Supplementary Information

Activation of Src and transformation by a RPTPα splice mutant found in human tumors

Jian Huang, Lin Yao, Rongting Xu, Huacheng Wu, Min Wang, Brian S. White, David Shalloway and Xinmin Zheng

Includes: Supplementary Discussion; Figs. S1 and S2; Tables S1–S3; Supplementary References

SUPPLEMENTARY DISCUSSION

Alternatively spliced isoforms of RPTPα found in normal cells

Two alternatively spliced mRNAs, differing in their 5' untranslated regions, are translated into the ubiquitously expressed 793 amino acid RPTPα isoform that is found in all the cancer cell types studied here. A third mRNA variant expresses RPTPα802, an isoform with a nine-amino acid insertion in the extracellular domain, which is expressed in a few tissues such as brain and fat (Daum et al, 1994; Kaplan et al, 1990; Krueger et al, 1990; Matthews et al, 1990; Norris et al, 1997; Sap et al, 1990). Tyr798 in RPTPα802 is homologous to Tyr789 in RPTPα; for clarity we to refer to this as Tyr798/789 in the discussion below.

Effect of Tyr789→Phe mutation in different biological systems

Care is required in interpreting experiments testing the effect of Y789F mutation, which is quantitative and biological-system dependent: Vacaru and den Hertog (2010a) analyzed overexpressed Src and RPTPα in Src/Yes/Fyn-deficient cells that were immortalized by SV40 large T antigen and showed that Y798/789F mutation of RPTPα or, conversely, SH2 mutation of Src, reduces but does not eliminate RPTPα-Src association (Fig. 7B, C of that paper). Thus, it is not surprising that some ability of RPTPα802(Y798/789F) to dephosphorylate Src can be detected when it is expressed at high levels by a retroviral vector (Kapp et al, 2007). However, even in this case it does not transform NIH3T3 cells unless Src is also overexpressed (Kapp 2007). Interestingly, wt RPTPα802 cannot be stably expressed in this system, consistent with the possibility that increased activity relative to RPTPα802(Y798/789F) renders it toxic (Lammers, 2000).

Potential cellular roles of cryptic exon cx95

Cx95 it is flanked by excellent U2 splice signals (Fig. 2A), its 3' splice boundary is contained in a bioinformatic database of predicted splice junctions (Wang et al, 2008), and the ExonScan algorithm (Wang et al, 2004), which analyzes a sequence for 5' and 3' splice sites and the presence/absence of exonic splicing enhancer/silencer motifs, predicts that it is an internal exon ($p = 0.001$). In addition, the high conservation within placental mammals of cx95 and the even higher conservation of its flanking regions (Fig. 7) suggests that it is an alternatively spliced exon: As previously noted and verified in Fig. 7B, the flanking regions of alternatively spliced exons are generally more conserved than those of constitutive exons (Kaufmann et al, 2004; Sorek & Ast, 2003; Sugnet et al, 2006; Yeo et al, 2005); the flanking regions of cx95 are more conserved than most known alternatively spliced exons. Moreover, RankVista analysis (Frazer et al, 2004; Prabhakar et al, 2006; Wang et al, 2007; http://genome.lbl.gov/vista/index.shtml) identifies the 317 nt region which includes cx95 as an "extremely conserved region" (ECR; $p=0.06$). Visel et al. (2008) have reported that expression of roughly half of such ECRs enhanced reporter gene expression in various tissues of the developing mouse embryo, particularly in the central nervous system. Intriguingly, the primary defects observed in RPTPα-

knockout mice involve brain and neuronal function (Kostic et al, 2007; Petrone et al, 2003; Skelton et al, 2003), brain has the highest proportion of alternatively spliced genes (Yeo et al, 2004), and brain expresses large amounts of a RPTPα splice variant having a small insertion in its extracellular domain (Kapp et al, 2007).

This suggests that cx95 might normally function as an alternative exon in brain and/or other tissues, particularly during development. For example, RPTPα245 itself might be expressed to activate wt RPTPα in restricted tissue types and/or at special times. This could have escaped detection, particularly since many anti-RPTPα antibodies might not react with this truncated form. Indeed, a transcript [FLJ Human cDNA Database (http://flj.lifesciencedb.jp) FLJ33424; GenBank AK090743.1] in which cx95 is inserted exactly as in RPTPα245 has been found in a screen of full-length cDNAs from normal human brain (Ota et al, 2004). (This transcript has additional splice variations that preclude translation of cx95, but it supports the existence of a normal mechanism for cx95 inclusion in an mRNA.)

