Oncogene (2000) 19, 5548–5557

www.nature.com/onc

The protein tyrosine kinase family of the human genome

Dan R Robinson*,[1], Yi-Mi Wu[1] and Su-Fang Lin[1]

[1]*Department of Biological Chemistry, UC Davis School of Medicine, UC Davis Cancer Center, Sacramento, California, CA 95817, USA*

As the sequencing of the human genome is completed by the Human Genome Project, the analysis of this rich source of information will illuminate many areas in medicine and biology. The protein tyrosine kinases are a large multigene family with particular relevance to many human diseases, including cancer. A search of the human genome for tyrosine kinase coding elements identified several novel genes and enabled the creation of a nonredundant catalog of tyrosine kinase genes. Ninety unique kinase genes can be identified in the human genome, along with five pseudogenes. Of the 90 tyrosine kinases, 58 are receptor type, distributed into 20 subfamilies. The 32 nonreceptor tyrosine kinases can be placed in 10 subfamilies. Additionally, mouse orthologs can be identified for nearly all the human tyrosine kinases. The completion of the human tyrosine kinase family tree provides a framework for further advances in biomedical science. *Oncogene* (2000) **19,** 5548–5557.

Keywords: tyrosine kinase; human genome; receptor; non-receptor; mouse; ortholog

Introduction

The protein tyrosine kinases (PTKs) are a large and diverse multigene family found only in Metazoans. Their principal functions involve the regulation of multicellular aspects of the organism. Cell to cell signals concerning growth, differentiation, adhesion, motility, and death, are frequently transmitted through tyrosine kinases. In contrast, many of the serine/threonine kinase families, such as cyclin dependent kinases and MAP kinases, are conserved throughout eukaryotes and regulate processes in both unicellular and multicellular organisms. In humans, tyrosine kinases have been demonstrated to play significant roles in the development of many disease states, including diabetes and cancer. Historically, tyrosine kinases define the prototypical class of oncogenes, involved in most forms of human malignancies. Tyrosine kinase genes have also been linked to a wide variety of congenital syndromes (Robertson *et al*., 2000). Intensive study of this relevant gene family over the past 20 years has produced numerous insights into the structure, regulation, and function of these genes and their products. Many excellent reviews of protein tyrosine kinase structure and function have been published recently and are listed in Table 1.

Tyrosine kinases contain highly conserved catalytic domains similar to those in protein serine/threonine and dual-specificity kinases but with unique subdomain motifs clearly identifying members as tyrosine kinases (Hanks and Quinn, 1991). The high degree of conservation of the tyrosine kinase motifs has allowed the identification of tyrosine kinase genes in most metazoan phyla. Tyrosine kinase genes have been characterized in poriferans, cnidarians, nematodes, annelids, arthropods, echinoderms, and chordates (Suga *et al*., 1999; Muller *et al*., 1999; Miller *et al*., 2000; Rikke *et al*., 2000; Lucini *et al*., 1999; Sakuma *et al*., 1997). These sequences provide an extensive set of probes to investigate the genomic sequences of other species. Protein kinase catalogs of *Saccharomyces cerevesiae*, *Caenorhabditis elegans*, and *Drosophila melanogaster* have been compiled from the completed genomic sequences of these organisms (Hunter and Plowman, 1997; Plowman *et al*., 1999; Popovici *et al*., 1999; Morrison *et al*., 2000). Thanks to the efforts of all involved in the Human Genome Project, the draft sequence of the euchromatic regions of the human genome is nearly complete, an achievement of immeasurable significance. The accessibility of this vast information to researchers in the biological sciences should greatly facilitate further advances. The large well-studied family of protein tyrosine kinases provides a fitting framework for an initial analysis of the human genome sequence.

Searching the genome

To create a catalog of protein tyrosine kinase coding regions in the human genome, we performed iterative BLAST searches against the six-frame translations of the human genomic sequences available in the public databases (Altschul *et al*., 1990; Walchli *et al*., 2000). As probes, we used the amino acid sequences from the canonical tyrosine kinase domain of known and predicted tyrosine kinase genes of vertebrates, *C. elegans*, and *D. melanogaster*. A single accession was selected for each unique sequence and the translation was examined for protein kinase motifs and reciprocal BLAST searches were done to verify the sequence had highest similarity to protein tyrosine kinases and not other protein kinase families. The results are shown in Table 2a and b. The human genome, as currently sequenced, contains 90 tyrosine kinase genes and five presumed tyrosine kinase pseudogenes. Of the 90 tyrosine kinase genes, 58 are of the receptor type as defined by encoding a protein with a predicted transmembrane domain. These 58 receptor tyrosine kinases can be grouped into 20 subfamilies based on kinase domain sequence. The 32 non-receptor tyrosine kinases fall into 10 subfamilies based on kinase domain sequence. The remaining five sequences are classified as

*Correspondence: D Robinson, UC Davis Cancer Center, 4645 Second Ave., Sacramento, California, CA 95817, USA

Table 1 Overview of human tyrosine kinase literatures

Gene family	*References*
Summary	Hubbard and Till, 2000; Hunter, 2000; Robertson *et al*., 2000; Cole *et al*., 1999; Fischer, 1999; Schenk and Snaar-Jagalska, 1999; Hubbard *et al*., 1998; Porter and Vaillancourt, 1998; Hunter, 1997; Johnson *et al*., 1996; Hanks and Hunter, 1995
Non-receptor families	Neet and Hunter, 1996
ABL	Lanier and Gertler, 2000; Laneuville, 1995
ACK	Mott *et al*., 1999
CSK	Sondhi and Cole, 1999; Klages *et al*., 1994
FAK	Avraham *et al*., 2000; Girault *et al*., 1999; Schlaepfer and Hunter, 1998; Weisberg *et al*., 1997
FES	Smithgall *et al*., 1998; Pendergast, 1996
FRK	Lee, 1998; Vasioukhin and Tyner, 1997; Kohmura *et al*., 1994; Lee *et al*., 1994
JAK	Danial and Rothman, 2000; Imada and Leonard, 2000; Aringer *et al*., 1999; Leonard and O'Shea, 1998
SRC	Abram and Courtneidge, 2000; Schlessinger, 2000; Schwartzberg, 1998; Thomas and Brugge, 1997; Chow and Veillette, 1995
TEC	Schaeffer and Schwartzberg, 2000; Yang *et al*., 2000; Mano, 1999; Rawlings and Witte, 1995
SYK	Turner *et al*., 2000; Chu *et al*., 1998
Receptor families	Schlessinger and Ullrich, 1992
ALK	Ladanyi, 2000; Iwahara *et al*., 1997; Ladanyi, 1997
AXL	Crosier and Crosier, 1997; Neubauer *et al*., 1997
DDR	Weiner and Zagzag, 2000; Vogel, 1999; Schlessinger, 1997
EGFR	Klambt, 2000; Hackel *et al*., 1999; Kim and Muller, 1999; Wells, 1999
EPH	Mellitzer *et al*., 2000; Frisen *et al*., 1999; Holder and Klein, 1999; Bruckner and Klein, 1998; Flanagan and Vanderhaeghen, 1998; Xu and Wilkinson, 1997
FGFR	Gerwins *et al*., 2000; Klint and Claesson-Welsh, 1999
INSR	LeRoith, 2000; Whitehead *et al*., 2000; Baserga, 1999; Marino-Buslje *et al*., 1999
MET	van der Voort *et al*., 2000; Danilkovitch and Leonard, 1999; Birchmeier and Gherardi, 1998; Comoglio and Boccaccio, 1996
MUSK	Glass and Yancopoulos, 1997
PDGFR	Bourette and Rohrschneider, 2000; Shurin *et al*., 1998
PTK7	Park *et al*., 1996
RET	Mason, 2000; Eng, 1999; Eng and Mulligan, 1997
ROR	Takeuchi *et al*., 2000; Forrester *et al*., 1999
ROS	Yeung *et al*., 1998; Birchmeier *et al*., 1993
RYK	Serfas and Tyner, 1998
TIE	Partanen and Dumont, 1999; Tallquist *et al*., 1999
TRK	Hallbook, 1999; Kaplan and Miller, 1997; Barbacid, 1995; 1994
VEGFR	Gerwins *et al*., 2000; Veikkola *et al*., 2000; Neufeld *et al*., 1999
AATYK	Gaozza *et al*., 1997

