombination system (8).

Cell Line. HEK293T cells were maintained in DMEM (ATCC) supplemented with 10% FBS and 1% pen/strep/glutamine, and the appropriate selection antibiotics when required.

Coimmunoprecipitations. HEK293T cells were cotransfected with GFP-Scribble and Flag-tagged constructs (described above). Cells were lysed 48 h after transfections with radioimmune precipitation assay buffer [50 mM Tris·HCl, pH 7.4, 1% Nonidet P-40, 150 mM NaCl, 1 mM EDTA, 10 mM Na_3VO_4, 10 mM sodium pyrophosphate, 25 mM NaF, 1× protease inhibitor mixture (Sigma)] for 30 min at 4 °C and and coimmunoprecipited with a GFP specific antibody (Abcam), as described previously (9). The resulting immunocomplexes and whole-cell lysates were analyzed by Western blot using the antibodies indicated in Fig. 4*B*. Protein samples were separated on a NuPage Bis·Tris 10% SDS/PAGE gel (Invitrogen) and transferred to nitrocellulose or PVDF membranes. Transferred samples were immunoblotted with primary anti-Flag antibodies, followed by incubation with horseradish peroxidase-conjugated goat anti-rabbit secondary antibodies (Santa Cruz Biotechnology) and detected using enhanced chemiluminescence (GE Healthcare).

Immunofluorescence. HEK293T cells were cotransfected with GFP-Scribble and Flag-tagged target constructs. Forty-eight hours after transfection, cells were fixed with 100% methanol for 20 min. Anti-Flag antibodies (1:400 Sigma) were incubated at room temperature for 1 h. Z-stack images were captured at room temperature by the Leica DMI6000B confocal microscope with a Leica 20×/0.40 NA objective lens and a Hamamatsu EM-CCD digital camera (C9100-13), and imported into Volocity software. The imaging medium was PBS.

Supplemental Network Analysis. We created a protein–protein interaction network of the four PDZ-containing proteins with their 78 putative binding using Cytoscape (10). The disk large homolog 1 (DLG1) part of the network contains previously known interactions with anion transporters, potassium channels, and G protein-coupled receptors. Consistent with the role of DLG1 in neuronal signaling, there are also known interactions with proteins involved in neuronal transmission, such as the motor protein KIF1β (11) and the microtubule-binding protein CRIPT (12). Among the new ligands we predict for DLG1, we highlight the Ras association domain-containing protein 6 (RASSF6), which interacts with the mammalian Ste20-like kinases (MST1/2), which are core kinases of the Hippo pathway (13). The suggested interaction between DLG1 and RASSF6 may add to the growing list of links between the cell polarity proteins and the Hippo signaling pathway (14). In addition, our predicted interactions between DLG1 and the E3 ubiquitin ligases DCNL1, RNF12, and MARCH3 may suggest unexplored connections between the ubiquitin system and the DLGs. Overall, the putative ligands appear relevant to the functions of DLG1.

Consistent with previous studies and roles in cell polarity and adhesion, the network of the LAP proteins Densin-180, Erbin, and Scribble contains interactions with the catenin family members PKP4, δ-catenin, and ARVCF, proteins that are found at the adherens junctions where they are involved in cell polarity and motility, but are also found in the nucleus where they are involved in transcriptional regulation (15–17). We also confirmed the interaction between Scribble and ARGH7, which is involved in cell migration, attachment, and cell spreading (18), and suggest novel interactions with a set of organic anion transporters and potassium channels as well as some nuclear proteins involved in transcriptional regulation, such as ATD2B. Scribble is not known to localize to the nucleus but it cannot be excluded that the proteins interact under specific circumstances. For example, ATD2B has been detected in the cytoplasm in some cancer cells (19). Under normal conditions, however, it is possible that the ATD2B C terminus is recognized by other class I PDZ proteins, such as NHERF2, that shuttle between the cytoplasm and the nucleus (20).

1. Tonikian R, et al. (2008) A specificity map for the PDZ domain family. *PLoS Biol* 6(9): e239.
2. Rajan S, Sidhu SS (2012) Simplified synthetic antibody libraries. *Methods Enzymol* 502: 3–23.
3. McLaughlin ME, Sidhu SS (2013) Engineering and analysis of peptide-recognition domain specificities by phage display and deep sequencing. *Methods Enzymol* 523: 327–349.
4. Quail MA, et al. (2008) A large genome center's improvements to the Illumina sequencing system. *Nat Methods* 5(12):1005–1010.
5. Kim T, et al. (2012) MUSI: An integrated system for identifying multiple specificity from very large peptide or nucleic acid data sets. *Nucleic Acids Res* 40(6):e47.
6. Monera OD, Sereda TJ, Zhou NE, Kay CM, Hodges RS (1995) Relationship of sidechain hydrophobicity and alpha-helical propensity on the stability of the single-stranded amphipathic alpha-helix. *J Pept Sci* 1(5):319–329.

