Advances in Gene, Molecular and Cell Therapy ©2007 Global Science Books

Comparative Genomics on the Basal *Mammalian* Urokinase Receptor-Promoter and Impact for Potential *in Vivo*/Clinical Relevance in Cancer

Jens Schumacher • Heike Allgayer*

Department of Experimental Surgery and Molecular Oncology, Medical Faculty Mannheim, Ruprecht Karls University Heidelberg and Joint Unit Molecular Oncology of Solid Tumors - DKFZ (German Cancer Research Center) Heidelberg, Klinikum Mannheim, 68135 Mannheim, Germany

Corresponding author: * **heike.allgayer@chir.ma.uni-heidelberg.de**

ABSTRACT

This review focuses on the molecular gene regulation of the human urokinase-receptor (u-PAR; gene: PLAUR; plasminogen activator receptor, urokinase-type). U-PAR has important functions in mediating tumor-associated proteolysis, invasion and metastasis. In particular, the present article prioritises the comparative basal promoter sequence alignment of u-PAR genes from diverse *mammalian* organisms and discusses the functional importance of the different promoter motifs and their associated transcription factors (TF) not only in the light of the prognostic relevance of the human u-PAR promoter, but also of the potential influence of these promoter sites on u-PAR gene regulation in all species analyzed.

Keywords: comparative genomics, PLAUR, promoter, sequence alignment, u-PAR

CONTENTS

INTRODUCTION

One of the critical hallmarks of cancer is invasion and metastasis (Hanahan and Weinberg 2000). The metastatic cascade involves several distinct processes. Initial crucial steps include the disruption of the physical cell-cell-adhesion and/or cell-matrix-attachment, and the activation of proteases which thereby lead to local invasion and intravasation. Especially, cell-surface-associated proteases degrading extracellular matrix proteins have been linked to metastasis.

One protease system which is linked to, and upregulated in several types of cancer is the u-PA-system (Blasi 1993). It should be emphasized that this protease is not restricted to cancer. Indeed, this system is involved in several tissue remodeling processes such as wound healing, fibrinolysis, inflammation, embryogenesis, and angiogenesis (Dano *et al.* 1985; Blasi 1988; Liotta *et al.* 1991; Blasi 1993). For example, u-PAR is expressed at the leading edge in re-epithelializing wounds (Romer *et al.* 1994). Characteristic components of the u-PA-system are the urokinase-type plasminogen activator (u-PA), the u-PA-specific inhibitors plasminogen-activator-inhibitor (PAI) -1, -2, and the cell surface receptor u-PAR (**Fig. 1**). U-PAR is a globular molecule, which consists of 3 disulfide-linked similar repeats (D1-D3) of approximately 90 residues each, the last (D3) of them being anchored to the cell membrane *via* a glycosyl-phosphatidylinositol chain (Behrendt *et al.* 1990;

Received: 11 March, 2007. Accepted: 25 April, 2007.