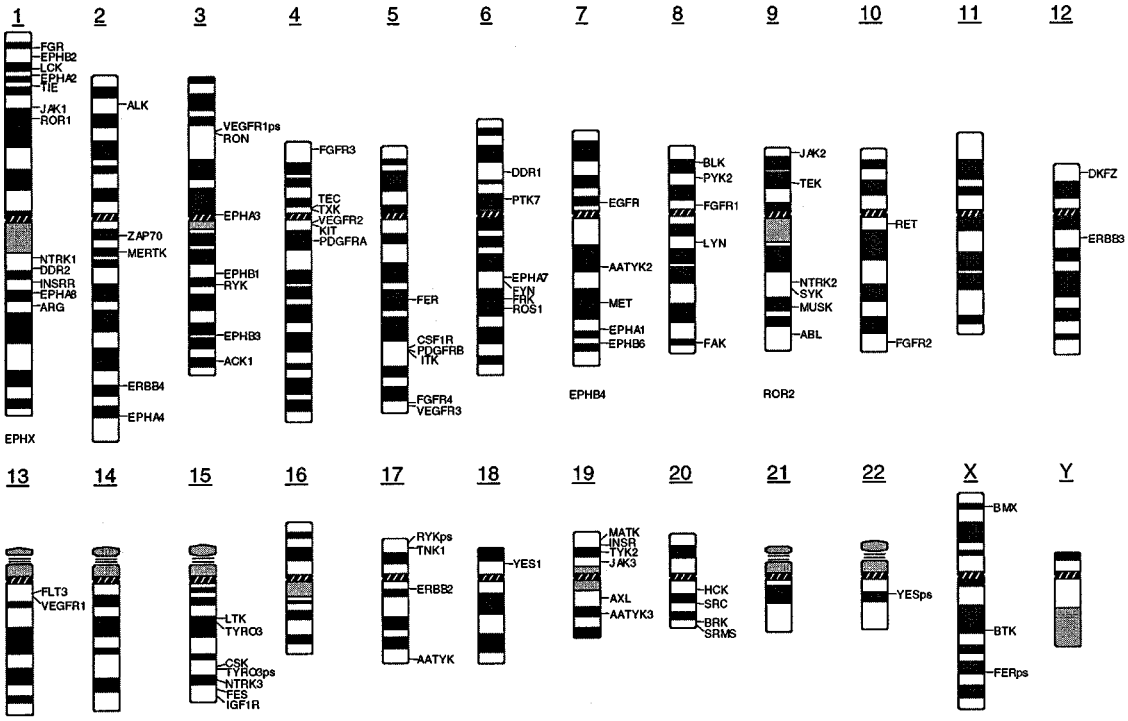

Figure 1 Distribution of protein tyrosine kinase genes in the human genome (G-banded chromosome ideograms, courtesy of David Adler, Department of Pathology, University of Washington, USA)

Table 2a Summary and classification of human non-receptor tyrosine kinases

Gene	*Synonyms*	*Hs NT ACC #*	*Hs PROT ACC#*	*Hs Unigene*	*Genomic ACC#*	*Chromosome*
ABL family						
ABL1	ABL	NM_005157	NP_005148	Hs.146355	AC073953	9q34.1
ARG	**ABL2**, ABLL	NM_07314	NP_005149	Hs.121521	AC024503	1q24-q25
ACK family						
ACK1	ACK2(b), Cdgip(m)	NM_005781	NP_005772	Hs.153937	AC009099	3q28-qter
TNK1		NM_003985	NP_003976	Hs.203420	AC002348	17p13.1
CSK family						
CSK	CYL	NM_004383	NP_004374	Hs.77793	AC020705	15q23-q24
MATK	CTK, HYL, CHK, LSK, Ntk(m)	NM_002378	NP_002369	Hs.274	AC005777	19p13.3
FAK family						
FAK	**PTK2**, Fadk(m)	NM_005607	NP_005598	Hs.740	AC067931	8q24-qter
PYK2	**PTK2B**, CAKbeta, RAFTK, FAK2, PKB	NM_004103	NP_004094	Hs.20313	AC010856	8p21.1
FES family						
FER	TYK3, Fert1/2(m)	NM_005246	NP_005237	Hs.121558	AC034224	5q21
FES	FPS	NM_002005	NP_001996	Hs.7636	AC003004	15q25-q26
FRK family						
BRK	**PTK6**, Sik(m)	NM_005975	NP_005966	Hs.51133	AL121829	20q13.3
FRK	RAK, Bsk(m), IYK(m)	NM_002031	NP_002022	Hs.89246	AC013613	6q21-q22
SRMS	SRM	AI373274 est		Hs.164442	AL121829	20q13.3
JAK family						
JAK1		NM_002227	NP_002218	Hs.50651	AL354886	1p32.3-p31.3
JAK2		NM_004972	NP_004963	Hs.115541	AL161450	9p24
JAK3	L-JAK	NM_000215	NP_000206	Hs.99877	AC007201	19p13.1
TYK2	JTK1	NM_003331	NP_003322	Hs.75516	AC011557	19p13.2
SRC-A family						
FGR	SRC2	NM_005248	NP_005239	Hs.1422	AL031729	1p36.2-p36.1
FYN	SLK, SYN	NM_002037	NP_002028	Hs.169370	AC073580	6q21
SRC		NM_005417	NP_005408	Hs.198298	AL133293	20q12-q13
YES1		NM_005433	NP_005424	Hs.194148	AC021474	18p11.31-p11.21
SRC-B family						
BLK		NM_001715	NP_001706	Hs.2243	AC022239	8p23-p22
HCK	JTK9, Bmk(m), HCTK	NM_002110	NP_002101	Hs.89555	AL049539	20q11-q12
LCK	Tck(m)	NM_005356	NP_005347	Hs.1765	AC022307	1p35-p34.3
LYN		NM_002350	NP_002341	Hs.80887	AC046176	8q13
TEC family						
BMX	ETK, PSCTK2	NM_001721	NP_001712	Hs.27372	AC003669	Xp22.2
BTK	ATK, PSCTK1, AGMX1, IMD1	NM_000061	NP_000052	Hs.159494	AL035422	Xq21.33-q22
ITK	EMT, Tsk(m), PSCTK2	NM_005546	NP_005537	Hs.211576	AC008389	5q31-q32
TEC	PSCTK4	NM_003215	NP_003206	Hs.89656	AC032007	4p12
TXK	PSCTK5, BTKL, Rlk(m)	NM_003328	NP_003319	Hs.29877	AC032007	4p12
SYK family						
SYK		NM_003177	NP_003168	Hs.74101	AC021581	9q22
ZAP70	SRK, STD	L05148	A44266	Hs.234569	AC016699	2q12
Pseudo genes						
FERps					AL030996	Xq25-q26
RYKps					AC012479	17p13.3
TYRO3 ps					AC010868	15q23-q24
VEGFR1ps					AC024739	3p21.3
YESps					AL022329	22q12.1