A different possibility follows from the suggestion (Ni et al, 2007) that the inclusion of extremely conserved alternatively spliced stop-codon-exons such as cx95 is used to regulate expression of normal mRNAs by diverting pre-mRNAs to futile endpoints that are degraded by nonsense-mediated decay (NMD). Indeed, the RPTPα245 mRNA satisfies the prediction rule for NMD (Chang et al, 2007) and might normally be degraded. In that case its survival might result from changes in the tumor microenvironment such as hypoxia and amino acid deprivation that have been shown to inhibit NMD (Gardner; 2010). However, there is currently no direct *in vivo* evidence linking inhibition of NMD to tumor formation.

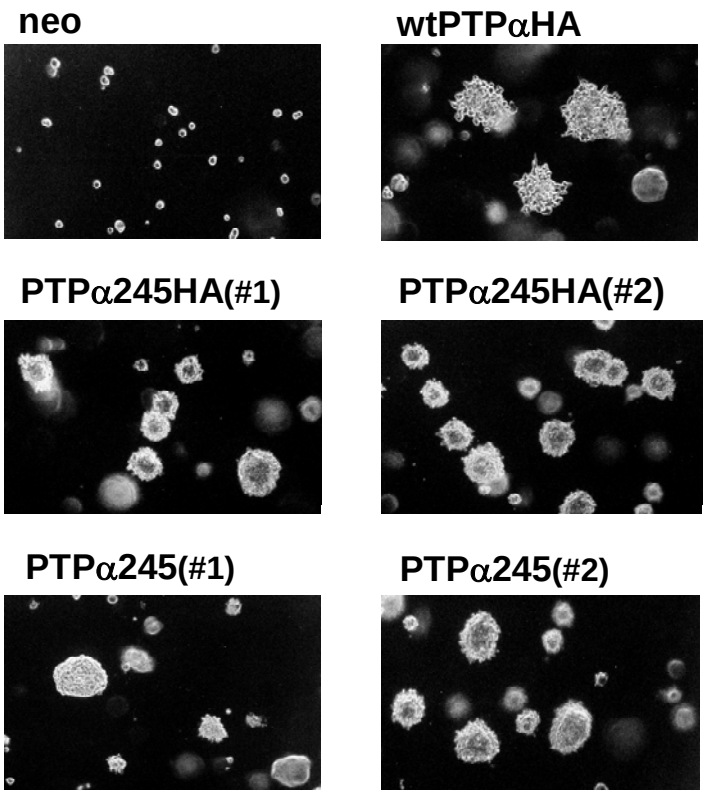

Figure S1. Anchorage-independent growth induced by RPTPα245. Inducible REFs expressing no exogenous RPTPα (neo), RPTPα245HA, or RPTPα245 were suspended in media containing 0.3% soft agarose without doxycycline and grown for 21 d.

