## Embryonic stem cells express multiple Eph-subfamily receptor tyrosine kinases

JASON D. LICKLITER\*, FIONA M. SMITH, JANE E. OLSSON, KAREN L. MACKWELL, AND ANDREW W. BOYD

Lions Cancer Research Laboratory, Walter and Eliza Hall Institute of Medical Research, P.O. Royal Melbourne Hospital, Victoria 3050, Australia

Communicated by G. J. V. Nossal, The Walter and Eliza Hall Institute of Medical Research, Victoria 3050, Australia, September 29, 1995

ABSTRACT Eph and its homologues form the largest subfamily of receptor tyrosine kinases. Normal expression patterns of this subfamily indicate roles in differentiation and development, whereas their overexpression has been linked to oncogenesis. This study investigated the potential role of Eph-related molecules during very early embryonic development by examining their expression in embryonic stem (ES) cells and embryoid bodies differentiated from ES cells in vitro. By use of a strategy based on reverse transcriptase-mediated PCR, nine clones containing Eph-subfamily sequence were isolated from ES cells. Of these, eight were almost identical to one of four previously identified molecules (Sek, Nuk, Eck, and Mek4). However, one clone contained sequence from a novel Eph-subfamily member, which was termed embryonic stemcell kinase or Esk. Northern analysis showed expression of Esk in ES cells, embryoid bodies, day 12 mouse embryos, and some tissues of the adult animal. Levels of expression were similar in ES cells and embryoid bodies. By comparison, Mek4 showed no significant transcription in the ES cell cultures by Northern analysis, whereas Eck displayed stronger signals in ES cells than in the embryoid bodies. These results suggest that Eph-subfamily molecules may play roles during the earliest phases of embryogenesis. Furthermore, the relative importance of different members of this subfamily appears to change as development proceeds.

The receptor tyrosine kinases (RTKs) are transmembrane molecules which transduce signals from the extracellular environment into the cytoplasm. They include well-studied regulators of cell proliferation and differentiation, such as c-Kit and the receptors for epidermal growth factor, platelet-derived growth factor, and macrophage colony-stimulating factor (1). Signaling by an RTK is initiated when its ligand binds to the extracellular domain of the receptor. This leads to the formation of receptor dimers, which activate their catalytic domains by reciprocal phosphorylation on tyrosine residues (2). Once activated, RTKs can bind and phosphorylate specific intracellular proteins that act as second messengers.

Eph was the first isolated member of what is currently the largest subfamily of RTKs (3–20). This group is distinguished by a cysteine-rich region and two fibronectin type III repeats in the extracellular domain (4). Eph-subfamily kinases have been found in diverse species, including zebrafish (17), frogs (19), chickens (4, 8, 12), mice (8, 11, 14–16, 20), rats (5, 7), and humans (3, 6, 10, 13, 18). Features of their expression pattern suggest key functions during embryonic development. First, strong expression in the embryo is characteristic (4, 8, 11, 12, 15, 21). Second, *in situ* hybridization and immunolocalization studies show associations between the expression of specific Eph-subfamily molecules and particular events in morphogenesis. For example, Eck is transiently expressed in cells adjacent to the primitive streak during gastrulation, and later its transcripts are found in specific rhombomeres of the developing

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

hindbrain and in the ectoderm of the second and third branchial arches (22). The expression of Nuk protein on growing peripheral nervous system axons, which disappears when the axons have ceased migrating, and the segmentrestricted pattern of Nuk and Sek expression during hindbrain morphogenesis are other examples (11, 15). Preferential expression at interfaces between embryonic cell populations and in intercellular junctions has led to the suggestion that Ephsubfamily molecules influence embryonic differentiation and cellular migration by interactions involving direct cell-cell contact (15, 22). The recent finding that ligands for some members of this group are cell membrane bound supports this notion (23-26). Eph-related molecules also have a potential role in oncogenesis. Eph and Eck overexpression was reported in epithelial tumor cell lines and some human carcinomas (3, 27), and Hek overexpression occurs sporadically in leukemia (9). Furthermore, artificial overexpression of Eph transforms NIH 3T3 cells, allowing them to form colonies in agar and produce tumors in nude mice (28). In transgenic models of murine mammary cancer, overexpression of the Eph-subfamily members Myk-1 and Myk-2 correlates with the development of poorly differentiated and invasive tumors, suggesting a role in tumor progression (20).

In this study, reverse transcriptase-mediated PCR (RT-PCR) was used to identify Eph-subfamily RTKs expressed by embryonic stem (ES) cells. These are undifferentiated, totipotent cells derived from the inner cell mass of the blastocyst (29). We speculated that Eph-related molecules expressed by ES cells may be involved in the initial differentiation and organization of embryonic tissues, thus helping to establish the blueprint for later development. Nine positive clones were isolated by this method, including one which encoded sequence from a novel molecule.<sup>†</sup> We also cultured ES cells under conditions which induced *in vitro* differentiation into embryoid bodies containing a variety of primitive tissue lineages (30). Expression of selected Eph-related molecules was then examined in both the ES cells and embryoid bodies by Northern analysis.