Special Feature

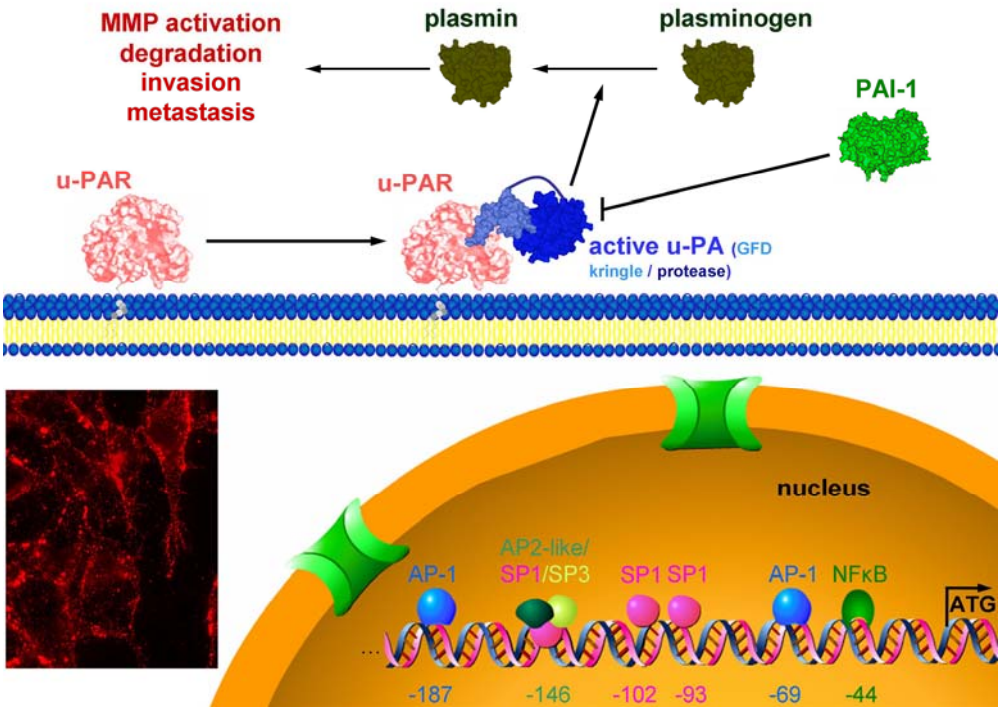

Fig. 1 Schematic overview of the proximal human u-PAR promoter, TF (transcription factors) and some functions and interactions of the u-PAR system. The molecular structure of **u-PAR** has been adapted from the crystal structure (Llinas *et al.* 2005). The u-PAR/u-PA (**GFD- and kringle-domain**) complex is shown as recently published from Huai *et al.* the arrangement of the **protease domain** (Sperl *et al.* 2000) is hypothetically shown via a linker. The structure of **PAI-1** is assumed from pdb: 1LJ5. Note that only the catalytic domain of **plasmin** is drawn (Wang *et al.* 2000). The immunofluorescence (with the monoclonal antibody 3936) image shows a human colon cancer cell line (RKO) overexpressing u-PAR.

Ploug *et al.* 1991; Llinas *et al.* 2005). The one-chain proenzyme u-PA binds the receptor specifically and with high affinity (Stoppelli *et al.* 1986). U-PAR binds the A-chain (1-135aa, including the growth factor-like domain (GFD) and the kringle domain) of u-PA and converts u-PA by cleavage into the active two-chain u-PA form. A recent structural paper (Huai *et al.* 2006) reports of the co-crystallisation at 1.9 Å, of the urokinase receptor complexed with the urokinase amino-terminal fragment (including the growth factor-like domain (GFD) and the kringle domain), and an antibody against the receptor. All three u-PAR domains (D1-D3) and both u-PA domains (GFD and kringle), are necessary for high-affinity u-PA/u-PAR interaction. Cleaved u-PA activates ubiquitously available plasminogen, and initiates the proteolytic cascade by the catalytic B-chain (proteases domain) (Stoppelli *et al.* 1986). Receptor-bound u-PA is inactivated by PAI-1 (-2), and the trimeric complex u-PAR/u-PA/PAI is internalized into the cell. Free u-PAR is recycled to the cell surface, and binding and activation of a second u-PA-molecule can occur (Allgayer 2006 and references therein). Furthermore, u-PAR is glycosylated at N-residues of glucosamine and sialic acid within the binding site, thereby regulating its affinity (K_D of 0.1-1.0 nM) for u-PA (Behrendt *et al.* 1990). Receptor-bound u-PA, when compared to the fluid phase enzyme, activates plasminogen much more efficiently, this being reflected by a 40-fold decrease in K_m of urokinase for its substrate (Ellis *et al.* 1991). The GPI-anchor of u-PAR is hypothesized to enable a high intramembrane mobility (Stoppelli *et al.* 1986; Behrendt *et al.* 1990).

For many different human tumors like breast, lung, kidney, liver, rectum and colon cancer, an overexpression of the u-PAR gene as compared to the normal tissue has been shown (Pyke *et al.* 1991; Jankun *et al.* 1993; Pyke *et al.* 1993; Wagner *et al.* 1995; Morita *et al.* 1997; Morita *et al.* 1998; overview in de Bock and Wang 2004). This increase is linked to an elevated invasive capacity of malignant tumor cells (Hollas *et al.* 1991; Bianchi *et al.* 1994), and to intravasation. Clinical up-regulation of the u-PA-system is an independent parameter predicting poor prognosis of patients with different cancers such as breast, gastric, or colorectal carcinoma (for an overview see Fuchs and Allgayer 2003). An up-regulation of u-PAR gene expression seems to be mainly due to an increase in gene transcription (Lund *et al.* 1995; Lengyel *et al.* 1996; Hapke *et al.* 2001a), although additional means of regulation such as mRNA stability, receptor recycling and post-translational modifications can occur (Lund *et al.* 1995; Wagner *et al.* 1995; Shetty *et al.* 1997). The human u-PAR gene (PLAUR) is located on chromosome 19q13 and spans seven exons transcribed into a 1,4 kbp mRNA (Borglum *et al.* 1992; **Table 1**). The human u-PAR promoter, first described by Wang *et al.* (1994) and Soravia *et al.* (1995), lacks TATA and CAAT boxes and contains a GC-rich proximal sequence with multiple Sp1 consensus elements. Up to now, at least nine different transcription factor binding sites (TFBSs) for the proximal human promoter have been described by others and our previous work in detail (**Fig. 1**). For some of them, we have recently suggested that they allow a further differentiation into a new prognostic high-risk group (Schewe *et al.*