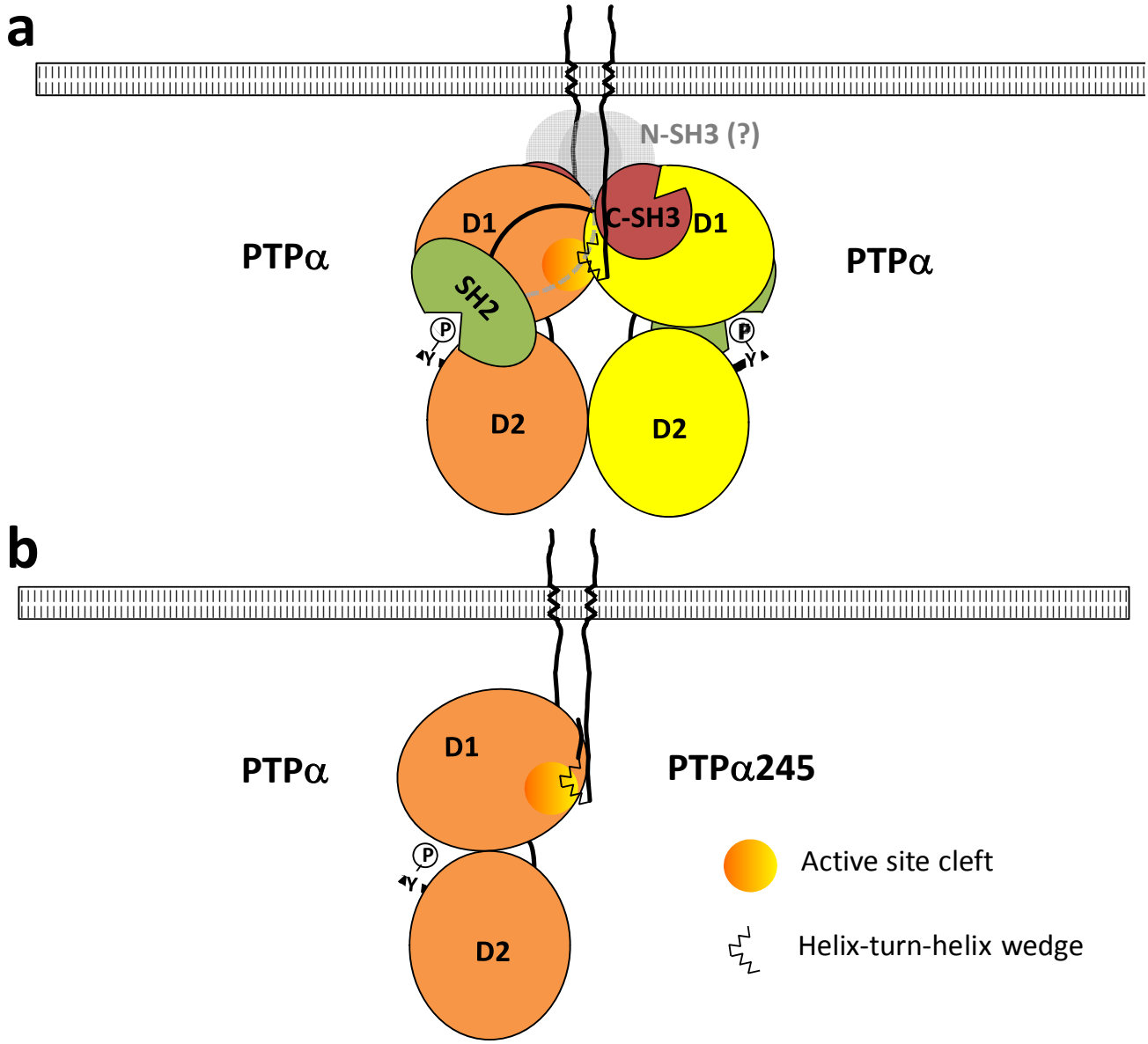

Figure S2. Hypothetical Grb2-RPTPα Cross-Binding Model. a, The positions of the D1 domains of homodimerized RPTPα are based on the dimer crystal structure which shows how the amino-proximal helix-turn-helix wedge of one D1 (spanning amino acids 214–-242) inserts into the active site cleft of the other D1 (Bilwes et al, 1996). (The left wedge-right cleft interaction is hidden on the back of the dimer.) Under normal conditions, the Grb2 SH2 domain is bound to RPTPα pTyr789 (den Hertog et al, 1994; Su et al, 1994) and its C-terminal SH3 domain is bound to RPTPα D1 by an interaction involving Arg469 (den Hertog & Hunter, 1996; Su et al, 1996). The position of the N-terminal Grb2 SH3 domain (shown in gray) is unknown; it is located here according to the suggestion (den Hertog & Hunter, 1996) [based on the Grb2 crystal structure (Maignan et al, 1995) and the interference of RPTPα-Grb2 binding with SOS binding] that it is near the C-SH3. We hypothesize that the Grb2 SH2 domain binds to pTyr789 from one RPTPα monomer while the C-SH3 domain binds to D1 from the other monomer. While the drawing supposes that the dimer has the "head-to-head" conformation suggested by the RPTPα D1-dimer crystal structure (Bilwes et al, 1996), the model is also consistent with the "head-to-toe" conformation suggested by analogy to the RPTPγ-dimer structure (Barr et al, 2009) and an analogous drawing can be made for that case. **b**, When RPTP245 displaces one RPTPα monomer, the binding of Grb2 to the bound RPTPα is weakened or eliminated because of the lack of a cross-binding site for C-SH3. The drawing assumes that RPTPα245 and RPTPα form a heterodimer analogous to the homodimer. However, the model only requires that RPTPα245 binding disrupts the RPTPα homodimer by any means.