## **MATERIALS AND METHODS**

ES Cell Cultures. The murine 129/Sv-derived ES cell line W9.5 (provided by F. Koentgen of the Walter and Eliza Hall Institute, Melbourne, Australia) was routinely passaged on underlayers of irradiated embryonic fibroblasts in Dulbecco's modified Eagle's medium supplemented with leukemia inhibitory factor (LIF, 1000 units/ml; AMRAD, Melbourne), 0.1 mM 2-mercaptoethanol, and 15% fetal bovine serum. Cultures were incubated in a 10% CO<sub>2</sub> atmosphere at 37°C. In preparation for the studies described below, ES cells were subcul-

Abbreviations: RTK, receptor tyrosine kinase; ES cells, embryonic stem cells; RT-PCR, reverse transcriptase-mediated PCR; LIF, leukemia inhibitory factor; GAPDH, glyceraldehyde-3-phosphate dehydrogenase.

<sup>\*</sup>To whom reprint requests should be addressed.

<sup>&</sup>lt;sup>†</sup>The sequence reported in this paper has been deposited in the GenBank database (accession no. MMU18084).

tured into gelatin-coated flasks, and four passages without a feeder layer were performed to deplete the embryonic fibroblasts. In some of the cultures, LIF was withdrawn 11 days prior to harvesting, to allow differentiation into embryoid bodies to take place (30). Control cultures of embryonic fibroblasts alone were also prepared.

RT-PCR. Prior to RNA extraction, cultures of undifferentiated ES cells were disrupted with trypsin and the cells were washed in phosphate-buffered saline. Cell pellets were resuspended in buffer containing guanidinium isothiocyanate and total RNA was extracted with the use of organic solvents (31). cDNA was then synthesized from 1  $\mu$ g of total RNA by use of an oligo(dT) primer and avian myeloblastosis virus reverse transcriptase (Promega). PCR amplifications were performed with degenerate primers derived from three regions of sequence which are relatively conserved in the Eph subfamily (13) (Fig. 1). Reaction mixtures (30  $\mu$ l) contained 50 mM KCl, 10 mM Tris HCl (pH 8.3), 1.25 mM MgCl<sub>2</sub>, 0.2 mM each dNTP, 2.5 units of Taq polymerase (Perkin-Elmer), 30 pmol of sense primer (P1 or P2), 30 pmol of antisense primer (P4), and 3  $\mu$ l of the ES cell cDNA synthesis reaction. In addition, mRNA was directly extracted from ES cells with oligo(dT)coated magnetic beads (Dynal, Oslo) and cDNA was synthesized from 1  $\mu$ g of mRNA with an oligo(dT) primer and Superscript II reverse transcriptase (Life Technologies, Grand Island, NY). Half of the reverse transcription reaction was then amplified with primers P1 and P4, with other reaction conditions identical to those described above. PCR products were electrophoretically separated, purified with the Geneclean II kit (Bio 101), and then subjected to a second round of PCR. Products initially amplified with primers P1 and P4 were reamplified with the same primers, while those amplified with P2 and P4 were reamplified with P2 and P3. All reactions were carried out in a PTC-100 programmable thermal controller (MJ Research, Cambridge, MA) employing cycle programs specific to each primer combination, as follows: 1 min at 95°C, 2 min at 70°C, and 3 min at 72°C for 30 cycles (primers P1 and P4); 1 min at 95°C, 1 min at 51°C, and 1 min at 72°C for 35 cycles (primers P2 and P4); 1 min at 95°C, 1 min at 41°C, and 1 min at 72°C for 35 cycles (primers P2 and P3).

**Cloning and Sequencing.** Reamplified, gel-purified PCR fragments were cloned into the *Sma* I site of pUC18 with the SureClone ligation kit (Pharmacia). Recombinant clones were sequenced with the *Taq* DyeDeoxy Terminator Cycle sequencing kit (Applied Biosystems); a Perkin–Elmer GeneAmp PCR System 2400 was used to perform the sequencing reactions and an Applied Biosystems 373 DNA sequencer was used for their subsequent analysis. Sequences were compared against se-



FIG. 1. Position of the four degenerate oligonucleotide primers used for RT-PCR. Relatively conserved peptide motifs on which the primers were based are shown above a schematic representation of the basic domain structure of Eph-subfamily molecules. ECD, extracellular domain; ICD, intracellular domain. The particular amino acid sequences shown are from Eph. Two sense primers (P1 and P2) and two antisense primers (P3 and P4) were synthesized. Their sequences were as follows: P1, 5'-GTAGGCATGCAAGGAGAC(A/C)TT(C/ T)AACC-3'; P2, 5'-GCGATGATCAT(C/G)AC(A/G/T)GA(A/ G)TA(C/T)ATGG-3'; P3, 5'-GTAGGAATTCCA(C/G/T)A-CATC(A/G)CT(A/G)GC-3'; P4, 5'-CCA(T/A)A(A/G)CTCCA(C/ T)ACATC(A/G)CT-3'.

quence databank entries with the FASTA sequence analysis program.

