Supplementary info	rmation S4   S	ignificance of the n	umber of prote	eins with Pfam	domains w	ithin specific	cancer sub-cla	ISSES		
Pfam domain	Dominant	versus recessive	Germline versu	is somatic	Cell type				Translocation ve translocation	Prsus no
	Dominant	Recessive	Germline	Somatic	Epithelial	Leukaemia/ lymphoma	Mesenchymal	Other	Translocation	No translocation
Pkinase	S	I	I	S	S	I	I	S	I	S
НГН	ഗ	I	Ι	S	S	S	I	I	S	I
Homeobox	ഗ	I	I	S	I	S	I	I	S	I
Ets	ഗ	I	I	S	I	I	S	I	S	I
PAX	S	I	I	S	I	I	S	I	S	I
PHD	ഗ	I	I	S	I	S	I	I	S	I
DNA_mis_repair	I	S	S	I	S	I	I	I	I	S
HATPase_c	I	S	S	I	S	I	I	I	I	S
Myc_N_term	ഗ	I	I	S	S	I	I	I	I	S
AF-4	S	I	I	S	I	S	I	I	S	I
Bromodomain	I	I	I	S	S	I	1	I	S	1
DUF902	I	I	I	S	I	S	I	I	S	I
KIX	I	I	I	S	I	S	I	I	S	I
SH2	ഗ	I	I	S	I	I	I	I	S	I
PWWP	I	I	I	I	I	I	I	I	I	I
DUF906	I	I	I	S	I	S	I	I	S	I
HRDC	I	S	S	I	S	S	I	I	I	S
zf-TAZ	I	I	I	S	I	S	I	I	S	I
ETS_PEA3_N	ഗ	I	I	S	I	I	S	I	S	I
SAM_PNT	T	I	I	S	I	I	S	I	S	I
MutS_I	I	S	S	I	S	I	I	I	I	S
Exostosin	T	S	S	I	I	I	S	I	Ι	S
SSXT	T	I	I	I	Ι	T	S	I	I	I
Nucleoporin_FG	T	I	I	I	Ι	T	Ι	I	Ι	I
MutS_II	I	S	S	I	S	I	I	I	I	S
SH3	S	I	I	S	I	S	I	I	S	I
GTP_CDC	I	I	I	I	I	S	I	I	I	I
DEAD	I	I	T	I	Ι	T	I	I	I	I
ERM	I	I	I	I	I	I	I	I	I	I
YEATS	I	I	I	I	I	I	I	I	I	I
HSP90	I	I	I	I	I	I	I	I	I	I
MutS_IV	I	S	S	I	S	I	I	I	I	S
AT_hook	T	I	I	I	Ι	T	I	I	I	1
Helicase_C	I	S	S	I	I	I	I	I	I	I
MutS_III	T	S	S	I	S	T	I	I	I	S
Recep_L_domain	I	I	I	I	I	I	I	I	I	S

PDGF_N	Topoisomer_I_N	Topoisomerase_I	HNF-1B_C	APC_basic	CBF_beta	Menin	CARD	Myosin_tail	LIM	Fork_head	Fanconi	XPA_C	BRCA2	UPF0023	Myc-LZ	VHL	HNF-1A_C	WT1	Cdc73	XPA_N	EB1_binding	Hamartin	CybS	GMP_synt_C	APC_15aa	Fanconia	Tuberin	Ribosomal_L22e	AICARFT_IMPCHas	COX6C	Armadillo_seg	Myosin_N	ENTH	MutS_V	Tropomyosin	MOZ_SAS	7tm_1	zf-RanBP	Furin-like
I	Ι	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	1
I	I	I	I	I	I	I	I	I	T	I	I	I	I	I	I	I	I	I	I	I	T	I	T	I	I	I	I	I	I	I	I	I	I	S	I	I	I	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	T	I	T	I	I	I	I	I	I	I	I	I	I	S	I	I	T	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	T	I	T	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
I	I	I	I	I	I	I	I	I	I	T	I	I	I	I	I	I	I	I	I	I	T	I	T	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
I	I	Ι	I	I	I	I	I	I	1	I	I	I	I	I	1	I	I	I	I	I	T	I	T	I	I	I	I	I	I	I	I	I	I	I	I	I	I	S	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	S	I	S	I	I	S	S	S	I	S	I	S	I	I	I	S	I	I	I	I	I	I	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	T	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	T	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	S