Patient number	Tissue of origin	Gender/Age/Cancer Stage*	Variants sequenced†
2	breast	F/49/IV	wt (10)
63	breast	F/37/II	wt (10)
68	breast	F/56/III	wt (10)
76	breast	F/51/III	wt (10)
84	breast	F/43/IV	wt (10)
90	breast	F/58/III	wt (10)
93	breast	F/67/IV	wt (10)
69	breast	F/62/III	wt (5)/RPTPα245 (5)
128	breast	F/55/III	wt (7)/RPTPα652 (3)
8	colon	F/68/II	wt (10)
10	colon	M/76/III	wt (10)
36	colon	M/74/II	wt (10)
46	colon	M/73/II	wt (10)
65	colon	M/72/III	wt (10)
97	colon	M/61/II	wt (5)/RPTPα245 (5)
106	colon	M/59/III	wt (6)/RPTPα245 (4)
223	colon	M/67/IV	wt (8)/RPTPα445 (2)
62	liver	M/53/IV	wt (10)
58	liver	M/48/III	wt (5)/RPTPα245 (5)
61	lung	M/53/III	wt (10)
103	lung	F/59/II	wt (10)
104	lung	F/53/II	wt (10)
127	lung	F/45/III	wt (10)
250	lung	F/57/II	wt (10)
158	lung	M/44/III	wt (10)
160	lung	M/62/II	wt (10)
64	thyroid	M/37/I	wt (10)
74	thyroid	F/52/II	wt (10)
86	thyroid	M/30/II	wt (10)
121	thyroid	F/33/II	wt (10)
123	thyroid	F/47/III	wt (10)
176	thyroid	F/27/II	wt (10)
186	thyroid	F/34/I	wt (10)
206	thyroid	F/59/III	wt (10)
218	thyroid	F/71/II	wt (10)
247	thyroid	F/49/III	wt (10)

Table S1. Related to Figure 1 and Table 1. RPTPα mutations in tumors from human cancer patients.
*Cancer stage reported by pathologists at Ruijin Hospital, Branch of Shanghai First Hospital, Shanghai Jiaotong University School of Medicine.

†Wild-type (wt) sequences exactly matched the "short" (793 amino acid) RPTPα isoform sequence [Vega (Deloukas et al, 2001) transcripts PTPRA-001/002]. The numbers in parentheses indicate the number of tumor clones that had each sequence. All sequences from the patient-matched normal tissues (5-10 replicates) were wt.

Protein	Plasmid	Cell line	siRNA	% foci formation in cell mixing	% colony formation in soft agarose	% reduction in colony formation
–	pTet-SPLICE	REF(pTet-SPLICE)#1		≤ 0.1	$\leq 10^{-2}$	
		REF(pTet-SPLICE)#2		≤ 0.1	$\leq 10^{-2}$	
RPTPαHA	pTPTPα	REF(RPTPαHA)#1		13 ± 1	17 ± 2	
		REF(RPTPα HA)#2		16 ± 0.3	14 ± 2	
RPTPα245	pNTPTPα245	REF(RPTPα245)#1	–	10 ± 1	16 ± 0.5	
			rRPTPα-1			92 ± 5
			rRPTPα-1 (mutated)			5 ± 7
			rRPTPα-2			92 ± 2
			rRPTPα-2 (mutated)			12 ± 19
			rSrc-1			87 ± 1
			rSrc-1 (reversed)			1 ± 1
			rSrc-1 (mutated)			2 ± 2
			rSrc-2			94 ± 1
			rSrc-2 (mutated)			3 ± 5
RPTPα245HA	pTPTPα245	REF(RPTPα245)#1	–	11 ± 0.4	12 ± 2	
		REF(RPTPα245)#2		15 ± 1	11 ± 1	
			rRPTPα-1			92 ± 0.4
			rRPTPα-1 (mutated)			6 ± 3
			rRPTPα-2			88 ± 3
			rRPTPα-2 (mutated)			2 ± 10
		REF(RPTPα245HA)#1		14 ± 0.3	15 ± 1	
		REF(RPTPα245HA)#2		12 ± 1	11 ± 1	

Table S2. Related to Figure 4. Transforming activities of wild-type and mutant RPTPα overexpressor cells.