Northern Blots. Poly(A)<sup>+</sup> RNA from ES cells, embryoid bodies, embryonic fibroblasts, day 12 mouse embryos, and adult mouse tissues (5  $\mu$ g per sample) was electrophoresed through 1.2% agarose gels containing 2.2 M formaldehyde and transferred to nylon membranes (Zeta-Probe, Bio-Rad). The Northern blots were then probed with cDNA inserts from selected ES cell recombinants cloned by the methods described earlier. Sequence analysis (see Table 1) showed that clone 35C4 encoded murine Eck (21), clone 35C11 encoded Mek4 (8), and clone 35C15 encoded an apparently novel sequence. Inserts were digested from these clones with EcoRI and Xba I and used as templates for synthesizing <sup>32</sup>P-labeled probes with the Prime-It II random primer labeling kit (Stratagene). A labeled probe was also made from glyceraldehyde-3phosphate dehydrogenase (GAPDH) cDNA. All blots were initially hybridized to the 35C15 probe. Subsequently, the membrane containing RNA from ES cells, embryoid bodies, and embryonic fibroblasts was reprobed with the Eck and Mek4 probes. Between each reprobing, hybridized probe was stripped from the membrane by pouring boiling 0.1% SDS on the membrane and cooling to room temperature. Exposure to x-ray film overnight confirmed effective removal of the probe. Hybridization to the GAPDH probe was performed last. In all cases, hybridizations took place in 50% formamide at 42°C and washes were performed under stringent conditions, with the final wash at 65°C in 0.1× SSC/0.1% SDS (SSC is 0.15 M NaCl/0.015 M sodium citrate, pH 7.6). Autoradiographs were exposed at -70°C.

## RESULTS

**Eph-Subfamily Molecules Expressed by ES Cells.** Primers P2 and P4 were expected to amplify  $\approx$ 350 bp from the catalytic domain of Eph-subfamily molecules. In contrast, it was anticipated that primers P1 and P4 would amplify about 2.1 kb, including much of the extracellular and intracellular domains (Fig. 1). RT-PCR was initially performed on total RNA derived from ES cells. Reactions using primers P2 and P4 amplified a band of the expected size, which reamplified with primers P2 and P3. However, no PCR product of the anticipated size was observed when primers P1 and P4 were used. Subsequently, when magnetically separated poly(A)<sup>+</sup> RNA was substituted for total RNA, reactions with primers P1 and P4 successfully amplified a 2.1-kb product. Cloning of the 350-bp product resulted in four recombinants containing Eph-subfamily sequences (Table 1). These were highly homol-

Table 1. Eph-subfamily cDNA clones isolated from ES cells by RT-PCR

	Closest	Homology,					
Clone	homologue	* %					
	Primers P2 and	P4*					
25C5	Sek	97.8					
33C1.1	Nuk	97.2					
33C1.2	Eck	99.1					
33C1.5	Eck	100					
-	Primers P1 and	P4*					
35C4	Eck	97.2					
35C6	Mek4	97.3					
35C10	Mek4	99.7					
35C11	Mek4	99.7					
35C15	Eph	83					

Homologues were identified by screening nucleic acid databanks with the FASTA program. Sek (11), Nuk (15), Eck (21), and Mek4 (8) are murine Eph-subfamily molecules; Eph (3) is a human molecule. \*Refers to the primers used in the initial amplification during RT– PCR. ogous to either Sek (11), Nuk (15), or Eck (21)—members of the Eph subfamily previously isolated from murine sources. When the 2.1-kb product was cloned, five recombinants containing sequences of this subfamily were identified. Four of these showed high levels of homology with either murine Eck or Mek4 (8), but one clone (35C15) contained 1602 bp of sequence which appeared novel after comparative databank analysis (GenBank and GenBank Cummulative Updates, November 1995). The novel molecule was termed <u>embryonic stem</u> cell kinase or Esk.

Sequence and Structural Features of Esk. The sequence of the partial Esk cDNA is shown in Fig. 2. Its shortfall of about 500 bp from the expected length of 2.1 kb is due to a 5'-end deletion, which presumably occurred during cloning. The deduced amino acid sequence of Esk was obtained by using an open reading frame spanning the entire cDNA fragment. This encoded a sequence of 533 amino acids (Fig. 2). Sequence homology strongly suggests that Esk is a new Eph-subfamily member. Its closest relative is the human kinase Eph, which displays 83.0% identity in its nucleotide sequence and 78.8% identity at the amino acid level. Eck is the closest murine relative, showing 44.9% amino acid identity. An alignment of the peptide sequences of Esk, Eph, and Eck is shown in Fig. 3. Included in the Esk sequence is a group of hydrophobic residues (amino acids 269-289) which form a putative transmembrane domain, dividing the molecule into extracellular and intracellular portions. Two fibronectin type III repeats are arranged in series in the C-terminal part of the extracellular domain. Similar repeats are found in all other Eph subfamily molecules (4, 10, 18). Also highly conserved in this group is a cysteine-rich box (6, 18) in the N-terminal region. All of the expected cysteine residues were encoded by the 35C15 sequence, except for the anticipated cysteine at position 104, which was replaced by an arginine. Since 35C15 was cloned after a total of 60 cycles of amplification using Taq polymerase,