Table 1 u-PAR gene (PLAUR) localization in different species.

Classification		Species	Chromosome*	Position ATG*	Myas[1]
Mammalia	Prim	*Homo sapiens* (human)	19q13.31	48866127	---
		Pan troglodytes (chimp)	20	45802351	5
		Macaca mulatta (Rhesus)	19	501298797	23
	Rod	*Mus musculus* (mouse)	7qA3	24171269	91
		Rattus norvegicus (rat)	1q21	79708657	91
	Lau	*Bos taurus* (cow)	18	4617587	92
		Canis familiaris (dog)	1	114369251	92

Prim.: *Primates*; Rod.: *Rodentia*; Lau.: *Laurasiatheria*.
*: Chromosome and gene position from UCSC Genome Bioinformatics (http://genome.ucsc.edu/) (release Aug. 10, 2006)
[1]: Evolutionary distances to *Homo sapiens* in million years (Myas) (Ureta-Vidal *et al.* 2003)

2003, 2005; Beyer *et al.* 2006; Maurer *et al.* 2007).

In this review we analyze in particular the *cis*-regulatory-modules (CRMs) of the basal promoter in detail by a species sequence comparison, trying to implicate a cross-species relevance for major promoter sites. Moreover, we discuss the TFBSs in the context of their potential *in vivo* and prognostic relevance.

COMPARATIVE GENOMICS

We analyzed human versus (vs) other organisms' sequences by comparative genomics. This is a method to predict biologically meaningful regions out of genomic sequences by assuming that functionally important DNA elements, such as exons and promoter elements, have a slower mutation rate than the genomic background. Therefore, these elements are evolutionary conserved across species, as opposed to non-conserved ones, which are supposed to be subjected to fewer functional constraints (Hardison 2000; Boguski 2002). Hence, we performed multiple proximal promoter (+66/-200) sequence alignments from different species (**Fig. 2**) to identify evolutionary conserved TFBSs of homologues to the human urokinase plasminogen activator receptor (u-PAR) gene PLAUR (Roldan *et al.* 1990; Min *et al.* 1992). Subsequently, these sites were analyzed with the known CRMs of the human gene.

Identification of homologous PLAUR genes

First, we searched for homologous PLAUR genes in different species (**Table 1**). The identification of *mammalian* sequences orthologous to the human PLAUR gene, was obtained using a combination of BLASTN, the NCBI Map Viewer (http://www.ncbi.nlm.nih.gov/mapview/), ENSEMBL (http://www.ensembl.org/index.html) and the UCSC Genome Bioinformatics (http://genome.ucsc.edu/). For the detection of orthologous genes, the intron-exon-structure, the protein sequence and the chromosomal context (flanking genes) were used as criteria (Nardone *et al.* 2004). By searching with different *mammalian* conserved amino acid residues of u-PAR as consensus sequences, we could not detect orthologous PLAUR genes in more distantly related species such as bird (*Gallus gallus*), amphibians (*Xenopus laevis*) or fish (*Danio rerio*), up to now. Only paralogues of the Ly-6/urokinase-type plasminogen activator receptor superfamily could be detected. As an exception, for the following species, homologous peptides for parts of the u-PAR protein were found: *Xenopus tropicalis* gene*:* ENSXETG00000015123, a 248 amino acid (aa) long peptide, *Tetraodon nigroviridis* gene: GSTENG00038774001, a 39