7. Lam HY, et al. (2010) MOTIPS: Automated motif analysis for predicting targets of modular protein domains. *BMC Bioinformatics* 11:243.
8. Colwill K, et al. (2006) Modification of the Creator recombination system for proteomics applications—Improved expression by addition of splice sites. *BMC Biotechnol* 6:13.
9. Mak AB, et al. (2010) A lentiviral functional proteomics approach identifies chromatin remodeling complexes important for the induction of pluripotency. *Mol Cell Proteomics* 9(5):811–823.
10. Cline MS, et al. (2007) Integration of biological networks and gene expression data using Cytoscape. *Nat Protoc* 2(10):2366–2382.
11. Mok H, et al. (2002) Association of the kinesin superfamily motor protein KIF1Balpha with postsynaptic density-95 (PSD-95), synapse-associated protein-97, and synaptic scaffolding molecule PSD-95/discs large/zona occludens-1 proteins. *J Neurosci* 22(13): 5253–5258.
12. Cai C, Coleman SK, Niemi K, Keinänen K (2002) Selective binding of synapse-associated protein 97 to GluR-A alpha-amino-5-hydroxy-3-methyl-4-isoxazole propionate receptor subunit is determined by a novel sequence motif. *J Biol Chem* 277(35):31484–31490.
13. Ikeda M, et al. (2009) Hippo pathway-dependent and -independent roles of RASSF6. *Sci Signal* 2(90):ra59.
14. Boggiano JC, Fehon RG (2012) Growth control by committee: Intercellular junctions, cell polarity, and the cytoskeleton regulate Hippo signaling. *Dev Cell* 22(4):695–702.
15. Laura RP, et al. (2002) The Erbin PDZ domain binds with high affinity and specificity to the carboxyl termini of delta-catenin and ARVCF. *J Biol Chem* 277(15):12906–12914.
16. Izawa I, et al. (2002) ERBIN associates with p0071, an armadillo protein, at cell-cell junctions of epithelial cells. *Genes Cells* 7(5):475–485.
17. Zhang Y, et al. (2006) Convergent and divergent ligand specificity among PDZ domains of the LAP and zonula occludens (ZO) families. *J Biol Chem* 281(31):22299–22311.
18. Audebert S, et al. (2004) Mammalian Scribble forms a tight complex with the betaPIX exchange factor. *Curr Biol* 14(11):987–995.
19. Leachman NT, Brellier F, Ferralli J, Chiquet-Ehrismann R, Tucker RP (2010) ATAD2B is a phylogenetically conserved nuclear protein expressed during neuronal differentiation and tumorigenesis. *Dev Growth Differ* 52(9):747–755.
20. Thevenet L, et al. (2005) NHERF2/SIP-1 interacts with mouse SRY via a different mechanism than human SRY. *J Biol Chem* 280(46):38625–38630.

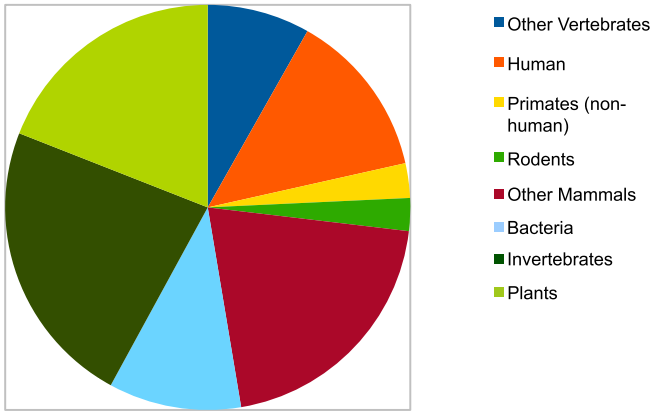

Fig. S1. Overview of the viral library design based on host organism.

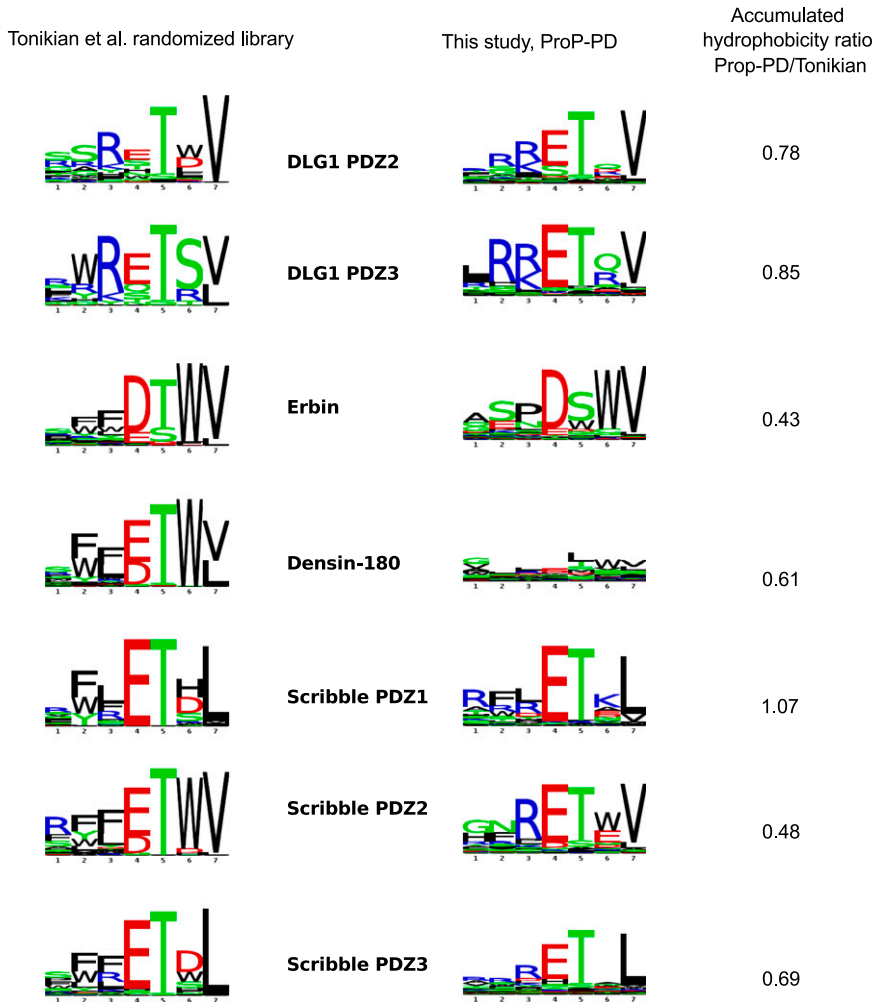

Fig. S2. Comparison between logos derived from ProP-PD and conventional peptide-phage display. On the left logos as derived from the Tonikian et al. (1) study using a combinatorial peptide-phage library, on the right logos derived from ProP-PD experiments. "Accumulated hydrophobicity ratio ProP-PD/Tonikian" gives the ratio of the accumulated and normalized hydrophobicity (see *Methods*). Values smaller than one indicate a more hydrophobic PWM for the data obtained from the conventional phage library.