a small number of base misincorporations would be expected. To check whether this was responsible, cDNA was reverse transcribed from ES cell  $poly(A)^+$  RNA and subjected to 35 cycles of PCR using *Taq* polymerase and two Esk-specific primers 519 bp apart. A product of the expected size was amplified and directly sequenced with the PCR primers. Unambiguous sequence data were obtained over a region corresponding to bases 45–397 of clone 35C15, and this contained four base differences compared with the cloned PCR fragment. The differences included a thymine instead of a cytosine at base 311, which would result in a cysteine at residue 104. All analyses and figures in this report employ the bases observed by direct sequencing of PCR product where they are different from those in clone 35C15.

Several features of the deduced peptide sequence of Esk imply protein-tyrosine kinase activity. The putative intracellular domain contains a kinase consensus region, beginning at residue 343 (32). This includes the presumed ATP-binding-site motif Gly-Xaa-Gly-Xaa-Xaa-Gly at residues 352-357 and a lysine at position 377, thought to be involved in the phosphotransfer reaction and highly conserved among protein-tyrosine kinases (32). However, the partial cDNA sequence ends before a complete catalytic domain is encoded. Esk also displays the motifs Asp-Leu-Ala-Ala-Arg-Asn (positions 470-475) and Pro-Ile-Arg-Trp-Thr-Ala-Pro (positions 510-516), which indicate substrate specificity for tyrosine (32). However, the latter sequence is followed by a conserved glutamate in other protein-tyrosine kinases, while in Esk this residue is a glycine. This could indicate another PCR artifact. Finally, the tyrosine residue at position 502 is a potential autophosphorylation site.

**Esk Expression.** Northern analysis revealed expression of a 4.2-kb Esk transcript in day 12 mouse embryo and adult mouse thymus, liver, kidney, lung, and placenta (Fig. 4). Faint bands at  $\approx 6.0$  kb observed in the liver, kidney, and lung samples may be due to alternatively spliced transcripts, although weak