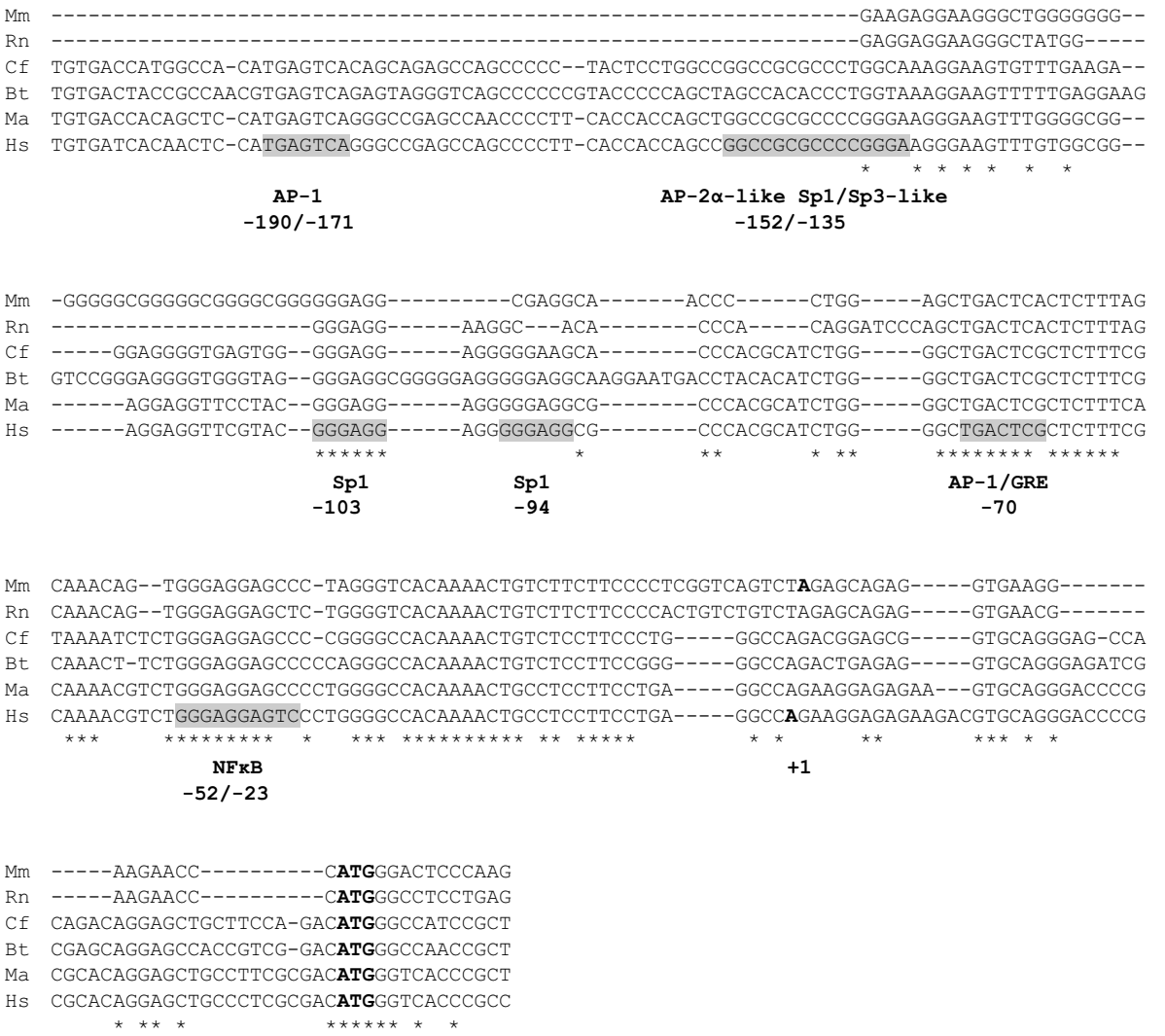

Fig. 2 Multiple DNA-sequence-alignments of u-PAR promoters (-200/+66). Sequence comparison of the ECR (evolutionary conserved region), including the basal promoters from various species such as *Mus musculus* (Mm), *Rattus norvegicus* (Rn), *Canis familaris* (Cf), *Bos taurus* (Bt), *Macaca mulatta* (Ma) and *Homo sapiens* (Hs). Note that the *Pan troglodytes* sequence is absolutely conserved with the *Homo sapiens* sequence in that particular region. In gray are known *cis*-regulated TFBSs for the human u-PAR-promoter. Bold: trancriptional start sites of the human and the murine u-PAR-transcript (Suh *et al.* 1994; Soravia *et al.* 1995); bold and underlined: Start-Met (ATG). Absolutely conserved nucleotides of the alignment are indicated with *. Numbers represent the positions in the human genome sequence (according to Soravia *et al.* 1995). The alignment was done with three different programs: MULAN (Ovcharenko *et al.* 2005; http://www.dcode.org/), CLUSTAL W (Thompson *et al.* 1994) and MUSCLE (Edgar 2004).