Table 2b Summary and classification of human receptor tyrosine kinases

ALK family						
ALK	Ki1	NM_004304	NP_004295	Hs.278572	AC019019	2p23
LTK	TYK1	NM_002344	NP_002335	Hs.210	AC011785	15q15.1-q21.1
AXL family						
AXL	UFO, Tyro7(r) Ark(m)	NM_001699	NP_001690	Hs.83341	AC011510	19q13.1
MER	**MERTK**, NYK, Eyk(ch)	NM_006343	NP_006334	Hs.78941	AC067761	2q14.1
TYRO3	RSE, SKY, BRT, DTK, TIF	NM_006293	NP_006284	Hs.301	AL353642	15q15.1-q21.1
DDR family						
DDR1	CAK, TRKE, NEP, NTRK4, EDDR1, PTK3	NM_013993	NP_001945	Hs.75562	AB023050	6p21.3
DDR2	TKT, TYRO10, NTRKR3	NM_006182	NP_006173	Hs.71891	AC021552	1q21-q22
EGFR family						
EGFR	ERBB, ERBB1	NM_005228	NP_005219	Hs.77432	AC073324	7p12
ERBB2	HER2, Neu(r), NGL	NM_004448	NP_004439	Hs.173664	AC025531	17q11.2-q12
ERBB3	HER3	NM_001982	NP_001973	Hs.199067	AC025162	12q13
ERBB4	HER4	NM_005235	NP_005226	Hs.1939	AC012069	2q33.3-q34
EPH family						
EPHA1	EPH, EPHT	NM_005232	NP_005223	Hs.89839	AF101169	7q32-q36
EPHA2	ECK, Sek2(m), Myk2(m)	NM_004431	NP_004422	Hs.171596	AC025928	1p34
EPHA3	HEK, ETK1, Tyro4(r), Mek4(m), Cek4(ch)	NM_005233	NP_005224	Hs.123642	AC048381	3p11.2
EPHA4	HEK8, Tyro1(r), Sek1(m), Cek8(ch)	NM_004438	NP_004429	Hs.73964	AC010899	2q36-qter
EPHA5	HEK7, Ehk1(r), Bsk(r), Cek7(ch)	L36644	P54756	Hs.31092	AC018683	

Continued

Table 2b (*Continued*)

Gene	*Synonyms*	*Hs NT ACC #*	*Hs PROT ACC#*	*Hs Unigene*	*Genomic ACC#*	*Chromosome*
EPHA6	DKFZp434C1418, Ehk2(r)	AL133666		Hs.145172	AC011783	
EPHA7	HEK11, Mdk1(m), Ebk(m), Ehk3(r), Cek11(ch)	NM_004440	NP_004431	Hs.73962	AC023873	6q21
EPHA8	HEK3, KIAA1459, Eek(r), Cek10(ch)	AB040892	CAB81612	Hs.145731	AL035703	1q23-q24
EPHB1	NET, EPHT2, HEK6, Elk(r), Cek6(ch)	NM_004441	NP_004432	Hs.78436	AC026606	3q21-q23
EPHB2	HEK5, ERK, DRT, EPHT3, Tyro5(r), Nuk(m), Sek3(m), Cek5(ch)	AF025304	AAB94602	Hs.125124	AL035704	1p36.1-p35
EPHB3	HEK2, Tyro6, Mdk5(m), Sek4(m)	NM_004443	NP_004434	Hs.2913		3q21-qter
EPHB4	HTK, Tyro11(r), Mdk2(m), Myk1(m)	NM_004444	NP_004435	Hs.155227	AC011895	7
EPHB6	HEP, Mep(m), Cek1(ch)	NM_004445	NP_004436	Hs.3796	AF107256	7q33-q35
EPHX					AC023225	1
FGFR family						
FGFR1	FLT2, bFGFR, FLG, N-SAM	M34641	AAA35835	Hs.748	AC011237	8p11.2
FGFR2	KGFR, K-SAM, Bek(m), CFD1, JWS, Cek3(ch)	NM_000141	NP_000132	Hs.278581	AC009988	10q26
FGFR3	HBGFR, ACH, Cek2(ch)	NM_000142	NP_000133	HS.1420	AC016773	4p16.3
FGFR4		NM_002011	NP_002002	Hs.165950	AC008570	5q35.1-qter
INSR family						
IGF1R	JTK13	NM_000875	NP_000866	Hs.239176	AC069029	15q25-q26
INSR	IR	NM_000208	NP_000199	Hs.89695	AC010606	19p13.3-p13.2
INSRR	IRR	J05046	AAC31759	Hs.248138		1q21-q23
MET family						
MET	HGFR	NM_000245	NP_000236	Hs.81688	AC002080	7q31
RON	**MST1R**, CDw136, Fv2(m), STK(m), SEA(ch)	NM_002447	NP_002438	Hs.2942	AC068701	3p21.3
MUSK family						
MUSK	Nsk2(m), Mlk1(m), Mlk2(m)	NM_005592	NP_005583	Hs.156465	AL157881	9q31.3-q32
PDGFR family						
CSF1R	FMS, C-FMS, CD115	NM_005211	NP_005202	Hs.174172	AC010899	5q31-q32
FLT3	FLK2, STK1, CD135	NM_004119	NP_0041110	Hs.385	AL356915	13q12
KIT	Sfr(m), CKIT	NM_000222	NP_000213	Hs.81665	AC006553	4q11-q12
PDGFRA		NM_006206	NP_006197	Hs.74615	AC025013	4q11-q13
PDGFRB	PDGFR, JTK12	NM_002609	NP_002600	Hs.76144	AC010899	5q31-q32
PTK7 family						
PTK7	CCK4, KLG(ch)	NM_002821	NP_002812	Hs.90572	AL355385	6p21.1-p12.2
RET family						
RET	MEN2A/B, HSCR1, MTC1	X12949	P07949	Hs.251479	AC010864	10q11.2
ROR family						
ROR1	NTRKR1	NM_005012	NP_005003	Hs.274273	AL138793	1p32-p31
ROR2	NTRKR2	NM_004560	NP_004551	Hs.155585		9
ROS family						
ROS1	MCF3	NM_002944	NP_002935	Hs.1041	AL354945	6q22
RYK family						
RYK	Vik(m), Mrk(m)	S59184	AAB26341	Hs.79350		3q22
TIE family						
TEK	TIE2	NM_000459	NP_000450	Hs.89640	AL133411	9p21
TIE	TIE1, JTK14	NM_005424	NP_005415	Hs.78824	AL139289	1p34-p33
TRK family						
NTRK1	TRK, TRKA	NM_002529	NP_002520	Hs.85844	AL158169	1q21-q22
NTRK2	TRKB	NM_006180	NP_006171	Hs.47860	AC007524	9q22.1
NTRK3	TRKC	NM_002530	NP_002521	Hs.26776	AC011966	15q25
VEGFR family						
VEGFR1	**FLT1**	NM_002019	NP_002010	Hs.138671	AL138712	13q12
VEGFR2	**KDR**, FLK1		AAB88005	Hs.12337	AC021220	4q11-q12
VEGFR3	**FLT4**, PCL	NM_002020	NP_002011	Hs.74049	AC022095	5q34-q35
AATYK family						
AATYK	**AATK**, KIAA0641	NM_004920	NP_004911	Hs.128316		17q25.3
AATYK2	**KIAA1079**	NM_014916	NP_055731	Hs.122708	AC025605	7q21-q22
AATYK3					AC008403	19q13.2-q13.3
Uncharacterized						
DKFZp761P1010		NM_018423	NP_060893	Hs.24979	AC021049	12p13

pseudogenes by the lack of introns in the sequence, the truncation of the coding regions compared to other members of the family, the presence of in-frame termination codons, and the absence of evidence for expression. Genomic sequences for all but five of the known tyrosine kinase genes can be found in the current GenBank databases, consistent with the predicted coverage of the human genome with the present BAC (Bacterial Artificial Chromosome) tiling and sequencing status. This allows the possibility that a very small number of protein tyrosine kinase coding elements remain undetected at present.