Coselected Fisher rat embryo fibrobast (REF) lines inducibly expressing the indicated RPTPα variants in the absence of doxycycline were tested for focus-formation by mixing 300 test cells with 10^5 REFs or were transfected with the indicated rat siRNAs and mutated siRNAs (Supplementary Table 3) and tested for growth in 0.3% soft agarose. The percentage of cells forming foci or colonies (e.g., see Figs. 4A, 4C, and S1) was determined after 21 d growth. Average values and SEs are shown (n=2 to 4).

RPTPα RT-PCR primers*		**Codons sequenced**
N portion: 5’	5’-AGC**ATG**GATTCCTGGTTCATTCTTGTTCTG-3’	1–361
3’	5’-CTCTACAGACACCCGAATATTCCCATAG-3’	
C portion: 5’	5’-AGCCACCATCGTCATGGTTACCAACCT-3’	344–793
3’	5’-CGC**TTA**CTTGAAGTTGGCATAATCTGA-3’	
Sequencing primers (matching flanking pCR2.1-TOPO)		
5’	5’-TTGGTACCGAGCTCGGATCCACTAGTAAC-3’	
3’	5’-GCCAGTGTGATGGATATCTGCAGAATTC-3’	
RPTP245α cloning primers[†]		
5’	5’-CGCCAAGCTTGGCCACC**ATG**GATTCCTGGTTCATTCTTGT-3’	
3’ (HA)	5’-TGTCAAGCTT**CA***TGCGTAGTCTGCACGTCGTATGGGTAC*GC TTGTCAAGTTCCACAAAAAAACGTT-3’	
3’ (no HA)	5’-TGTCAAGCTT**CA**GCTTGTCAAGTTCCACAAAAAAACGTT-3’	
PPLL mutagenesis oligomer[‡]	5’-AGCACCAACAGGAAATAC**CTCCTC**CTGCCCGTGGACAAG-3’	
siRNA sequences (sense strand)		**Start location** [§]
rRPTPα-1	5’-AGCCUGUAACCCUCAGUAUdTdT-3’	1266
rRPTPα-1(mutated)	5’-AGC**U**UGUAA**U**CCUCA**A**UAUdTdT-3’	1266
rRPTPα-2	5’-GGCCCUUCUGGAGCAUUAUdTdT-3’	1470
rRPTPα-2 (mutated)	5’-GGC**U**CUUCU**A**GAGCA**C**UAUdTdT-3’	1470
rSrc-1	5’-GGCUGCAGAUUGUCAAUAAdTdT-3’	329
rSrc-1(mutated)	5’-GGC**C**GCAGAU**C**GUCAA**C**AAdTdT-3’	329
rSrc-2	5’-GGCUCAUAGAGGACAACGAdTdT-3’	1235
rSrc-2 (mutated)	5’-GGC**C**CAUAGA**A**G*A*CAAC**A**AdTdT-3’	1235

Table S3. Related to Figures 1, 2, and 4. Primer, oligomer, and siRNA sequences.

*The nucleotides corresponding to the start and termination codons are in boldface. ATG start and (reversed) TAA stop codons are in bold.

[†]The cloning primers were used to amplify the RPTPα245 coding region from plasmid pCR2.1RPTPα245 while simultaneously modifying the translational start to a GCCACCATGG Kozak consensus (Kozak, 1987), inserting flanking *Hind*III sites (underlined), and, in the case of the 3’ (HA) primer, inserting a downstream HA epitope coding sequence YPYDVPDYA (italics; Wilson et al, 1984). The ATG start codon and downstream (reversed) TGA stop codon are in bold. The only RPTPα-derived sequence in the resultant cloning insert is the RPTPα245 coding region.

[‡] The Pro210→Leu/Pro211→Leu oligomer spanned nucleotides 610-648 of the human RPTPα mRNA sequence (Vega transcripts PTPRA-001/002) with the beginning of the coding sequence numbered 1.

[§]Locations indicate the 5’ positions of the homologous rat mRNA sequences as per GenBank Locus NM_012763.2 (RPTPα) or AF130457.1 (Src) with the coding sequence starts numbered 1. The nucleotides in the mutated siRNAs that do not match the wt rat sequences are in bold. The siRNA sequences were selected according to criteria suggested by Tuschl and coworkers (Elbashir et al, 2001) and synthesized by Shanghai GenePharma Co. rRPTPα-1 and -2 are located within coding sequence exons 11 and 13, respectively. Both of these exons are deleted in the RPTPα245 coding sequence included in the expression plasmids.