aa long peptide with a perfect homology to *Mus musculus*, *Fugu rubripes* gene: SINFRUG00000157419, a 157 aa long peptide (ENSEMBL database). The fact that we were not able to find orthologous PLAUR genes in those species may explain that PLAUR is an emerging gene, and that the listed species have a larger evolutionary distance to the common ancestor to *Homo sapiens*, e.g. for *Gallus gallus* 310 million years (Myas), 360 Myas for *Xenopus laevis* and 450 Myas for *Danio rerio* (Ureta-Vidal *et al.* 2003).

When we compared the marked start-methionine of u-PAR-proteins from different species with human u-PAR as a consensus sequence, some discrepancies in the NCBI database were found. The variation between the marked start in NCBI of *Rattus norvegicus* (rat) compared to *Homo sapiens* were 21 base pairs (bp) downstream of the start, 228 bp for *Canis familaris* (dog) and 42 bp for *Pan troglodytes* (chimpanzee).

The identifed homologous sequences were multiply aligned with Mulan (Ovcharenko *et al.* 2005; http://www.dcode.org/) or MUSCLE (Edgar 2004) to discover common *mammalian*-specific gene regulatory elements (**Fig. 2**).

Analysis of the proximal promoter

Human to *mammalian* sequence comparison identified conserved *mammalian*-specific regulatory 5′-UTR elements (Venkatesh and Yap 2005). Downstream to the human PLAUR gene, the IGSF4C gene (immunoglobulin superfamily, member 4C) and upstream the IRGC1 gene (immunity-related GTPase family, cinema 1) are localized. As we mentioned above (section: Identification of homologous PLAUR genes), the surrounding genes are the same in all species analyzed. It is possible that the already identified CRM also have an influence on these or other neighbouring genes. Although regulatory sequences can be widely dispersed within the genome, they may lie in the intronic regions, or they may be located within neighbouring genes in the surrounding areas, and the respective RNA may be able to act as a gene regulator.

With the phylogenetic footprinting method, only conserved regions are identified, which are of general biological importance. However, species-specific binding sites may not be conserved, so they should be analyzed with phylogenetic shadowing (Boffelli *et al.* 2003). Furthermore, bioinformatic prediction of regulatory regions is only an adjunct to biological experiments. Therefore, the next step should be to evaluate the biological relevance of the conserved elements in different cell lines, tissues and species, a task which has been initiated by us and others since many years (see final chapters).

SPECIES-SPECIFIC PROXIMAL PROMOTER ANALYSIS

To get an estimate of the sequence importance of the respective nucleotides in the proximal promoter motif, we performed sequence alignments of human PLAUR gene with the orthologous genes from various species (**Fig. 2**). Accordingly, in this particular region (-200/+66) the nucleotide conservation in *Pan troglodytes* (chimpanzee) is 100%, and in *Macaca mulatta* (rhesus monkey) 96%. The sequence identity in human vs. dog is 82.4%, human vs. cow 80.3%, human vs. rat 69.3% and human vs. mouse 68.7% (according to Soravia *et al.* 1995). **Fig. 2** shows a sequence alignment of a section spanning 266 bp including 66 bp of the first exon of the previously characterized proximal promoter region -190/-1 of the human u-PAR, which drives the basal expression of the PLAUR gene (Lengyel *et al.* 1996; Dang *et al.* 1999).

Kozak sequence

In general, the optimal site for initiation of translation (kozak sequence) in *mammals* is GCC**R**CCaug**G** (Kozak 2005; R indicates A or G). This consensus sequence fits well, notably the two most important sites for initiation, the -3 R and +4 G site show 100% match to the u-PAR region of *primates* and *laurasiatheria* but not to *rodentia* sequences. In these species (mouse and rat), the -3 R site is exchanged to C, which is normally associated with a reduction of translation as compared to the "optimal" initiation site. For example, a pathological consequence of the G>C replacement at this specific position occurs in the human BRCA1 5'-UTR. In *in vitro* transcription/translation assays, this point mutation in the 5′-UTR of BRCA1 leads to 30-50% reduction in translation efficiency as compared to the wild type BRCA1 5′-UTR (Signori *et al.* 2001).