The information in Table 2a and b is presented in a way to facilitate access to the tools and databases at the NCBI Web site (http://www.ncbi.nlm.nih.gov). Gene symbols in the first column are those most

Table 3a Assignment of human non-receptor tyrosine kinases and vertebrate orthologs

Human	*MUS MGD*	*MUS ACC#*	*% Ident/%SIM*	*MUS Unigene*	*Rattus Name/ACC#*	*Gallus Name/ACC#*
ABL family						
ABL1	**Abl**	L10656	98/99	Mm.1318		ABL/AF041794
ARG	Arg	AAA98465(P)	99/100			
ACK family						
ACK1	**Cdgip**	NM_016788	98/99	Mm.1483		
TNK1	Tnk1	AA727750 est		Mm.28874		
CSK family						
CSK	**Csk**	NM_007783	97/100	Mm.21974	CSK/X58631	CSK/M85039
MATK	**Matk**	NM_010768	92/97	Mm.2918	BATK/L34542	
FAK family						
FAK	**Fadk**	NM_007982	99/99	Mm.29674	PTK2/NM_013081	FAK/M86656
PYK2	**Ptk2b**	AA537970	98/100	Mm.21613	CAKbeta/NM_017318	
FES family						
FER	Fer	U76762	93/100	Mm.5175	FLK/X13412	
FES	**Fes**	X12616	96/97	Mm.3394		FPS/M11611
FRK family						
BRK	**Ptk6**	NM_009184	82/94	Mm.4497		
FRK	**Frk**	NM_010237	95/100	Mm.4953	FRK/U09583	
SRMS	**Srms**	NM_011481	84/94	Mm.4752		
JAK family						
JAK1	**Jak1**	S63728	97/99	Mm.28598		JAK1/AF096264
JAK2	**Jak2**	NM_008413	97/99	Mm.809	JAK2/U13396	
JAK3	**Jak3**	NM_010589	89/94	Mm.4181	JAK3/NM_012855	JAK/AF034576
TYK2	**Tyk2**	NM_018793	91/95	Mm.20249		TYK2/AF041801
SRC-A family						
FGR	**Fgr**	NM_010208	91/98	Mm.1338	FGR/X57018	
FYN	**Fyn**	NM_008054	96/98	Mm.4848	FYN/NM_012755	FYN/X52841
SRC	**Src**	NM_009271	99/99	Mm.22845	SRC/AF130457	SRC/V00402 YRK/X67786
YES1	**Yes**	NM_009535	99/100	Mm.4558		YES/X12461
SRC-B family						
BLK	**Blk**	NM_007549	93/99	Mm.3962		
HCK	**Hck**	NM_010407	94/99	Mm.715	HCK/NM_013185	
LCK	**Lck**	X03533	97/98	Mm.142		
LYN	**Lyn**	M64608	98/98	Mm.1834	LYN/L14951	
TEC family						
BMX	**Bmx**	NM_009759	94/99	Mm.504		
BTK	**Btk**	NM_013482	98/100	Mm.4475		
ITK	**Itk**	ITK/NM_010583	96/97	Mm.16009		
TEC	**Tec**	NM_013689	96/96	Mm.2350		PROKINC/L20624
TXK	**Txk**	NM_013698	87/94	Mm.3264		
SYK family						
SYK	**Syk**	NM_011518	98/100	Mm.4708	SYK/NM_012758	
ZAP70	**Zap70**	NM_009539	94/98	Mm.8038		

Table 3b Assignment of human receptor tyrosine kinases and vertebrate orthologs

ALK family						
ALK	**Alk**	NM_007439	98/99	Mm.2536		
LTK	**Ltk**	NM_008523	88/98	Mm.1740		
AXL family						
AXL	**Axl**	NM_009465	97/99	Mm.4128	AXL/AF04688	
MER	**Mer**	NM_008587	93/97	Mm.4582	MERTK/AF208235	EYK/L21719
TYRO3	**Tyro3**	NM_019392	98/98	Mm.2901	TYRO3/NM_017092	REK/U70045
DDR family						
DDR1	**Cak**	NM_007584	92/99	Mm.5021	CAK/NM_013137	
DDR2	**Ddr2**	NM_008746	93/98	Mm.4999	TYRO10/AF016247	TYRKINA/L20622
EGFR family						
EGFR	**Egfr**	X78987	98/100	Mm.8534	EGFR/M37394	ERBB/M10066
ERBB2	**Erbb2**	L47239	99/99 (rat)		ERBB2/NM_017003	
ERBB3	**Erbb3**	AF059175	95/98 (rat)	Mm.57112	ERBB3/NM_017218	
ERBB4	**Erbb4**	AF059177	96/99 (rat)	Mm.57113	ERBB4/AF041838	ERBB4/AF121963
EPH family						
EPHA1	**Epha1**	U18084	93/98	Mm.133330		
EPHA2	**Epha2**	NM_010139	94/97	Mm.2581		
EPHA3	**Epha3**	M68513	97/100	Mm.1977	REK4/U69278	CEK4/M68514
EPHA4	**Epha4**	NM_007936	100/100	Mm.3249		CEK8/Z19059
EPHA5	**Epha5**	NM_007937	98/100	Mm.4466	EHK1/X78689	CEK7/U03910
EPHA6	**Epha6**	NM_007938	98/99	Mm.4264	EHK2/P54758	
EPHA7	**Epha7**	X79082	99/100	Mm.4806	EHK3/U21954	CEK11/Y14271
EPHA8	**Epha8**	NM_007939	93/98	Mm.1390	EEK/X59290	CEK10/Z19061
EPHB1	**Ephb1**	AA058194 est	99/100 (rat)	Mm.125507	ELK/M59814	CEK6/Z19110
EPHB2	**Ephb2**	L25890	100/100	Mm.4652		CEK5/M62325
EPHB3	**Ephb3**	Z49086	99/100	Mm.6972		
EPHB4	**Ephb4**	NM_010144	99/100	Mm.34533		
EPHB5						CEK9/U23783

Continued

Table 3b (*Continued*)