		10								20										30													
	CI	, т	G	F	Y	R	v	D	м	N	т	LP	с	L	ĸ	с	Ρ	0	н	s	I	Α	Е	s	Е	G	s	т	I	С	т	с	Е
1	CTGT	TAC	TIGG	<b>1.1.1</b>	CTA	TCG	AGT	GGA	CATG	AAT	ACA	CTCC	GTTC	TOT	(TAA	GTG	cca	CAZ	ACA'	TAGO	AT	AGC	AGAG	TC	IGA	300	TC	PAC	CAT	CIG	TACO	TG	TGA
-				40	• • • •								50										50										70
	NI C		v			в	~	F	c	ъ	~ ,		~	m	ъ	ъ	в	0	2	ъ	^	N	Ť	c	P	c	T	c	C	Ŧ	0	τ.	c
100	14 (		- <u>1</u>	~~~~	. ^ _					F	ě.					- 			<u>~</u> ~		¥	14	~~~~			~~~					<u>~</u>		~
106	GAATO	GAC	VI.I.A	TUG	AGCI	CCC	IGG	GGA	3661	CCC	CAG	GTAG	CAT	CAC	ACG	100	CCC	AICC	360	rccu	CA	AAA:	ICIG	AG	crr	CICC	ACA	and a	166	GAC	ICA	ACTO	CIC
								1	80									90									1	100					
	LF	2 W	Е	P	P	R	D	т	G	G	R I	нс	I	R	Y	s	v	Е	с	L	Q	с	R	G	Ι	А	Q	D	G	G	P	с	Q
211	CCTGC	GCT	<b>GGA</b>	GCC	CCC	CAG	AGA	TAC/	AGGG	GGA	CGC	CATG	ATAT	CAG	ATA	CAG	CGT	GAC	TG	CTT	GCA	GTG.	rcgg	GG	CAT	IGC/	ACAC	GA	IGG	GGG	TCCC	CTG	CCA
				110									120	)									130										140
	PC		к	G	v	н	F	s	P	A	A :	5 6	T.	т	т	S	т	v	0	v	0	G	τ.	E	Р	Y	А	N	v	т	F	т	v
316	ACCC		2888	<u>.</u>	тст т	202	-		$\frac{1}{2}$	<u>сст</u>	200		<u>~</u>	ra.	cac	2 Tr	- 	- TOTA	2		202		-	-	-	- 	200	אמי	- 	20	- 		COT
510	150 160 170											inc.																					
						~	~	. '	100	~	~ .			~		~			-			~				~	. 1	20	~		~		v
401	A 1		IN .		. ~	~~~~				5					~~~~	~~~		3		N		-G	п 2016	A			Сторок С Сторок Сторок С С С С С С С С С С С С С С С С С С С		-G	<u>_</u>			~~~
421	CAAAC			CAG	AGIV	GIC	AGG	ACIU	GAC	AGI	icc	AGCC	CIAC	CAG	CGC	CIC	cen	CAG'	LA1	CAAC	AIG	GGG	CAC	GC	AGA	31C/	ACIC	CIC	IGG	CCI	GIC	ACIO	GAA
				180									190										200										210
	гι	K	к	Е	Р	R	Q	L	Е	L	тι	N A	G	s	R	P	R	N	Ρ	G	G	N	L	s	Y	Е	L	н	v	г	N	Q	D
526	GCTGC	TGA	GAA	AGA	ACC	GAG	GCA	GCT	GAG	CTG	ACT	IGGG	CAGO	GTC	CCG	ACC	CCG	AA7	rcc.	TGG/	AGG	GAA!	ICIG	AG	CTA	IGAC	CTC	GCA	CGT	GCT	GAA	ICA	GGA
									220									230									2	240					
	EE	: W	н	0	м	v	L	Е	Ρ	R	V J	LR	т	к	L	0	P	D	т	т	Y	I	v	R	v	R	т	L	А	P	L	G	P
631	CGAAC	AATO	GCA	CCA	GAT	GGT	GTIN	GGA	ACCC	AGG	GTC	FTGC	GGAC	AAA	ACT	TCA	GCC	AGAT	CAC	CAC	ATA	CAT	IGTO	AG	AGT	GCGJ	ACA	ACTO	GC	ccc	ACTO	GGG	GCC
				250									260										270										280
	~ •		~		n	u	F	F	n	-	c 1			· ~		c		-	~	~	F	-	17		17	+		~				~	
726	- G 2		~~~~	~~~~				-	~~~~	1	200		~~~~		~~~		~~~~~	1	~		- -		~~~~	СС	~~~~~		r	-G				-G	
130	TGGCC	CTT	CIC	ccc	IGA	CCA.	IGAG	5111		ACA	AGC	LUAC	CAGI	TIC	CAG	AAG	CI	JUAC	.66	AGG	GAG	GAT.	GIG	GC	JGI	AIG			41-14	SCI	scr.	1GG4	AA1
									290									300									-	510					
	AI	, L	I	G	I	Y	v	F	R	s	RI	RG	Q	R	Q	R	Q	Q	R	Q	R	Е	R	т	т	N	v	G	R	Е	D	к	L
841	AGCTO	TGCT	GAT	CGG	GAT	TTA'	IGI	CTT	CCGT	TCA	AGG	AGAG	GCC2	GAG	ACA	GAG	ACA	SCAC	JAG	GCAC	GCG	TGA/	ACGC	ACO	CAC	CAAT	GIC	CG.	ICG.	AGA	GAC	CAA	CT
				320									330									3	340									:	350
	WI	K	Ρ	Y	v	٠D	L	Q	Α	Y	E I	DP	A	Q	G	А	L	D	F	Α	Q	Е	L	D	Р	Α	W	L	I	v	D	т	v
946	GTGGG	TAAZ	ACC	CTA	TGT	GGA	CCTN	CCAC	GCC	TAT	GAG	GACC	CTGC	ACA	GGG	AGC	CTT	AGAC	TT.	TGCO	CAC	GGA	ACTG	GA	ccc	AGCO	TGO	GCT	GAT	IGIN	GAG	CAC	IGT
								3	360									370									3	380					
	то	E	G	E	Y	G	Е	v	v	R	G	A L	R	τ.	P	s	0	D	c	к	т	v	А	т	к	т	т. Т	R	п	т	s	P	р
1051	CATAC	GAGZ	AGG	GGA					- 	nna	COM	2000	TGAC		rrn	പ്പ	າດີລະ	-	- Tra	CAAC	200	Tr:Tr	and the second	ידייני	האמ	200	-	280	-	TAC	amrr	-	
1031				200		100.							400							and			110								1100		420
	~ `			350						_	÷.,		-				-		+		-	. `		~		-	-		<b>n</b>		~	÷.	120
1150			~~~~	N	г				~~~	1	1 1		~~~~		N	<u>п</u>	P		. <u>.</u>		ĸ			6	~	1	T		R		P		
1120	IGGCI	ACIO	GIG	GAA	T-P-P	CCP.	ICG	AGA	JGCA	ACT	AIC	AIGG	GCCF	GPI	CAA	CCA	ccc	ACAC	AT	IC17	ACG	CCI	<b>GAA</b>	GG.	IGI	CATC	ACA		AG	AAA	GCCC	CAT	CAT
								4	130									40									4	150					
	II	Т	Е	F	м	Е	N	G	A	L	D	A F	L	ĸ	Е	R	Е	G	Q	L	А	P	G	Q	L	v	А	м	L	L	G	Ι	А
1261	GATC	TCAC	AGA	GTT	TAT	GGA	AAA	TGG/	AGCC	CIG	GAT	CCI	TICI	GAA	GGA	ACG	GAG	GGG	CA	ACT7	ACC.	ICC.	IGGI	CAC	<b>GCT</b>	AGT	GCI	TAT	<b>SCT</b>	ACTO	GGG	CAT	AGC
				460									470									4	80										490
	S G	M	N	С	L	s	G	н	N	Y	vı	H R	D	L	А	А	R	N	I	L	v	N	0	N	L	С	С	к	v	s	D	L	G
1366	ATCAG	GCAT	GAA	CTG	CCT	CAG	IGG	CCAC	CAAT	TAT	GTC	CATA	GAGA	CCI	GGC	TGC	CAG	AAC	ATC	CTR	GT	GAAD	CAG	AA	CT	TG	TGC		GT	TC	IGAC	CT	IGG
									500									510										20					
	т. п		т	T	Ð	Б	F		0	m	vi		0	0	c	v	- <sup>`</sup>	'n	т	ъ	1.7	-		ъ	~		÷ -	20	ы	ъ	τ.	F	m
1471	00000	$\sim$		~~~~	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~				~~~~	200	- 		$\sim$	~~~	·~		~~~~~~	÷~~		~~~·		~. ~	~~~~	÷				~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	~~~~	~~~		-	-
74/1	CIIGA		~ C I	CC1	GGH.	1 GMC		- GAU		ALC	INIC	3MMA	CUA	فحاصف	nuu	nna	3410		-MI(	دريغا	191	JUME	5500	ι u	100	100	un 1		<b>א</b> ت.	ICG(	34.10	.140	
			-	230		-	-																										
	TA	S	D	v	W	S	F																										
	~ ~ ~ ~ ~	0030		mom	~~~~	~~~~		$\mathbf{n}$																									
15/6	CACAG	CCAG	TGA	IGI	316	JUAC	.111	199																									