FYRN	death	RhoGAP	TIG	P53	Hox9_act	XPG_I	Band_41	HNF-1_N	CbI_N	lyase_1	Cbl_N3	Cbl_N2	FHA	YLP	ERCC4	BH4	APC_crr	MGS	Runt	Nucleoporin2	Rabaptin	RB_B	Caudal_act	RB_A	TCL1_MTCP1	SAMP	22	zf-C4	hormone_rec	zf-MYND	Rad4	Sdh_cyt	Clathrin_propel	MoeA_C	Fip1	SNF5	MoeA_N	SUFU	Nucleoplasmin
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	T	I	I	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	T	I	I	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	T	I	I	I	I
1	1	I	1	I	1	I	1	1	1	1	1	I	1	I	1	I	1	I	1	I	1	1	I	I	1	1	1	1	1	1	1	1	1	1	1	1	1	1	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	1	I	1	1	1	I	I	T	I	I	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	1	I	I	I	I	I	1	1	I	I	1	1	I	T	I	I	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	1	1	1	1	1	1	I	I	I	I	I.	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	1	I	I	I	I	I	I	I	I	I	I	I	I	T	I	I	1	I

Activin_recp	RHO	BIR	RIIa	PDGF	IRF	SET	UBA	Nebulin	WH1	Peptidase_M28	FATC	HHH	NAC	myosin_head	Anti_proliferat	BTK	MH2	DIL	3_5_exonuclease	ras	Clathrin	FAT	RBD	Patched	XPG_N	fer2	zf-B_box	PAS	BCL_N	TFR_dimer	PBX	GATase	notch	Glypican	I_LWEQ	pKID	HhH-GPD	FYRC	MoCF_biosynth
T	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
T	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
T	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
I	I	I	I	I	I	I	T	I	T	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	Ι	I	I	I	I	Ι	I	I	I	I	I	1
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
T	I	I	I	I	I	I	T	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
I	I	I	Ι	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
T	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I

cyclin	Q	SAP	Fz	BRCT	G-alpha	PI3_PI4_kinase	PSI	LRR	PBD	TNFR_c6	Cadherin_C_term	NUDIX	LRRCT	FCH	RhoGEF	PA	WH2	NACHT	MAM	UBX	PB1	BcI-2	BTB	zf-CXXC	DAG_PE-bind	COLFI	WD40	GATA	NAP	RasGAP	ig	CAP_GLY	Rap_GAP	Peptidase_C14	MH1	OAR	Frizzled	cyclin_C	zf-C3HC4
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	1
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	I	1	1
	1				•		•						•				•					•	•								•						•		
I	1	I	Ι	1	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	T	I	I	Ι	I	I	I	I	I	I	I	I	Ι	I	I
I	I	I	Ι	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	Ι	I	I
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	Ι	I	I	I	I	I	I	I	I	I	I	I	I

The probabilities of the	KRAB	efhand	Collagen	C2	PH	PDZ	сн	UCH	WW	HMG_box	ank	cadherin	Y_phosphatase	LRRNT	bZIP	histone	EGF	pkinase_C	RA	zf-C2H2	Sema	kazal	rrm	AMP-binding	cNMP_binding
e same or a greater n	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
umber of proteins with	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
a given domain occur	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
ing in a randomly sele	I	Ι	I	I	I	I	I	Ι	I	I	I	Ι	I	I	I	I	I	I	I	I	I	I	I	I	I
cted set of human p	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
roteins of the same s	I	Ι	Ι	Ι	Ι	I	ļ	Ι	Ι	Ι	I	Ι	Ι	I	Ι	I	Ι	I	I	I	I	I	ļ	I	Ι
ize under a biromial	I	I	I	I	I	I	I	I	I	I	I	I	1	I	I	I	I	I	1	I	I	I	I	I	I
distribution that were	I	I	I	I	I	I	Ι	Ι	I	I	Ι	I	I	I	I	I	I	I	I	I	I	I	Ι	I	I
e significant at 5% a	I	I	I	I	I	I	I	Ι	I	I	Ι	I	Ι	I	I	I	I	I	I	I	I	I	I	I	I
re shown.	I	Ι	Ι	I	Ι	I	I	I	I	I	Ι	Ι	I	I	I	I	I	I	I	I	I	I	I	I	I