Human	*MUS MGD*	*MUS ACC#*	*% Ident/%SIM*	*MUS Unigene*	*Rattus Name/ACC#*	*Gallus Name/ACC#*
EPHB6	**Ephb6, Cek1**	NM_007680	93/96	Mm.1480		
EPHX						
FGFR family						
FGFR1	**Fgfr1**	NM_010206	100/100	Mm.3157	FGFR1/D1249	CEK1/M24637
FGFR2	**Fgfr2**	NM_010207	99/100	Mm.16340	FGFR2/Z35138	
FGFR3	**Fgfr3**	NM_008010	97/99	Mm.6904	FGFR3/AF277717	CEK2/M35195
FGFR4	**Fgfr4**	NM_008011	97/98	Mm.4912	FGFR4/M91599	
INSR family						
IGF1R	**lgf1r**	AF056187	95/97	Mm.10226	IGF1R/L29232	IGF1R/AJ223164
INSR	**Insr**	NM_010568	96/99	Mm.46791	INSR/NM_017071	
INSRR	**Insrr**	NM_011832	94/98	Mm.42041		INSRR/AF041799
MET family						
MET	**Met**	NM_008591	98/99	Mm.86844	HGFR/U65007	MET/X84044
RON	**Ron**	NM_009074	89/93	Mm.3901		SEA/L12024
MUSK family						
MUSK	**Nsk2**	NM_010944	97/98	Mm.16148	MUSK/U34985	
PDGFR family						
CSF1R	**Csf1r**	NM_007779	93/99	Mm.22574	CSF1R/X61479	FMS/L20625
FLT3	**Flt3**	NM_010229	92/99	Mm.194		
KIT	**Kit**	Y00864	94/96	Mm.4394	KIT/D12524	KIT/D13225
PDGFRA	**Pdgfra**	NM_011058	98/99	Mm.2924	APDGFR/M63837	PDGFRA/AF188842
PDGFRB	**Pdgfrb**	NM_008809	94/99	Mm.4146		
PTK7 family						
PTK7		BE306618 est	83/97 (ch)			KLG/M63437
RET family						
RET	**Ret**	NM_009050	95/97	Mm.57199		RET/Z49898
ROR family						
ROR1	**Ntrkr1**	NM_013845	98/99	Mm.57252		
ROR2	**Ntrkr2**	NM_013846	95/98	Mm.86922		
ROS family						
ROS1	**Ros1**	U15443	92/97	Mm.4155	ROS1/NM_012874	ROS/X06770
RYK family						
RYK	**Ryk**	M98547	97/98	Mm.3860		RYK/AF041796
TIE family						
TEK	**Tek**	NM_013690	99/100	Mm.14313		
TIE	**Tie1**	NM_011587	99/100	Mm.4345		
TRK family						
NTRK1	**Ntrk1**	AI429726 est	94/94 (rat)		TRK/M85214	TRKA/X93581
NTRK2	**Ntrk2**	NM_008745	99/100 (rat)	Mm.3993	NTRK2/NM_012731	TRKB/X77251
NTRK3	**Ntrk3**	AF035399	99/99 (rat)	Mm.20467	NTRK3/NM_019248	TRKC/S74248
VEGFR family						
VEGFR1	**Flt1**	NM_010228	91/97	Mm.3464	FLT1/D28498	
VEGFR2	**Kdr**	NM_010612	96/99	Mm.285	KDR/NM_013062	
VEGFR3	**Flt4**	NM_008029	91/92	Mm.3291		VEGFR3/AF041795
AATYK family						
AATYK	**Aatk**	NM_007377	94/97	Mm.6826		
AATYK2					X94514 est	
AATYK3						
Uncharacterized						
DKFZp761P1010		AI386314 est				

frequently used recently in the PubMed literature database. Synonyms used in the literature for human and vertebrate orthologs are listed in the second column. Symbols used in non-human species are followed by a species designation of (m) *Mus musculus*, (r) *Rattus norvegicus*, and (ch) *Gallus gallus*. Gene symbols in bold are those approved by HUGO Gene Nomenclature Committee (White *et al*., 1997). These approved symbols are used in the LocusLink curated database of linked information on defined genes (Maglott *et al*., 2000; Pruitt *et al*., 2000). Where possible, reference sequences (RefSeq) accession numbers are given for nucleotide and protein sequences. Unigene cluster numbers are given for access to Unigene EST database and related expression links (Boguski and Schuler, 1995). Genomic accession numbers can be used in conjunction with Mapviewer for graphical access to the human genome data.

Newly identified human tyrosine kinases

Five novel human kinase sequences were identified in the genome-wide search. The first is the human ortholog of the EphA6 gene of rodents. Interestingly, the human transcript as defined by several ESTs contains a truncated reading frame produced by a cryptic splice in kinase subdomain VIII. The genomic locus confirms the validity of the EST transcripts and contains putative unused exons for subdomains IX–XI. Whether the altered human EPHA6 gene product retains tyrosine kinase activity is an open question. A second potential new member of the Eph family is found on chromosome 1 and is designated as EPHX. The EPHX sequences in the genome retain the basic intron/exon structure of the Eph family kinase domains, and the predicted protein sequence is similar to the other members of the Eph family, although

divergent at several conserved elements. No evidence for the transcription of this locus exists in the databases and EPHX may be a vestigial gene sequence. Another novel sequence uncovered, AATYK3, is a third member of a class of tyrosine kinases defined by the AATYK and KIAA1079 genes. AATYK3 is highly similar to the other members of the family in both kinase and non-kinase domains. The fourth novel sequence is the human ortholog of the murine Srms non-receptor tyrosine kinase and the predicted reading frame encodes a protein with high similarity to Srms in both kinase and non-kinase domains. Finally, a novel sequence with tyrosine kinase homology is found on chromosome 12 and a partial cDNA sequence exists, DKFZp761P1010. The predicted reading frame encodes a protein with weak similarity to the kinase domains of fibroblast growth factor receptors, a transmembrane domain, but lacking a signal peptide. No other potential members of this class of tyrosine kinases were detected in the human genome.

There are several tyrosine kinases found in other species for which human orthologs could not be found. The large Kin 15/Kin 16 class of receptor tyrosine kinase genes in the *C. elegans* genome has no identifiable orthologs in the human genome. In chickens, the src family member Yrk and the eph family member EphB5 (Cek9) lack identified mammalian orthologs. One possibility is that the EPHX sequences in the human genome are from the ancestral EphB5 gene. Finally, no sequences for the proposed human ZRK zona pelucida kinase can be detected in the human genome and the ZRK clone is most likely derived from the human MER gene (Tsai and Silver, 1996).

Distribution of human tyrosine kinase genes

The cytogenetic distribution of tyrosine kinase genes in the human genome is depicted in Figure 1. Tyrosine kinase genes can be found on 19 of the 24 human chromosomes. At least three locations show evidence of more recent gene duplication events. On chromosome 5, the highly similar PDGFRB and CSF1R genes are adjacent and contained within a single BAC clone. On chromosome 20, the FRK family members BRK and SRMS are contained within a single BAC. On chromosome 4, the TEC and TXK genes can also be found within a single BAC clone.

Classification of human tyrosine kinases

A phylogenetic analysis of the amino acid sequences of the kinase domains from the identified human tyrosine kinases is shown in Figure 2. The human tyrosine kinases may be grouped into 20 receptor and 10 non-receptor classes as marked on the right of the figure. Grouping tyrosine kinase genes by intron/exon structure closely parallels the results obtained from phylogenetic sequence analysis. For example, phylogenetic analysis of amino acid sequence places the two CSK family members apart from the three FRK family members and the eight SRC family members, even though all three families share the same overall protein domain structure (Figure 3a) (Lee *et al*., 1998). This distinction of the families is verified by a different intron/exon organization for each family. Within a tyrosine kinase family, the members of a single family exhibit a common intron/exon pattern, distinct from other families.

The clustering of tyrosine kinase genes into families based on kinase domain sequence also parallels the overall domain structure of the proteins. A diagram of the overall protein domain structure of a representative member of each tyrosine kinase family is shown in Figure 3a,b. The human tyrosine kinase families exhibit a wide spectrum of protein domains consistent with the numerous and varied interactions and functions of these molecules as reviewed by the publications in Table 1. The domain representations are not complete. Many families encode proteins for which some domains are not defined, given the current state of knowledge.