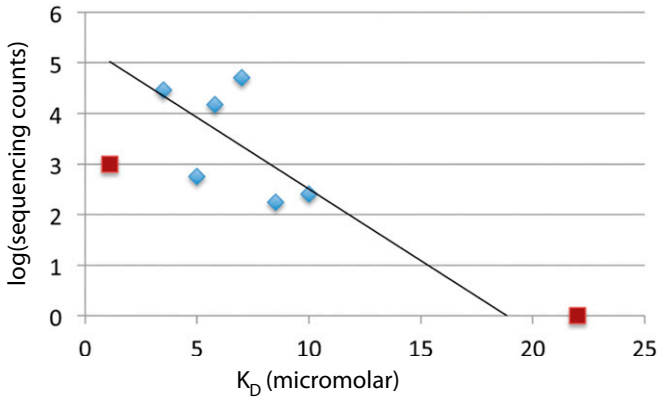

Fig. S3. Correlation between affinities and sequencing counts for Scribble PDZ3. Semilog scale plot of the sequencing counts versus affinities with a linear fit. The two red squares indicate outliers (the DNM1L peptide, to the left) and the GSPDSWV peptide (to the right).

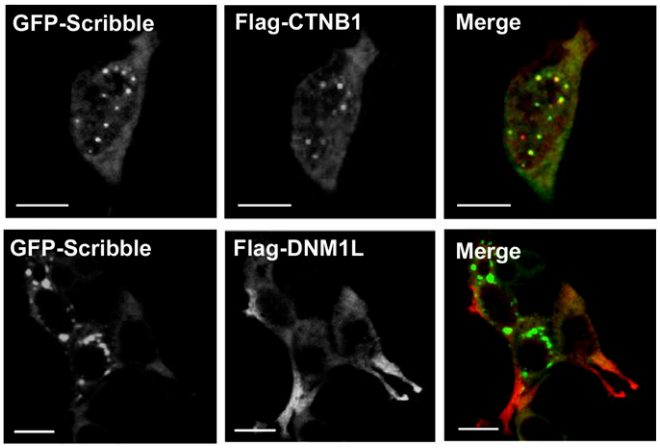

Fig. S4. Colocalization of Scribble with CTNB1 and DNM1L, respectively. (*A*) Colocalization of GFP-tagged full-length Scribble with Flag-tagged CTNB1 and DNM1L 48 h after cotransfection in HEK293T cells (confocal micrographs). (Scale bars, 15 μm.)

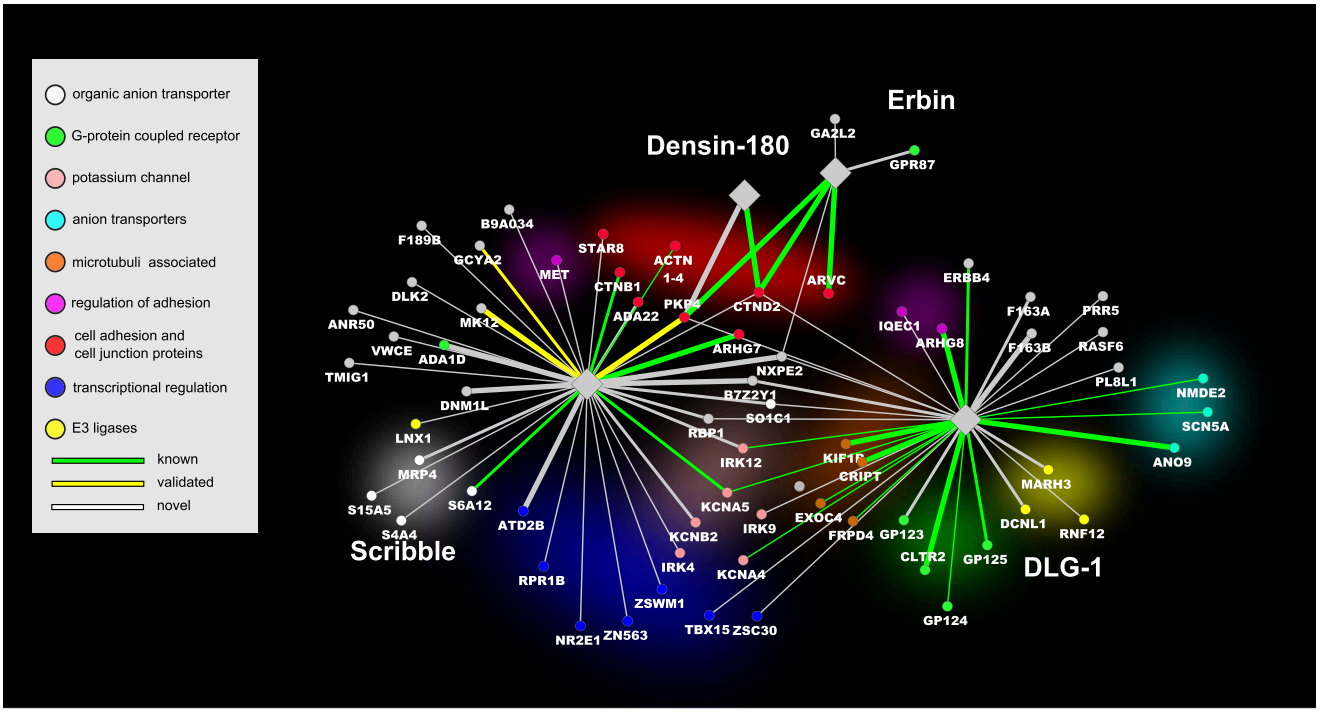

Fig. S5. Comprehensive network of identified interactions. The bait proteins (Densin-180 PDZ; Erbin PDZ; Scribble PDZ1, PDZ2, and PDZ3; and DLG1 PDZ1, PDZ2, and PDZ3) are indicated by gray diamonds. Ligands identified by ProP-PD experiments are indicated by circles, in which colors indicate their biological processes. The width of the connecting lines reflect the frequency of a ligand in the sequencing data, with the ligands divided into three categories: high [$\log_{10}$(counts) > 3], medium [3 > $\log_{10}$(counts) > 2], and low [$\log_{10}$(counts) < 2]. The color of the connecting branches indicates if the interactions are novel (gray), known (green), or here validated (yellow). The network was designed using the program Cytoscape (10).