FIG. 2. Partial nucleotide and deduced amino acid sequences of Esk. Nucleotides are numbered down the left side of the figure; amino acids are numbered above the peptide sequence. The putative transmembrane domain is overlined. Clone 35C15 showed the following differences when compared with directly sequenced PCR product over nucleotides 45–397: nt 86 = C, aa 29 = Pro; nt 232 = G, aa 77 = Arg (unchanged); nt 254 = G, aa 85 = Val; nt 311 = C, aa 104 = Arg.



FIG. 3. Peptide sequence of Esk aligned with two close homologues, Eph and Eck. Conserved cysteine residues are marked with arrowheads. The fibronectin type III repeats are indicated with a single overline and the transmembrane domain with a double overline. The bracket shows the beginning of the catalytic domain, and the ATP-binding motif is marked with dots.

hybridization to related molecules cannot be excluded. No signal was detected from lymph node, spleen, heart, brain, or skeletal muscle.

**Comparative Expression of Esk, Mek4, and Eck in ES Cells and Embryoid Bodies.** To determine whether expression of Eph-subfamily molecules occurred in ES cells at a significant level and to investigate potential changes in expression as ES cells differentiated *in vitro*, a Northern blot containing ES cell and embryoid body RNA was sequentially hybridized to Esk, Mek4, and Eck probes. Embryonic fibroblast RNA was included as a control on this blot, because some fibroblast



FIG. 4. Expression of Esk. Northern blots were prepared with  $poly(A)^+$  RNA (5 µg per lane) extracted from day 12 mouse embryo and the adult mouse tissues shown. Hybridization was performed in turn to a <sup>32</sup>P-labeled probe derived from clone 35C15 and a GAPDH probe. The positions of RNA size markers are indicated to the left of the blots.

contamination of the ES cell and embryoid body samples could not be excluded. As shown in Fig. 5, the Esk probe hybridized to all three samples, approximately in proportion to the amount of RNA present. This indicates significant levels of Esk expression in ES cells and embryoid bodies and could not be accounted for by fibroblast contamination alone. In contrast, expression of Mek4 was barely detectable in the ES cell



FIG. 5. Comparative expression of Esk, Mek4, and Eck during *in vitro* differentiation of ES cells. Cultures of ES cells were induced to differentiate into embryoid bodies by the withdrawal of LIF 11 days prior to harvesting. The Northern blot was prepared with poly(A)<sup>+</sup> RNA (5  $\mu$ g per lane) extracted from undifferentiated ES cells (ESC), embryoid bodies (EB), and embryonic fibroblasts (EF). <sup>32</sup>P-labeled probes synthesized from Esk, Mek4, Eck, and GAPDH were then sequentially hybridized to the membrane. The sizes of transcripts are indicated at right.

cultures. A 9.5-kb Mek4 transcript was expressed by the embryonic fibroblasts, however, suggesting that the faint bands seen in the ES cell lane may be due to contaminating fibroblast RNA. Finally, the Eck probe hybridized to all three samples, but expression was relatively greater in the undifferentiated ES cells. Transcripts were of slightly different sizes in the different lanes.