Mouse orthologs of human tyrosine kinase genes

The assignment of vertebrate orthologs to the human tyrosine kinase genes was done by reciprocal BLAST

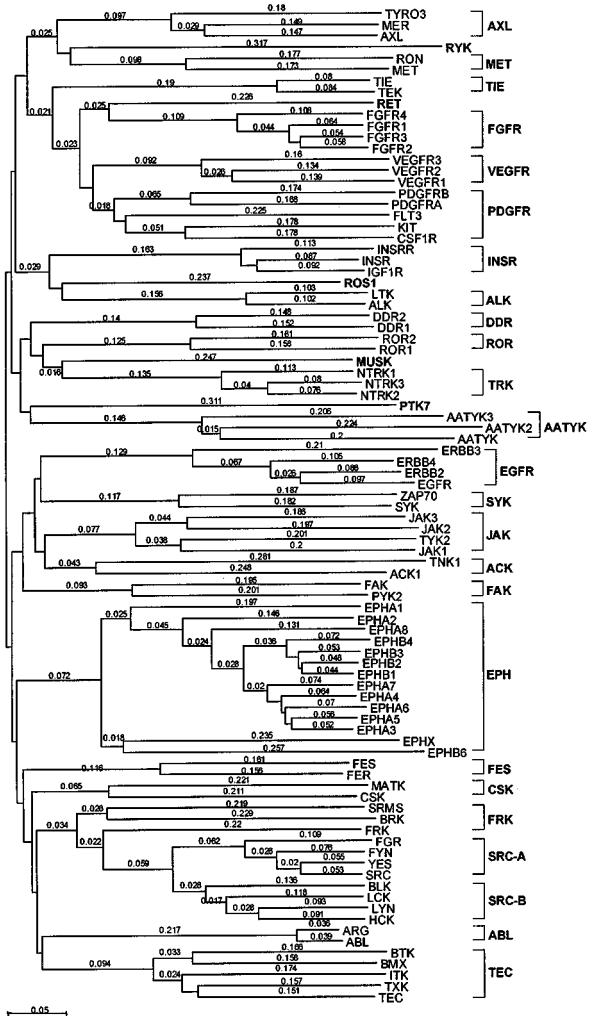

Figure 2 Phylogram of the human protein tyrosine kinase family inferred from amino acid sequences of the kinase domains. The tree is constructed by the N-J method (Saitou and Nei, 1987) and the evolutionary distance is calculated by Tamura-Nei algorithm (Tamura and Nei, 1993). Numbers on each node indicate the evolutionary distance. The tree is drawn to scale and is midpoint-rooted

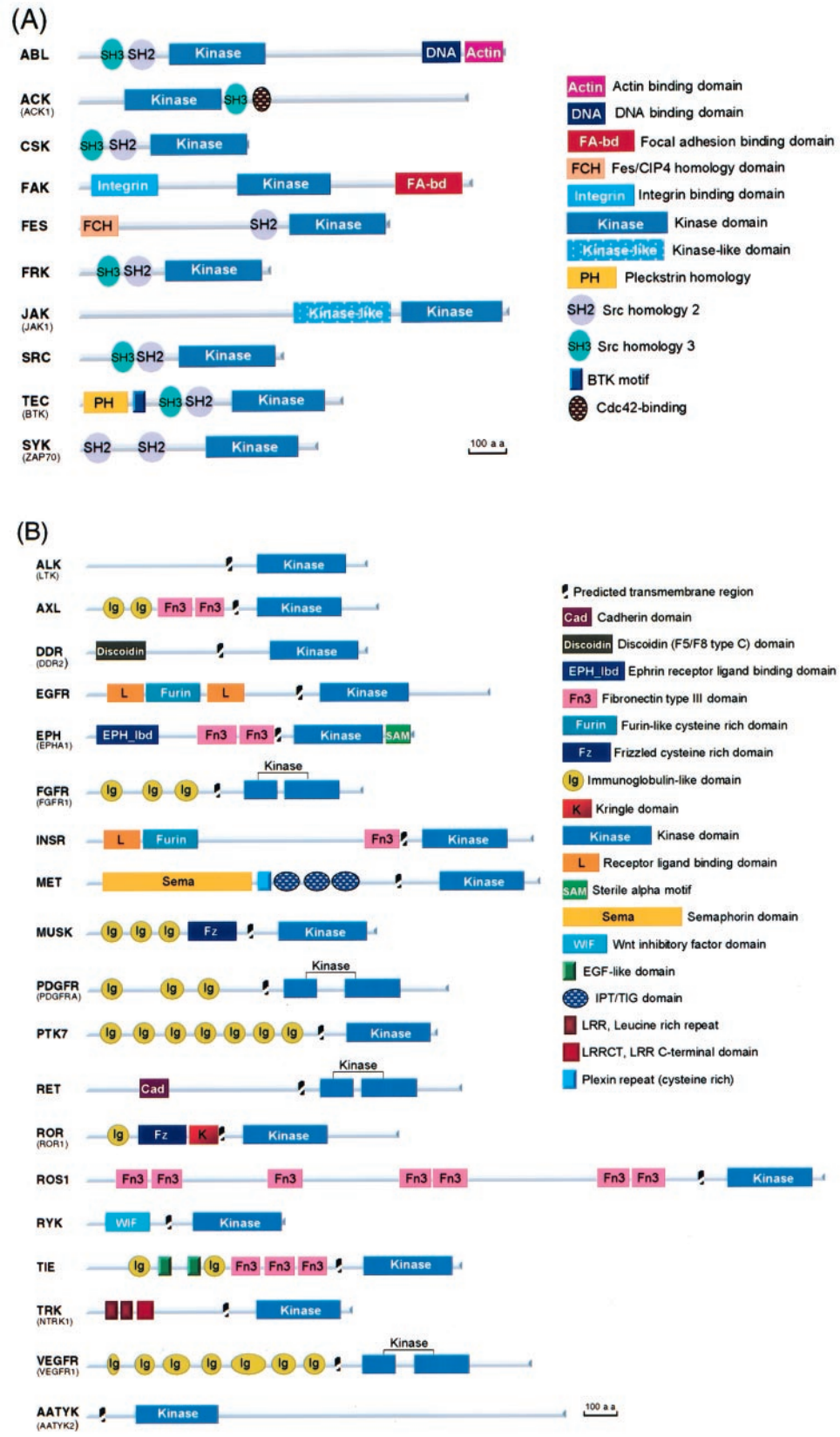

Figure 3 Domain structures of the human non-receptor tyrosine kinases (**a**) and receptor tyrosine kinases (**b**). Data were obtained by searching the amino acid sequences of the human tyrosine kinases against the latest version (5.5) of the Pfam database (Bateman *et al*., 2000). Domains sharing significant homologies to the Pfam-A alignments are used in this study. The schematics are shown to scale

searches between the human, mouse, rat, and chicken sets of tyrosine kinase genes. Pairs with reciprocal best scores were assigned as orthologs and the results are shown in Table 3a and b (Walchi *et al*., 2000). The mouse gene symbols used are in accordance with the MGD Nomenclature Committee and LocusLink (Blake *et al*., 2000). The per cent identity and similarity for the protein kinase domains of the mouse and human orthologs are shown in the fourth column. Cluster numbers for the murine Unigene EST database are given. Accession numbers for identified rat and chicken orthologs complete the tables. A nearly complete correspondence between the tyrosine kinase families of man and mouse exists. There are only three human PTKs for which an orthologous murine sequence has not been discovered, and this gap is likely to be filled quickly. Conversely, there are no identified rodent tyrosine kinases for which a human ortholog does not exist. For the non-mammalian vertebrates, such as chicken, a limited number of tyrosine kinases may not have human orthologs, as discussed previously. This correspondence between the tyrosine kinase gene families of humans and mice reinforces the validity of mouse models for human diseases, especially cancer.