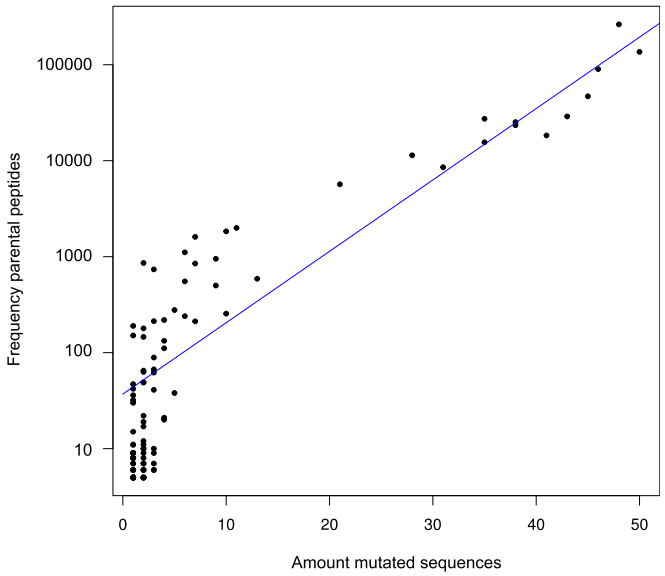

Fig. S6. Frequency of mutated peptides versus designed parental peptides after the fifth round of selection. The comparison between the amount of sequences with point mutations (*x* axis) to the frequency of their parental sequences ("Frequency parental peptides," *y* axis, log-scale) illustrates that the more selected a wild-type peptide is, the more mutants of it will accumulate during the phage propagation.

Table S1. Comprehensive list of selected targets for each domain with literature references when applicable

Protein	Peptide	Library A	Library B	Total	Uniprot	Source	PMID	Rank ProP-PD	Rank Tonikian
Scribble PDZ1	RFLETKL	67,514	45,351	112,865	B7Z2Y1_HUMAN			1	4
Scribble PDZ1	AWDETNL	3,009	1,552	4,561	ARHG7_HUMAN	Audebert et al. (1)	15182672	2	25
Scribble PDZ1	HMFETFL	286	5	291	ARHG8_HUMAN			3	15
Scribble PDZ1	TSRETDL	9	1	10	KCNA5_HUMAN	Zhang et al. (2)	16737968	4	186
Scribble PDZ1	RGEESTM	0	8	8	VWCE_HUMAN			5	12,927
Scribble PDZ1	IRETHLW	7	0	7	DNML1_HUMAN			6	41,469
Scribble PDZ2	GSPDSWV	0	618	618	PKP4_HUMAN			1	2,100
Scribble PDZ2	VQRHTWL	76	70	146	NXPE2_HUMAN			2	126
Scribble PDZ2	RFLETKL	35	41	76	B7Z2Y1_HUMAN			3	3
Scribble PDZ2	AWDETNL	29	29	58	ARHG7_HUMAN	Audebert et al. (1)	15182672	4	46
Scribble PDZ2	ASPDSWV	6	32	38	CTND2_HUMAN			5	523
Scribble PDZ2	PYEQVQL	20	1	21	ZSWM1_HUMAN			6	14,689
Scribble PDZ3	VQRHTWL	83,844	17,638	101,482	NXPE2 _HUMAN			1	55
Scribble PDZ3	AWDETNL	30,601	27,096	57,697	ARHG7_HUMAN	Audebert et al. (1)	15182672	2	322
Scribble PDZ3	RFLETKL	12,281	17,719	30,000	B7Z2Y1_HUMAN			3	13
Scribble PDZ3	HMFETFL	2	7,318	7,320	ATD2B_HUMAN			4	14
Scribble PDZ3	NLRETDI	784	494	1,278	ADA1D_HUMAN			5	112
Scribble PDZ3	IRETHLW	1	1,193	1,194	DNM1L_HUMAN			6	36,666
Scribble PDZ3	VSKETPL	782	359	1,141	MK12_HUMAN			7	179
Scribble PDZ3	PGKETQL	775	0	775	SO1C1_HUMAN			8	77
Scribble PDZ3	DRKETSI	302	0	302	RBP1_HUMAN			9	99
Scribble PDZ3	FLRETSL	1	252	253	GCYA2_HUMAN			10	1
Scribble PDZ3	RLWETSI	1	248	249	ADA22_HUMAN			11	1,230
Scribble PDZ3	PTRETSM	91	145	236	KCNB2_HUMAN			12	342
Scribble PDZ3	TIFETAL	203	12	215	MRP4_HUMAN			13	2
Scribble PDZ3	AWFDTDL	174	0	174	CTNB1_HUMAN	Xhang et al. (2)	16737968	14	228
Scribble PDZ3	GEKETHL	94	76	170	S6A12_HUMAN	Gfeller et al. (3)	21525870	15	23
Scribble PDZ3	YRRESEI	100	11	111	IRK12_HUMAN			16	3,906
Scribble PDZ3	TSRETDL	7	86	93	KCNA5_HUMAN	Xhang et al. (2)	16737968	17	5
Scribble PDZ3	ASFWETS	87	0	87	MET_HUMAN			18	30,860
Scribble PDZ3	LYGESDL	55	0	55	ACTN1,2,3,4_HUMAN	Xhang et al. (2)	16737968	19	877
Scribble PDZ3	YRRESAI	2	39	41	IRK4_HUMAN			20	42
Scribble PDZ3	PGKTTAL	0	39	39	DLK2_HUMAN			21	268
Scribble PDZ3	YKKETPL	30	1	31	ANR50_HUMAN			22	104
Scribble PDZ3	DLWETAL	3	26	29	S15A5_HUMAN			23	21
Scribble PDZ3	GSRETGL	0	26	26	F189B_HUMAN			24	58
Scribble PDZ3	GDLFSTD	0	21	21	RPR1B_HUMAN			25	9,794
Scribble PDZ3	PHSETAL	1	17	18	TMIG1_HUMAN			26	472
Scribble PDZ3	AGPETKL	0	15	15	STAR8_HUMAN			27	2,136
Scribble PDZ3	THWRETI	0	14	14	ZN563_HUMAN			28	48,049
Scribble PDZ3	KGTETTL	0	10	10	S4A4_HUMAN			29	352
Scribble PDZ3	MYKSSDI	0	8	8	NR2E1_HUMAN			30	577
Scribble PDZ3	SWPGTFL	0	7	7	LNX1_HUMAN			31	356
Densin-180	GSPDSWV	Failed	5,450	5,450	PKP4_HUMAN			1	170
Densin-180	ASPDSWV	Failed	2,794	2,794	CTND2_HUMAN	Izawa et al. (4)	11729199	2	35
Erbin	ASPDSWV	64,274	71,946	136,220	CTND2_HUMAN	Laura et al. (5)	11821434	1	2
Erbin	GSPDSWV	0	17,680	17,680	PKP4_HUMAN	Izawa et al. (6)	12047349	2	3
Erbin	QPVDSWV	1,747	88	1,835	ARVC_HUMAN	Laura et al. (5)	11821434	3	4
Erbin	YYDYTDV	114	2	116	GPR87_HUMAN			4	1,357
Erbin	VQRHTWL	21	1	22	NXPE2 _HUMAN			5	44
Erbin	PEEESWV	14	5	19	GA2L2_HUMAN			6	46
DLG1 PDZ1	RFLETKL	569	77	646	B7Z2Y1_HUMAN			1	721
DLG1 PDZ1	QMSVHMV	131	0	131	TBX15_HUMAN			2	1,230
DLG1 PDZ2	LRKETRV	58,320	29,164	87,484	CLTR2_HUMAN	Gfeller et al. (3)	21525870	1	24
DLG1 PDZ2	RSISTDV	14,224	24,273	38,497	F163B_HUMAN			2	2
DLG1 PDZ2	KRKETLV	21,909	11,998	33,907	ARHG8_HUMAN	Carr et al. (7)	19586902	3	3
DLG1 PDZ2	SARSTDV	4,643	16,686	21,329	ANO9_HUMAN	Gfeller et al. (3)	21525870	4	1
DLG1 PDZ2	AGRETTV	719	10,401	11,120	KIF1B_HUMAN	Mok et al. (8)	12097473	5	13
DLG1 PDZ2	NSKETVV	421	279	700	MARH3_HUMAN			6	5
DLG1 PDZ2	WKHETTV	483	90	573	GP125_HUMAN	Yamamoto et al. (9)	15021905	7	41
DLG1 PDZ2	RAISTDV	185	155	340	F163A_HUMAN			8	18
DLG1 PDZ2	WKNETTV	255	35	290	GP123_HUMAN			9	218
DLG1 PDZ2	GTKSTTV	0	269	269	DCNL1_HUMAN			10	43