## DISCUSSION

Nine Eph-subfamily cDNA clones were isolated from ES cells by use of RT-PCR with degenerate primers specific for this subfamily (Table 1). Of these, eight showed >97% homology to one of four previously reported molecules (Sek, Nuk, Eck, and Mek4). Given that clones were isolated after 60-70 amplification cycles using Taq polymerase, PCR artifacts are a likely reason for the less than complete identity with published sequences. In support of this, when sequence from Esk obtained by direct sequencing of PCR product was compared with clone 35C15, a misincorporation rate of 4/353 bases (1.1%) was observed in the latter. However, it remains possible that clones amplified from the catalytic domain with primers P2 and P4 could be derived from novel RTKs. This region is highly conserved within the Eph subfamily and may be close to identical in molecules which have markedly divergent sequence elsewhere (13).

A novel molecule, Esk, was also identified in ES cells. The deduced peptide sequence of Esk contains the conserved cysteine residues and fibronectin type III repeats characteristic of the Eph subfamily. It also shows typical protein-tyrosine kinase catalytic domain motifs. These findings define Esk as a member of the Eph subfamily of RTKs. Esk most closely resembles the human kinase Eph, with an overall amino acid sequence identity of 78.8%. When the most disparate region of Esk (residues 302-338) is excluded, the identity is 85.7%. However, this remains less than the similarity between other Eph-subfamily molecules considered to be interspecies homologues. For example, Mek4 shows 95.5% amino acid sequence identity to Hek, and murine Eck is 92% identical to its human homologue (21). Although a definitive answer awaits further data, it appears that Esk may not be the murine homologue of Eph.

Northern analysis of Esk revealed strong expression in ES cells, embryoid bodies, and day 12 embryo, suggesting that Esk, like other Eph-related kinases, has a role in development. Expression was also observed in adult mouse thymus, liver, kidney, and lung, which implies ongoing roles in these tissues. In contrast to the situation for the majority of its subfamily, Esk transcripts were not observed in adult brain. Eph and Eck show a similar propensity for expression in epithelial tissues and, together with Esk, may form a subgroup within the Eph subfamily.

When expression of two other Eph-related molecules in the ES cell cultures was investigated, contrasting patterns were observed. Mek4 demonstrated no significant expression by Northern analysis, while Eck was preferentially expressed by undifferentiated ES cells. These results suggest that Esk and Eck are likely to be more important than Mek4 in the *in vitro* differentiation of ES cells to embryoid bodies. A corollary is that they may also be more important in very early embryogenesis *in vivo*. The identification of Eck protein in gastrula-stage embryos (22) provides additional support for this conclusion.

While the specific functions of Eph-related RTKs in development require further study, several findings indicate involvement in signaling processes mediated by direct cell-cell contact (15, 23–26). Therefore, although stimulation of proteintyrosine kinase activity is the final common pathway of signal transduction, this may be critically linked to the adhesive interactions of these molecules. One possibility is that the extensive family of Eph-related proteins convey positional information between embryonic cells, which regulates their migration and differentiation. Given the oncogenic potential of some Eph-like RTKs (3, 9, 20, 27, 28), the attainment of correct position within the embryo might be communicated by signals which induce proliferation and/or prevent apoptosis. In this context, it is interesting that Esk, Eck, and Mek4 were all present in embryonic fibroblasts. Expressing multiple Ephrelated molecules may permit fibroblasts to migrate into all developing tissues.

The normal functions of Eph-related proteins in adult tissues remain unclear but may be an extension of their functions in development. For example, thymocyte maturation requires direct cell-cell interactions and a specific migration pathway within the thymus (33). Eph-subfamily molecules expressed in thymus, such as Esk, are possible mediators of some of these interactions. In this context, it is interesting that Esk transcripts were not observed in lymph node and spleen, suggesting expression by thymus-specific populations such as CD4<sup>+</sup>CD8<sup>+</sup> T-cell progenitors or thymic epithelial cells. Studies of the ligands of Eph-related RTKs, and potentially the outcome of gene targeting experiments, will be required to better define the biological function of these molecules.

We thank David Vaux, Christine Hawkins, and Steve Gerondakis for helpful discussions. We acknowledge the assistance of Vladimir Brusic with the sequence analysis and John Wilkins and Philip Vernon with the graphic artwork. This study was supported by the National Health and Medical Research Council, the Anti-Cancer Council of Victoria, and the Cooperative Research Centre for Cellular Growth Factors. J.D.L. is a National Health and Medical Research Council Postgraduate Scholar.