The application of the human genomic sequence information should greatly aid the study of protein tyrosine kinases. As a first step, the catalog of tyrosine kinase genes in the human genome should provide a foundation for further discovery of kinase involvement in disease progression and functional characterizations. The fact that the range of human protein tyrosine kinases was nearly identified prior to the genomic sequence information is a testament to the interest shown and quality of work performed by the numerous researchers devoted to protein kinases. Other less well-studied gene families are likely to have more surprises in the human genome. Continuing studies on expression patterns, functional characterizations, and disease associations of tyrosine kinases, as well as studies of genetic variations at tyrosine kinase loci, should provide a basis for the development of new therapeutics for the treatment of human disease.

Acknowledgments

We would like to thank Dr H-J Kung for valuable discussions and support. We acknowledge all those whose relevant contributions we were unable to cite in this review due to space limitations. This work was supported by NIH grant ROI CA82073.

References

Abram CL and Courtneidge SA. (2000). *Exp. Cell Res.,* **254,** 1–13.
Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ. (1990). *J. Mol. Biol.,* **215,** 403–410.
Aringer M, Cheng A, Nelson JW, Chen M, Sudarshan C, Zhou YJ and O'Shea JJ. (1999). *Life Sci.,* **64,** 2173–2186.
Avraham H, Park SY, Schinkmann K and Avraham S. (2000). *Cell Signal,* **12,** 123–133.
Barbacid M. (1994). *J. Neurobiol.,* **25,** 1386–1403.
Barbacid M. (1995). *Ann. NY Acad. Sci.,* **766,** 442–458.
Baserga R. (1999). *Exp. Cell Res.,* **253,** 1–6.
Bateman A, Birney E, Durbin R, Eddy SR, Howe KL and Sonnhammer EL. (2000). *Nucleic Acids Res.,* **28,** 263–266.
Birchmeier C and Gherardi E. (1998). *Trends Cell Biol.,* **8,** 404–410.
Birchmeier C, Sonnenberg E, Weidner KM and Walter B. (1993). *Bioessays,* **15,** 185–190.
Blake JA, Eppig JT, Richardson JE and Davisson MT. (2000). *Nucleic Acids Res.,* **28,** 108–111.
Boguski MS and Schuler GD. (1995). *Nat. Genet.,* **10,** 369–371.
Bourette RP and Rohrschneider LR. (2000). *Growth Factors,* **17,** 155–166.
Bruckner K and Klein R. (1998). *Curr. Opin. Neurobiol.,* **8,** 375–382.
Chow LM and Veillette A. (1995). *Semin. Immunol.,* **7,** 207–226.
Chu DH, Morita CT and Weiss A. (1998). *Immunol. Rev.,* **165,** 167–180.
Cole PA, Sondhi D and Kim K. (1999). *Pharmacol. Ther.,* **82,** 219–229.
Comoglio PM and Boccaccio C. (1996). *Genes Cells,* **1,** 347–354.
Crosier KE and Crosier PS. (1997). *Pathology,* **29,** 131–135.
Danial NN and Rothman P. (2000). *Oncogene,* **19,** 2523–2531.
Danilkovitch A and Leonard EJ. (1999). *J. Leukoc. Biol.,* **65,** 345–348.
Eng C. (1999). *J. Clin. Oncol.,* **17,** 380–393.
Eng C and Mulligan LM. (1997). *Hum. Mutat.,* **9,** 97–109.
Fischer EH. (1999). *Adv. Enzyme Regul.,* **39,** 359–369.
Flanagan JG and Vanderhaeghen P. (1998). *Annu. Rev. Neurosci.,* **21,** 309–345.
Forrester WC, Dell M, Perens E and Garriga G. (1999). *Nature,* **400,** 881–885.
Frisen J, Holberg J and Barbacid M. (1999). *EMBO J.,* **18,** 5159–5165.
Gaozza E, Baker SJ, Vora RK and Reddy EP. (1997). *Oncogene,* **15,** 3127–3135.
Gerwins P, Skoldenberg E and Claesson-Welsh L. (2000). *Crit. Rev. Oncol. Hematol.,* **34,** 185–194.
Girault JA, Costa A, Derkinderen P, Studler JM and Toutant M. (1999). *Trends Neurosci.,* **22,** 257–263.
Glass DJ and Yancopoulos GD. (1997). *Curr. Opin. Neurobiol.,* **7,** 379–384.
Hackel PO, Zwick E, Prenzel N and Ullrich A. (1999). *Curr. Opin. Cell Biol.,* **11,** 184–189.
Hallbook F. (1999). *Curr. Opin. Neurobiol.,* **9,** 616–621.
Hanks SK and Hunter T. (1995). *FASEB J.,* **9,** 576–596.
Hanks SK and Quinn AM. (1991). *Methods Enzymol.,* **200,** 38–62.
Holder N and Klein R. (1999). *Development,* **126,** 2033–2044.
Hubbard SR, Mohammadi M and Schlessinger J. (1998). *J. Biol. Chem.,* **273,** 11987–11990.
Hubbard SR and Till JH. (2000). *Annu. Rev. Biochem.,* **69,** 373–398.
Hunter T. (1997). *Cell,* **88,** 333–346.
Hunter T. (2000). *Cell,* **100,** 113–127.
Hunter T and Plowman GD. (1997). *Trends Biochem. Sci.,* **22,** 18–22.
Imada K and Leonard WJ. (2000). *Mol. Immunol.,* **37,** 1–11.
Iwahara T, Fujimoto J, Wen D, Cupples R, Bucay N, Arakawa T, Mori S, Ratzkin B and Yamamoto T. (1997). *Oncogene,* **14,** 439–449.
Johnson LN, Noble ME and Owen DJ. (1996). *Cell,* **85,** 149–158.