Table S1. Cont.

Protein	Peptide	Library A	Library B	Total	Uniprot	Source	PMID	Rank ProP-PD	Rank Tonikian
DLG1 PDZ2	RHRNTVV	42	147	189	ERBB4_HUMAN	Huang et al. (10)	12175853	11	12
DLG1 PDZ2	MTKDTLV	60	18	78	PL8L1_HUMAN			12	42
DLG1 PDZ2	QRTHTRV	73	5	78	ZSC30_HUMAN			13	372
DLG1 PDZ2	SGISTIV	27	33	60	IQEC1_HUMAN			14	29
DLG1 PDZ2	GSPDSWV	0	54	54	PKP4_HUMAN	Izawa et al. (4)	12047349	15	314
DLG1 PDZ2	WKSETTV	42	5	47	GP124_HUMAN	Yamamoto et al. (9)	15021905	16	148
DLG1 PDZ2	GNRESVV	0	45	45	RNF12_HUMAN			17	17
DLG1 PDZ2	GGRQSVV	0	38	38	PRR5_HUMAN			18	16
DLG1 PDZ2	SSIESDV	18	18	36	NMDE2_HUMAN	Inanobe et al. (11)	11997254	19	28
DLG1 PDZ2	RDRESIV	19	16	35	SCN5A_HUMAN	Petitprez et al. (12)	21164104	20	19
DLG1 PDZ2	PGKETQL	2	29	31	SO1C1_HUMAN			21	108
DLG1 PDZ2	KIKETTV	15	12	27	FRPD4_HUMAN	Lee et al. (13)	19118189	22	38
DLG1 PDZ2	IKTETTV	12	12	24	RASF6_HUMAN			23	270
DLG1 PDZ2	DKKITTV	13	9	22	EXOC4_HUMAN	Bolis et al. (14)	19587293	24	451
DLG1 PDZ2	VQRHTWL	15	3	18	NXPE2_HUMAN			25	33
DLG1 PDZ2	DRKETSI	8	9	17	RBP1_HUMAN			26	70
DLG1 PDZ2	TSRETDL	13	0	13	KCNA5_HUMAN	Mathur et al. (15)	16466689	27	4
DLG1 PDZ2	KAVETDV	5	5	10	KCNA4_HUMAN	Kim et al. (16)	7477295	28	9
DLG1 PDZ2	YRRESEI	7	0	7	IRK12_HUMAN	Leonoudakis et al. (17)	14960569	29	267
DLG1 PDZ2	ASPDSWV	1	6	7	CTND2_HUMAN			30	429
DLG1 PDZ3	LRKETRV	146285	89,316	235,601	CLTR2_HUMAN	Gfeller et al. (3)	21525870	1	2
DLG1 PDZ3	NYKQTSV	287	622	909	CRIPT_HUMAN	Cai et al. (18)	12070168	2	1
DLG1 PDZ3	KRKETLV	18	4	22	ARHG8_HUMAN	Carr (6)	19586902	3	5

Library A and Library B, the sequencing counts for a given peptide from the replicate selection; Peptide, the selected C-terminal peptides; Protein, the identity of the bait PDZ domain; Rank ProP-PD, the rank of a peptide based on the selection (1: sequence with the highest total sequencing counts); Rank Tonikian PWM, predicted rank of a selected peptide using position specific scoring matrices based on the data of Tonikian et al. (18) among all sequences in the designed human ProP-PD library Reference, reference to a supporting publication with Pubmed id in PMID; Total, total sequencing counts of a given peptide; Uniprot, the Uniprot entry corresponding to a selected peptide.