SUPPLEMENTARY REFERENCES

Chang YF, Imam JS, Wilkinson MF (2007) The nonsense-mediated decay RNA surveillance pathway. *Annu Rev Biochem* **76:** 51-74

Daum G, Regenass S, Sap J, Schlessinger J, Fischer EH (1994) Multiple forms of the human tyrosine phosphatase RPTPα. Isozymes and differences in glycosylation. *J Biol Chem* **269:** 10524-10528

Deloukas P, Matthews LH, Ashurst J, Burton J, Gilbert JG, Jones M, Stavrides G, Almeida JP, Babbage AK, Bagguley CL, Bailey J, Barlow KF, Bates KN, Beard LM, Beare DM, Beasley OP, Bird CP, Blakey SE, Bridgeman AM, Brown AJ, Buck D, Burrill W, Butler AP, Carder C, Carter NP, Chapman JC, Clamp M, Clark G, Clark LN, Clark SY, Clee CM, Clegg S, Cobley VE, Collier RE, Connor R, Corby NR, Coulson A, Coville GJ, Deadman R, Dhami P, Dunn M, Ellington AG, Frankland JA, Fraser A, French L, Garner P, Grafham DV, Griffiths C, Griffiths MN, Gwilliam R, Hall RE, Hammond S, Harley JL, Heath PD, Ho S, Holden JL, Howden PJ, Huckle E, Hunt AR, Hunt SE, Jekosch K, Johnson CM, Johnson D, Kay MP, Kimberley AM, King A, Knights A, Laird GK, Lawlor S, Lehvaslaiho MH, Leversha M, Lloyd C, Lloyd DM, Lovell JD, Marsh VL, Martin SL, McConnachie LJ, McLay K, McMurray AA, Milne S, Mistry D, Moore MJ, Mullikin JC, Nickerson T, Oliver K, Parker A, Patel R, Pearce TA, Peck AI, Phillimore BJ, Prathalingam SR, Plumb RW, Ramsay H, Rice CM, Ross MT, Scott CE, Sehra HK, Shownkeen R, Sims S, Skuce CD, Smith ML, Soderlund C, Steward CA, Sulston JE, Swann M, Sycamore N, Taylor R, Tee L, Thomas DW, Thorpe A, Tracey A, Tromans AC, Vaudin M, Wall M, Wallis JM, Whitehead SL, Whittaker P, Willey DL, Williams L, Williams SA, Wilming L, Wray PW, Hubbard T, Durbin RM, Bentley DR, Beck S, Rogers J (2001) The DNA sequence and comparative analysis of human chromosome 20. *Nature* **414:** 865-871

Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411:** 494-498

Kaplan R, Morse B, Huebner K, Croce C, Howk R, Ravera M, Ricca G, Jaye M, Schlessinger J (1990) Cloning of three human tyrosine phosphatases reveals a multigene family of receptor-linked protein-tyrosine-phosphatases expressed in brain. *Proc Natl Acad Sci U S A* **87:** 7000-7004

Kostic A, Sap J, Sheetz MP (2007) RPTPα is required for rigidity-dependent inhibition of extension and differentiation of hippocampal neurons. *J Cell Sci* **120:** 3895-3904

Krueger NX, Streuli M, Saito H (1990) Structural diversity and evolution of human receptor-like protein tyrosine phosphatases. *EMBO J* **9:** 3241-3252

Maignan S, Guilloteau JP, Fromage N, Arnoux B, Becquart J, Ducruix A (1995) Crystal structure of the mammalian Grb2 adaptor. *Science* **268:** 291-293

Matthews RJ, Cahir ED, Thomas ML (1990) Identification of an additional member of the protein-tyrosine-phosphatase family: evidence for alternative splicing in the tyrosine phosphatase domain. *Proc Natl Acad Sci U S A* **87:** 4444-4448

Ni JZ, Grate L, Donohue JP, Preston C, Nobida N, O'Brien G, Shiue L, Clark TA, Blume JE, Ares M, Jr. (2007) Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes Dev* **21:** 708-718

Norris K, Norris F, Kono DH, Vestergaard H, Pedersen O, Theofilopoulos AN, Moller NP (1997) Expression of protein-tyrosine phosphatases in the major insulin target tissues. *FEBS Lett* **415:** 243-248