- 1. Yarden, Y. & Ullrich, A. (1988) Annu. Rev. Biochem. 57, 443-478.
- 2. Schlessinger, J. & Ullrich, A. (1992) Neuron 9, 383-391.
- Hirai, H., Maru, Y., Hagiwara, K., Nishida, J. & Takaku, F. (1987) Science 238, 1717–1720.
- 4. Pasquale, E. B. (1991) Cell Regul. 2, 523-534.
- Letwin, K., Yee, S. P. & Pawson, T. (1988) Oncogene 3, 621-627.
  Lindberg, R. A. & Hunter, T. (1990) Mol. Cell. Biol. 10, 6316-
- 6324.
- 7. Chan, J. & Watt, V. M. (1991) Oncogene 6, 1057-1061.
- Sajjadi, F. G., Pasquale, E. B. & Subramani, S. (1991) New Biol. 3, 769–778.
- Boyd, A. W., Ward, L. D., Wicks, I. P., Simpson, R. J., Salvaris, E., Wilks, A., Welch, K., Loudovaris, M., Rockman, S. & Busmanis, I. (1992) J. Biol. Chem. 267, 3262–3267.
- Wicks, I. P., Wilkinson, D., Salvaris, E. & Boyd, A. W. (1992) Proc. Natl. Acad. Sci. USA 89, 1611–1615.
- Gilardi, H. P., Nieto, M. A., Frain, M., Mattei, M. G., Chestier, A., Wilkinson, D. G. & Charnay, P. (1992) Oncogene 7, 2499–2506.
- 12. Sajjadi, F. G. & Pasquale, E. B. (1993) Oncogene 8, 1807-1813.
- Bohme, B., Holtrich, U., Wolf, G., Luzius, H., Grzeschik, K. H., Strebhardt, K. & Rubsamen, W. H. (1993) Oncogene 8, 2857– 2862.
- 14. Maisonpierre, P. C., Barrezueta, N. X. & Yancopoulos, G. D. (1993) Oncogene 8, 3277-3288.
- Henkemeyer, M., Marengere, L. E., McGlade, J., Olivier, J. P., Conlon, R. A., Holmyard, D. P., Letwin, K. & Pawson, T. (1994) *Oncogene* 9, 1001–1014.
- Zhou, R., Copeland, T. D., Kromer, L. F. & Schulz, N. T. (1994) J. Neurosci. Res. 37, 129–143.
- 17. Xu, Q., Holder, N., Patient, R. & Wilson, S. W. (1994) Development (Cambridge, U.K.) 120, 287-299.
- Bennett, B. D., Wang, Z., Kuang, W. J., Wang, A., Groopman, J. E., Goeddel, D. V. & Scadden, D. T. (1994) *J. Biol. Chem.* 269, 14211–14218.
- 19. Winning, R. S. & Sargent, T. D. (1994) Mech. Dev. 46, 219-229.
- Andres, A. C., Reid, H. H., Zurcher, G., Blaschke, R. J., Albrecht, D. & Ziemiecki, A. (1994) Oncogene 9, 1461–1467.
- Ganju, P., Shigemoto, K., Brennan, J., Entwistle, A. & Reith, A. D. (1994) Oncogene 9, 1613–1624.
- 22. Ruiz, J. C. & Robertson, E. J. (1994) Mech. Dev. 46, 87-100.

- Bartley, T. D., Hunt, R. W., Welcher, A. A., Boyle, W. J., Parker, V. P., Lindberg, R. A., Lu, H. S., Colombero, A. M., Elliott, R. L., Guthrie, B. A., Holst, P. L., Skrine, J. D., Toso, R. J., Zhang, M., Fernandez, E., Trail, G., Varnum, B., Yarden, Y., Hunter, T. & Fox, G. M. (1994) Nature (London) 368, 558-560.
- Beckmann, M. P., Cerretti, D. P., Baum, P., Vanden, B. T., James, L., Farrah, T., Kozlosky, C., Hollingsworth, T., Shilling, H., Maraskovsky, E., Fletcher, F. A., Lhotak, V., Pawson, T. & Lyman, S. D. (1994) *EMBO J.* 13, 3757–3762.
- 25. Cheng, H. J. & Flanagan, J. G. (1994) Cell 79, 157-168.
- Shao, H., Lou, L., Pandey, A., Pasquale, E. B. & Dixit, V. M. (1994) J. Biol. Chem. 269, 26606-26609.

.

- Kiyokawa, E., Takai, S., Tanaka, M., Iwase, T., Suzuki, M., Xiang, Y. Y., Naito, Y., Yamada, K., Sugimura, H. & Kino, I. (1994) *Cancer Res.* 54, 3645–3650.
- 28. Maru, Y., Hirai, H. & Takaku, F. (1990) Oncogene 5, 445-447.
- Abbondanzo, S. J., Gadi, I. & Stewart, C. L. (1993) Methods Enzymol. 225, 803–823.
- Keller, G., Kennedy, M., Papayannopoulou, T. & Wiles, M. V. (1993) Mol. Cell. Biol. 13, 473–486.
- 31. Chomczynski, P. & Sacchi, N. (1987) Anal. Biochem. 162, 156-159.
- 32. Hanks, S. K., Quinn, A. M. & Hunter, T. (1988) Science 241, 42-52.
- 33. Anderson, G. & Jenkinson, E. J. (1995) Semin. Immunol. 7, 177-183.