1. Audebert S, et al. (2004) Mammalian Scribble forms a tight complex with the betaPIX exchange factor. *Curr Biol* 14(11):987–995.
2. Zhang Y, et al. (2006) Convergent and divergent ligand specificity among PDZ domains of the LAP and zonula occludens (ZO) families. *J Biol Chem* 281(31):22299–22311.
3. Gfeller D, et al. (2011) The multiple-specificity landscape of modular peptide recognition domains. *Mol Syst Biol* 7:484.
4. Izawa I, et al. (2002) Densin-180 interacts with delta-catenin/neural plakophilin-related armadillo repeat protein at synapses. *J Biol Chem* 277(7):5345–5350.
5. Laura RP, et al. (2002) The Erbin PDZ domain binds with high affinity and specificity to the carboxyl termini of delta-catenin and ARVCF. *J Biol Chem* 277(15):12906–12914.
6. Izawa I, et al. (2002) ERBIN associates with p0071, an armadillo protein, at cell-cell junctions of epithelial cells. *J Biol Chem* 277(15):12906–12914.
7. Carr HS, et al. (2009) Interaction of the RhoA exchange factor Net1 with discs large homolog 1 protects it from proteasome-mediated degradation and potentiates Net1 activity. *J Biol Chem* 284(36):24269–24280.
9. Yamamoto Y, et al. (2004) Direct binding of the human homologue of the Drosophila disc large tumor suppressor gene to seven-pass transmembrane proteins, tumor endothelial marker 5 (TEM5), and a novel TEM5-like protein. *Oncogene* 23(22):3889–3897.
10. Huang YZ, et al. (2002) Compartmentalized NRG signaling and PDZ domain-containing proteins in synapse structure and function. *Int J Dev Neurosci* 20(3-5):173–185.
11. Inanobe A, et al. (2002) Inward rectifier K+ channel Kir2.3 is localized at the postsynaptic membrane of excitatory synapses. *Am J Physiol Cell Physiol* 282(6):C1396–1403.
12. Petitprez S, et al. (2011) SAP97 and dystrophin macromolecular complexes determine two pools of cardiac sodium channels Nav1.5 in cardiomyocytes. *Circ Res* 108(3):294–304.
13. Lee HW, et al. (2008) Preso, a novel PSD-95-interacting FERM and PDZ domain protein that regulates dendritic spine morphogenesis. *J Neurosci* 28(53):14546–14556.
14. Bolis A, et al. (2009) Dlg1, Sec8, and Mtmr2 regulate membrane homeostasis in Schwann cell myelination. *J Neurosci* 29(27):8858–8870.
15. Mathur R, et al. (2006) A specific N-terminal residue in Kv1.5 is required for upregulation of the channel by SAP97. *Biochem Biophys Res Commun* 342(1):1–8.
16. Kim E, et al. (1995) Clustering of Shaker-type K+ channels by interaction with a family of membrane-associated guanylate kinases. *Nature* 378(6552):85–88.
17. Leonoudakis D, et al. (2004) A multiprotein trafficking complex composed of SAP97, CASK, Veli, and Mint1 is associated with inward rectifier Kir2 potassium channels. *J Biol Chem* 279(18):19051–19063.
18. Tonikian R, et al. (2008) A specificity map for the PDZ domain family. *PLoS Biol* 6(9):e239.

Table S2. Comprehensive list of identified viral targets stating interacting PDZ domains, peptide sequences, sequencing counts, and literature reference, when applicable