Ota T, Suzuki Y, Nishikawa T, Otsuki T, Sugiyama T, Irie R, Wakamatsu A, Hayashi K, Sato H, Nagai K, Kimura K, Makita H, Sekine M, Obayashi M, Nishi T, Shibahara T, Tanaka T, Ishii S, Yamamoto J, Saito K, Kawai Y, Isono Y, Nakamura Y, Nagahari K, Murakami K, Yasuda T, Iwayanagi T,

Wagatsuma M, Shiratori A, Sudo H, Hosoiri T, Kaku Y, Kodaira H, Kondo H, Sugawara M, Takahashi M, Kanda K, Yokoi T, Furuya T, Kikkawa E, Omura Y, Abe K, Kamihara K, Katsuta N, Sato K, Tanikawa M, Yamazaki M, Ninomiya K, Ishibashi T, Yamashita H, Murakawa K, Fujimori K, Tanai H, Kimata M, Watanabe M, Hiraoka S, Chiba Y, Ishida S, Ono Y, Takiguchi S, Watanabe S, Yosida M, Hotuta T, Kusano J, Kanehori K, Takahashi-Fujii A, Hara H, Tanase TO, Nomura Y, Togiya S, Komai F, Hara R, Takeuchi K, Arita M, Imose N, Musashino K, Yuuki H, Oshima A, Sasaki N, Aotsuka S, Yoshikawa Y, Matsunawa H, Ichihara T, Shiohata N, Sano S, Moriya S, Momiyama H, Satoh N, Takami S, Terashima Y, Suzuki O, Nakagawa S, Senoh A, Mizoguchi H, Goto Y, Shimizu F, Wakebe H, Hishigaki H, Watanabe T, Sugiyama A, Takemoto M, Kawakami B, Watanabe K, Kumagai A, Itakura S, Fukuzumi Y, Fujimori Y, Komiyama M, Tashiro H, Tanigami A, Fujiwara T, Ono T, Yamada K, Fujii Y, Ozaki K, Hirao M, Ohmori Y, Kawabata A, Hikiji T, Kobatake N, Inagaki H, Ikema Y, Okamoto S, Okitani R, Kawakami T, Noguchi S, Itoh T, Shigeta K, Senba T, Matsumura K, Nakajima Y, Mizuno T, Morinaga M, Sasaki M, Togashi T, Oyama M, Hata H, Komatsu T, Mizushima-Sugano J, Satoh T, Shirai Y, Takahashi Y, Nakagawa K, Okumura K, Nagase T, Nomura N, Kikuchi H, Masuho Y, Yamashita R, Nakai K, Yada T, Ohara O, Isogai T, Sugano S (2004) Complete sequencing and characterization of 21,243 full-length human cDNAs. *Nat Genet* **36:** 40-45

Petrone A, Battaglia F, Wang C, Dusa A, Su J, Zagzag D, Bianchi R, Casaccia-Bonnefil P, Arancio O, Sap J (2003) Receptor protein tyrosine phosphatase α is essential for hippocampal neuronal migration and long-term potentiation. *EMBO J* **22:** 4121-4131

Sap J, D'Eustachio P, Givol D, Schlessinger J (1990) Cloning and expression of a widely expressed receptor tyrosine phosphatase. *Proc Natl Acad Sci U S A* **87:** 6112-6116

Skelton MR, Ponniah S, Wang DZ, Doetschman T, Vorhees CV, Pallen CJ (2003) Protein tyrosine phosphatase alpha (PTPα) knockout mice show deficits in Morris water maze learning, decreased locomotor activity, and decreases in anxiety. *Brain Res* **984:** 1-10

Visel A, Prabhakar S, Akiyama JA, Shoukry M, Lewis KD, Holt A, Plajzer-Frick I, Afzal V, Rubin EM, Pennacchio LA (2008) Ultraconservation identifies a small subset of extremely constrained developmental enhancers. *Nat Genet* **40:** 158-160

Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature* **456:** 470-476

Wilson IA, Niman HL, Houghten RA, Cherenson AR, Connolly ML, Lerner RA (1984) The structure of an antigenic determinant in a protein. *Cell* **37:** 767-778

Yeo G, Holste D, Kreiman G, Burge CB (2004) Variation in alternative splicing across human tissues. *Genome Biol* **5:** R74