Kaplan DR and Miller FD. (1997). *Curr. Opin. Cell Biol.,* **9,** 213–221.
Kim H and Muller WJ. (1999). *Exp. Cell Res.,* **253,** 78–87.
Klages S, Adam D, Class K, Fargnoli J, Bolen JB and Penhallow RC. (1994). *Proc. Natl. Acad. Sci. USA,* **91,** 2597–2601.
Klambt C. (2000). *Curr. Biol.,* **10,** R388–R391.
Klint P and Claesson-Welsh L. (1999). *Front. Biosci.,* **4,** D165–D177.
Kohmura N, Yagi T, Tomooka Y, Oyanagi M, Kominami R, Takeda N, Chiba J, Ikawa Y and Aizawa S. (1994). *Mol. Cell. Biol.,* **14,** 6915–6925.
Ladanyi M. (1997). *Cancer Surv.,* **30,** 59–75.
Ladanyi M. (2000). *Am. J. Pathol.,* **157,** 341–345.
Laneuville P. (1995). *Semin. Immunol.,* **7,** 255–266.
Lanier LM and Gertler FB. (2000). *Curr. Opin. Neurobiol.,* **10,** 80–87.
Lee H, Kim M, Lee KH, Kang KN and Lee ST. (1998). *Mol. Cells,* **8,** 401–407.
Lee J, Wang Z, Luoh SM, Wood WI and Scadden DT. (1994). *Gene,* **138,** 247–251.
Leonard WJ and O'Shea JJ. (1998). *Annu. Rev. Immunol.,* **16,** 293–322.
LeRoith D. (2000). *Endocrinology,* **141,** 1287–1288.
Lucini C, Castaldo L, Lamanna C, Maruccio L, Vega JA and Gargiulo G. (1999). *Neurosci. Lett.,* **261,** 163–166.
Maglott DR, Katz KS, Sicotte H and Pruitt KD. (2000). *Nucleic Acids Res.,* **28,** 126–128.
Mano H. (1999). *Cytokine Growth Factor Rev.,* **10,** 267–280.
Marino-Buslje C, Martin-Martinez M, Mizuguchi K, Siddle K and Blundell TL. (1999). *Biochem. Soc. Trans.,* **27,** 715–726.
Mason I. (2000). *Pharm. Acta. Helv.,* **74,** 261–264.
Mellitzer G, Xu Q and Wilkinson DG. (2000). *Curr. Opin. Neurobiol.,* **10,** 400–408.
Miller MA, Malik IA, Shenk MA and Steele RE. (2000). *Oncogene,* **19,** 3925–3930.
Morrison DK, Murakami MS and Cleghon V. (2000). *J. Cell Biol.,* **150,** F57–F62.
Mott HR, Owen D, Nietlispach D, Lowe PN, Manser E, Lim L and Laue ED. (1999). *Nature,* **399,** 384–388.
Muller WE, Kruse M, Blumbach B, Skorokhod A and Muller IM. (1999). *Gene,* **238,** 179–193.
Neet K and Hunter T. (1996). *Genes Cells,* **1,** 147–169.
Neubauer A, Burchert A, Maiwald C, Gruss HJ, Serke S, Huhn D, Wittig B and Liu E. (1997). *Leuk. Lymphoma,* **25,** 91–96.
Neufeld G, Cohen T, Gengrinovitch S and Poltorak Z. (1999). *FASEB J.,* **13,** 9–22.
Park SK, Lee HS and Lee ST. (1996). *J. Biochem. (Tokyo),* **119,** 235–239.
Partanen J and Dumont DJ. (1999). *Curr. Top. Microbiol. Immunol.,* **237,** 159–172.
Pendergast AM. (1996). *Curr. Opin. Cell Biol.,* **8,** 174–181.
Plowman GD, Sudarsanam S, Bingham J, Whyte D and Hunter T. (1999). *Proc. Natl. Acad. Sci. USA,* **96,** 13603–13610.
Popovici C, Roubin R, Coulier F, Pontarotti P and Birnbaum D. (1999). *Genome Res.,* **9,** 1026–1039.
Porter AC and Vaillancourt RR. (1998). *Oncogene,* **17,** 1343–1352.
Pruitt KD, Katz KS, Sicotte H and Maglott DR. (2000). *Trends Genet.,* **16,** 44–47.
Rawlings DJ and Witte ON. (1995). *Semin. Immunol.,* **7,** 237–246.
Rikke BA, Murakami S and Johnson TE. (2000). *Mol. Biol. Evol.,* **17,** 671–683.
Robertson SC, Tynan J and Donoghue DJ. (2000). *Trends Genet.,* **16,** 368.
Saitou N and Nei M. (1987). *Mol. Biol. Evol.,* **4,** 406–425.
Sakuma M, Onodera H, Suyemitsu T and Yamasu K. (1997). *Zoolog. Sci.,* **14,** 941–946.
Schaeffer EM and Schwartzberg PL. (2000). *Curr. Opin. Immunol.,* **12,** 282–288.
Schenk PW and Snaar-Jagalska BE. (1999). *Biochim. Biophys. Acta,* **1449,** 1–24.
Schlaepfer DD and Hunter T. (1998). *Trends Cell Biol.,* **8,** 151–157.
Schlessinger J. (1997). *Cell,* **91,** 869–872.
Schlessinger J. (2000). *Cell,* **100,** 293–296.
Schlessinger J and Ullrich A. (1992). *Neuron,* **9,** 383–391.
Schwartzberg PL. (1998). *Oncogene,* **17,** 1463–1468.
Serfas MS and Tyner AL. (1998). *Oncogene,* **17,** 3435–3444.
Shurin MR, Esche C and Lotze MT. (1998). *Cytokine Growth Factor Rev.,* **9,** 37–48.
Smithgall TE, Rogers JA, Peters KL, Li J, Briggs SD, Lionberger JM, Cheng H, Shibata A, Scholtz B, Schreiner S and Dunham N. (1998). *Crit. Rev. Oncog.,* **9,** 43–62.
Sondhi D and Cole PA. (1999). *Biochemistry,* **38,** 11147–11155.
Suga H, Koyanagi M, Hoshiyama D, Ono K, Iwabe N, Kuma K and Miyata T. (1999). *J. Mol. Evol.,* **48,** 646–653.
Takeuchi S, Takeda K, Oishi I, Nomi M, Ikeya M, Itoh K, Tamura S, Ueda T, Hatta T, Otani H, Terashima T, Takada S, Yamamura H, Akira S and Minami Y. (2000). *Genes Cells,* **5,** 71–78.
Tallquist MD, Soriano P and Klinghoffer RA. (1999). *Oncogene,* **18,** 7917–7932.
Tamura K and Nei M. (1993). *Mol. Biol. Evol.,* **10,** 512–526.
Thomas SM and Brugge JS. (1997). *Annu. Rev. Cell. Dev. Biol.,* **13,** 513–609.
Tsai JY and Silver LM. (1996). *Science,* **271,** 1432–1434; (discussion) 1434–1435.
Turner M, Schweighoffer E, Colucci F, Di Santo JP and Tybulewicz VL. (2000). *Immunol. Today,* **21,** 148–154.
van der Voort R, Taher TE, Derksen PW, Spaargaren M, van der Neut R and Pals ST. (2000). *Adv. Cancer Res.,* **79,** 39–90.
Vasioukhin V and Tyner AL. (1997). *Proc. Natl. Acad. Sci. USA,* **94,** 14477–14482.
Veikkola T, Karkkainen M, Claesson-Welsh L and Alitalo K. (2000). *Cancer Res.,* **60,** 203–212.
Vogel W. (1999). *FASEB J.,* **13,** S77–S82.
Walchli S, Colinge J and Hooft van Huijsduijnen R. (2000). *Gene,* **253,** 137–143.
Weiner HL and Zagzag D. (2000). *Cancer Invest.,* **18,** 544–554.
Weisberg E, Sattler M, Ewaniuk DS and Salgia R. (1997). *Crit. Rev. Oncog.,* **8,** 343–358.
Wells A. (1999). *Int. J. Biochem. Cell Biol.,* **31,** 637–643.
White JA, McAlpine PJ, Antonarakis S, Cann H, Eppig JT, Frazer K, Frezal J, Lancet D, Nahmias J, Pearson P, Peters J, Scott A, Scott H, Spurr N, Talbot Jr C and Povey S. (1997). *Genomics,* **45,** 468–471.
Whitehead JP, Clark SF, Urso B and James DE. (2000). *Curr. Opin. Cell Biol.,* **12,** 222–228.
Xu Q and Wilkinson DG. (1997). *J. Mol. Med.,* **75,** 576–586.
Yang WC, Collette Y, Nunes JA and Olive D. (2000). *Immunity,* **12,** 373–382.
Yeung CH, Sonnenberg-Riethmacher E and Cooper TG. (1998). *J. Reprod. Fertil.* (Suppl.), **53,** 137–147.