Protein	Peptide	Count A	Count B	Total	Uniprot	Name	Source (for the protein, not always the exact variant)	PMID
Scribble PDZ1	RRRETAL	x	21,586	21,586	VE6_HPV33	Human papillomavirus type 33 protein E6	Nakagawa and Huibregtse (1)	11027293
Scribble PDZ1	TRRETQL	x	175	175	VE6_HPV16	Human papillomavirus type 16 protein E6	Nakagawa and Huibregtse (1)	11027293
Scribble PDZ1	PDTDWLV	x	62	62	LRP2_HHV1F	Human herpesvirus 1 latency-related protein 2		
Scribble PDZ2	HFRETEV	3,827	49,903	53,730	TAX_HTL1F	Human T-cell leukemia virus 1 protein Tax-1	Arpin-André and Mesnard (2)	17855372
Scribble PDZ2	HFHETEV	26	474	500	TAX_HTL1L	Human T-cell leukemia virus 1 protein Tax-1	Arpin-André and Mesnard (2)	17855372
Scribble PDZ3	RRRETAL	118,682	69,771	188,453	VE6_HPV33	Human papillomavirus type 33 protein E6	Nakagawa and Huibregtse (1)	11027293
Scribble PDZ3	TRRETQL	28,733	29,867	58,600	VE6_HPV16	Human papillomavirus type 16 protein E6	Nakagawa and Huibregtse (1)	11027293
Scribble PDZ3	HFRETEV	110	589	699	TAX_HTL1F	Human T-cell leukemia virus 1 protein Tax-1	Arpin-André (2)	17855372
Scribble PDZ3	TRRETEV	81	130	211	VE6_HPV35	Human papillomavirus type 35 protein E6	Nakagawa and Huibregtse (1)	11027293
Scribble PDZ3	TRRETQV	39	145	184	VE6_HPV39	Human papillomavirus type 39 protein E6	Nakagawa and Huibregtse (1)	11027293
Scribble PDZ3	AIFSTDI	1	81	82	YVDA_VACCW	Vaccinia virus uncharacterized 9.2 kDa protein		
Scribble PDZ3	SGGETRL	2	75	77	VGLG_RABVV	Rabius virus glycoprotein G		
Scribble PDZ3	TGRSTTL	8	64	72	VFUS_SHEVK	Sheeppox virus putative fusion protein		
Scribble PDZ3	IRRETQV	24	41	65	VE6_HPV70	Human papillomavirus type 70, protein E6	Nakagawa and Huibregtse (1)	11027293
Scribble PDZ3	RRRETQV	2	61	63	VE6_HPV45	Human papillomavirus type 45, protein E6	Nakagawa and Huibregtse (1)	11027293
Scribble PDZ3	HFHETEV	6	39	45	TAX_HTL1L	Human T-cell leukemia virus 1 protein Tax-1	Okajima et al. (3)	18661220
Scribble PDZ3	PFSSSDL	4	15	19	GAG_MLVAB	Abelson murine leukemia virus Gag polyprotein		
Scribble PDZ3	LNYETNL	1	13	14	ENV_HTL3P	Human T-cell leukemia virus 3 envelope glycoprotein gp63		
Erbin	FPPEDWV	7	3,235	3,242	VEMP_BCHK4	Bat coronavirus Envelope small membrane protein		
Erbin	YPPEDWV	5	699	704	VEMP_BCHK5	Bat coronavirus Envelope small membrane protein		
Erbin	RRRETAL	4	407	411	VE6_HPV33	Human papillomavirus type 33 protein E6		
Erbin	DKLDNWV	2	319	321	VPU_HV1YB	HIV type 1 group N protein Vpu		
Erbin	IDQDNWV	4	220	224	VPU_SIVEK	Simina immunodeficiency virus protein Vpu		
Erbin	TRRETQL	0	168	168	VE6_HPV16	Human papillomavirus type 16 protein E6		
Erbin	ATCTFTL	2	106	108	VP23_ELHVK	Elephantid herpesvirus 1, triplex capsid protein U56		
Erbin	IRRETQV	6	92	98	VE6_HPV70	Human papillomavirus type 70, protein E6		
Erbin	ATHLINA	5	83	88	1102L_ASFWA	African swine fever virus protein MGF 110–2L		
Erbin	APSVLTV	2	80	82	NS3D_BCHK5	Bat coronavirus KHU5 nonstructural protein 3d		
Erbin	AVNFSTL	1	74	75	OBP_HHV2H	Human herpesvirus 2 replication origin-binding protein		
Erbin	HFRETEV	0	74	74	TAX_HTL1F	Human T-cell leukemia virus 1 protein Tax-1	Song et al. (4)	19472191
DLG1 PDZ1	IRRETQV	283	539	822	VE6_HPV70	Human papillomavirus type 70, protein E6	Gardiol et al. (5)	10523825
DLG1 PDZ1	TRRETQV	406	132	538	VE6_HPV39	Human papillomavirus type 39 protein E6	Gardiol et al. (5)	10523825

Table S2. Cont.

Protein	Peptide	Count A	Count B	Total	Uniprot	Name	Source (for the protein, not always the exact variant)	PMID
DLG1 PDZ2	IRRETQV	102,877	71,200	174,077	VE6_HPV70	Human papillomavirus type 70, protein E6	Gardiol et al. (5)	10523825
DLG1 PDZ2	TRRETQV	17,861	24,016	41,877	VE6_HPV39	Human papillomavirus type 39 protein E6	Gardiol et al. (5)	10523825
DLG1 PDZ2	TRRETEV	4,089	4,716	8,805	VE6_HPV35	Human papillomavirus type 35 protein E6	Gardiol et al. (5)	10523825
DLG1 PDZ2	TGRSTEV	1,223	6,115	7,338	VU47_HHV6U	Human herpesvirus 6A glycoprotein U47	Blot et al. (6)	15286176
DLG1 PDZ2	RRRETQV	2,679	4,196	6,875	VE6_HPV18/45	Human papillomavirus type 18/45 protein E6	Gardiol et al. (5)	10523825
DLG1 PDZ2	RGIESEV	329	1,321	1,650	NS1_I63A1	Influenza A virus (Avian) nonstructural protein 1	Liu et al. (7)	20702615
DLG1 PDZ2	RRRETAL	382	706	1,088	VE6_HPV33	Human papillomavirus type 33 protein E6	Gardiol et al. (5)	10523825
DLG1 PDZ2	HFRETEV	100	534	634	TAX_HTL1F	Human T-cell leukemia virus 1 protein Tax-1	Lee et al. (8)	9192623
DLG1 PDZ2	TRRETQL	175	359	534	VE6_HPV16	Human papillomavirus type 16 protein E6	Gardiol et al. (5)	10523825
DLG1 PDZ2	TRQETQV	22	323	345	VE6_HPVME	Human papillomavirus type ME180 protein E6	Gardiol et al. (5)	10523825
DLG1 PDZ2	IRQETQV	46	293	339	VE6_HPV68	Human papillomavirus type 68 protein E6	Gardiol et al. (5)	10523825
DLG1 PDZ2	RHRETYV	25	133	158	US32_HCMVA	Human cytomegalovirus, uncharacterized protein HHRF7		
DLG1 PDZ2	PRTETQV	9	139	148	VE6_HPV31	Human papillomavirus type 31 protein E6	Gardiol et al. (5)	10523825
DLG1 PDZ2	RQTETQV	8	101	109	VE6_HPV26	Human papillomavirus type 26 protein E6	Gardiol et al. (5)	10523825
DLG1 PDZ2	QRNETQV	24	72	96	VE6_HPV51	Human papillomavirus type 51 protein E6	Gardiol et al. (5)	10523825
DLG1 PDZ2	RRIESEV	3	22	25	NS1_I49A1	Influenza A virus (Avian) nonstructural protein 1	Liu et al. (7)	20702615
DLG1 PDZ2	RRVESEV	3	18	21	NS1_I82A8	Influenza A virus (Avian) nonstructural protein 1	Liu et al. (7)	20702615
DLG1 PDZ2	RRRQTQV	0	11	11	VE6_HPV58	Human papillomavirus type 58 protein E6	Gardiol et al. (5)	10523825
DLG1 PDZ2	TGRSTTL	1	9	10	VFUS_SHEVK	Sheeppox virus putative fusion protein		
DLG1 PDZ3	IRRETQV	61,437	31,152	92,589	VE6_HPV70,	Human papillomavirus type 70, protein E6	Gardiol et al. (5)	10523825
DLG1 PDZ3	TRRETQV	19,459	10,283	29,742	VE6_HPV39,	Human papillomavirus type 39 protein E6	Gardiol et al. (5)	10523825
DLG1 PDZ3	RRRETQV	2,336	8,525	10,861	VE6_HPV45,	Human papillomavirus type 18/45 protein E6	Gardiol et al. (5)	10523825
DLG1 PDZ3	RRRETAL	1,213	1,297	2,510	VE6_HPV33,	Human papillomavirus type 33 protein E6	Gardiol et al. (5)	10523825
DLG1 PDZ3	TRRETEV	90	227	317	VE6_HPV35,	Human papillomavirus type 35 protein E6	Gardiol et al. (5)	10523825
DLG1 PDZ3	RHRETYV	26	49	75	US32_HCMVA	Human cytomegalovirus, uncharacterized protein HHRF7		

Column labels are as in Table S1, plus a "Name" describing the protein and virus.

1. Nakagawa S, Huibregtse JM (2000) Human scribble (Vartul) is targeted for ubiquitin-mediated degradation by the high-risk papillomavirus E6 proteins and the E6AP ubiquitin-protein ligase. *Mol Cell Biol* 20(21):8244–8253.
2. Arpin-André C, Mesnard JM (2007) The PDZ domain-binding motif of the human T cell leukemia virus type 1 tax protein induces mislocalization of the tumor suppressor hScrib in T cells. *J Biol Chem* 282(45):33132–33141.
3. Okajima M, et al. (2008) Human T-cell leukemia virus type 1 Tax induces an aberrant clustering of the tumor suppressor Scribble through the PDZ domain-binding motif dependent and independent interaction. *Virus Genes* 37(2):231–240.
4. Song C, et al. (2009) Tax1 enhances cancer cell proliferation via Ras-Raf-MEK-ERK signaling pathway. *IUBMB Life* 61(6):685–692.
5. Gardiol D, et al. (1999) Oncogenic human papillomavirus E6 proteins target the discs large tumour suppressor for proteasome-mediated degradation. *Oncogene* 18(40):5487–5496.
6. Blot V, et al. (2004) Human Dlg protein binds to the envelope glycoproteins of human T-cell leukemia virus type 1 and regulates envelope mediated cell-cell fusion in T lymphocytes. *J Cell Sci* 117(Pt 17):3983–3993.
7. Liu H, et al. (2010) The ESEV PDZ-binding motif of the avian influenza A virus NS1 protein protects infected cells from apoptosis by directly targeting Scribble. *J Virol* 84(21):11164–11174.
8. Lee SS, Weiss RS, Javier RT (1997) Binding of human virus oncoproteins to hDlg/SAP97, a mammalian homolog of the *Drosophila* discs large tumor suppressor protein. *Proc Natl Acad Sci USA* 94(13):6670–6675.

Table S3. Selected peptides of low interest to this study

PDZ domain	Peptide	Protein	# Library A	# Library B	Total count
Additional target peptides from the proteolytic TopFind set					
Scribble PDZ1	DRDYMGW	CCKN_HUMAN, CLEAVAGE-8141	343	0	343
Scribble PDZ2	GFYESDV	A2MG_HUMAN CLEAVAGE-593	1	328	329
Scribble PDZ3	GFYESDV	A2MG_HUMAN CLEAVAGE-593	0	242	242
Scribble PDZ3	WTTSTDL	AMPH_HUMAN INFERRED FROM CLEAVAGE-4706	1,823	62	1,885
Additional target peptides from the ENSEMBL-only set					
Scribble PDZ1	KTYETDL	ENSP00000447314	1,326	56	1,382
Scribble PDZ3	KTYETDL	ENSP00000447314	2,224	12,398	14,622
Scribble PDZ3	LLRETSL	ENSP00000420911	97	0	97
Scribble PDZ3	VSRETKL	ENSP00000415771	0	85	85
Scribble PDZ3	GIRESKL	ENSP00000399301	0	79	79
Scribble PDZ3	GVRKETA	ENSP00000451805	0	20	20
Scribble PDZ3	FSEGTDL	ENSP00000440057	0	10	10
Scribble PDZ3	AGKTTIL	ENSP00000450315	0	8	8
Erbin	QENDWWV	ENSP00000398110	28,359	1,117	29,476
Erbin	QHHWESW	ENSP00000270281	0	40	40
DLG1 PDZ2	FPKETQV	ENSP00000442101	0	2,275	2,275
DLG1 PDZ2	SGTAYLL	ENSP00000449745	8	7	1

The peptides correspond to protein C termini either only supported by an experiment in TopFind (resulting from proteolytic cleavage or COFRADIC-based complementary positional proteomics experiments) or only found in ENSEMBL. PDZ domain, the identity of the bait PDZ domain; Peptide, the selected C-terminal peptides; Protein, the identifier corresponding to a selected peptide and the cleavage site when applicable; # Library A and # Library B, the sequencing counts for a given peptide from the replicate selection.

Other Supporting Information Files

Dataset S1 (XLSX)
Dataset S2 (XLSX)