Transcriptional start and upstream regions

As reported previously, one of the human transcription start sites in HeLa and U937 cell lines coincides with the major transcriptional start site as determined by primer extension and RNase protection assay with mRNA of murine cells (Suh *et al.* 1994; Soravia *et al.* 1995). The surrounding area of the major transcriptional initiation site of human u-PAR correlates with rhesus monkey, dog and cow, but less with that of rat and mouse (**Fig. 2**, see below).

In the species analyzed, the corresponding sequence from -8 to -69 is highly conserved, especially the TFBS for NFκB, AP-1, as well as a region spanning from -8 to -30. Sequence analysis with TESS (http://www.cbil.upenn.edu/cgi-bin/tess/tess) reveals a potential E twenty-six (Ets), c-Ets-1 binding site [5′-CTTCCTG-3′] within -8 to -30. These factors are involved in the regulation of a tissue-specific inhibitor of metalloproteinase 1 and enhance transcription synergistically with AP-1 (Logan *et al.* 1996). Nevertheless, with MultiTF (http://www.dcode.org/), we identified MYB, PAX2, CAP, cDXA (all in region -16 to -26) and STATs (in -8 to -14) as potential binding sites within the -8 to -30 sequence.

Dubchak *et al.* (2000) showed that a three-way comparison of human, mouse and dog sequences helps to define putative CRMs in non-coding DNA. The hypothesis is that the conservation of the non-coding human sequence in this locus with species more distantly related than *primates* does not extend beyond the boundaries of functionally relevant elements. Therefore, the non-coding sequences conserved in "distant" species should highlight potential functional elements within this locus. These settings have been shown to provide a high sensitivity for analyzing human/mouse conervation profiles and thus potential, regulatory elements (Loots *et al.* 2000), since for example, when assuming no selective pressure according to the Jukes–Cantor equation (Jukes and Cantor 1969), human–mouse sequences are estimated to be 51% identical. PLAUR is a fast-diverging locus since the sequence conservation of the 5′-non-coding homologous u-PAR region in different *mammalians* is low (with the exception of other *primates*, where it is highly conserved).

NFκB (-52/-23)

The human NFκB motif -52/-23 (Wang *et al.* 2000b) is strongly conserved among the different species (**Fig. 2**). The decameric consensus sequence of NFκB is naturally subdivided in two blocks (A/B). The "A-block" [5′-GGGRN-3′, N indicates any nucleotide], especially the three G nucleotides in position 1-3, appear to be necessary for the binding of the p50 subunit (Parry and Mackman 1994), so this section is absolutely conserved [5′-GGGAG-3′]. This implicates that p50 might be one subunit of the NFκB dimeric complex that may bind to that site in different species. This observation is consistent with Wang *et al.* (2000b) who showed that a mutated construct in the "A-block", 5′-GGGA GGAGTC-3′ to 5′-tGtAGGAGTC-3′ impaired u-PAR promoter activity. In luciferase assays, this mutated construct of the NFκB motif reduced u-PAR promoter activity between 6-11 fold in human HCT116 colon cancer cells. The second "block B" [5′-GAGYY-3′, where Y indicates C or

T] does not match with the natural consensus sequence of the NFκB motif [5′-NYYCC-3′ N indicates any nucleotide], however the specific sequence motif variations [5′-GAG-3′] within the beginning of "block B" are conserved in contrast to the last two nucleotides [5′-YY-3′ instead of 5′-CC-3′]. These last nucleotides match the original NFκB motif in rhesus monkey, mouse, cow and dog only, but not in men or rat, indicating that, potentially, separate isoforms of NFκB may bind to that region in these different groups. Using electro-mobility-shift-assays (EMSA), two specific bands from nuclear extracts of human HCT116 cells were detected and separated, and it has been shown that p49 also binds to this NFκB site (Wang *et al.* 2000a). Until now, structural studies could not provide evidence for highly specific interactions between NFκB and its cognate DNA binding sites. X-ray structures of p50/p65 heterodimers demonstrate that they can bind to a variety of κB site sequences (Chen and Ghosh 1999; Berkowitz *et al.* 2002; Chen-Park *et al.* 2002; Escalante *et al.* 2002).

AP-1 (-70 and -190/-171)

Fos and Jun family bZIP proteins form homo- and heterodimers *via* a leucine zipper dimerization interface and bind to palindromic AP-1 recognition elements [5′-TGA(C/G) TCA-3′] *via* a basic DNA contact region. In the human proximal promoter region of u-PAR, two AP-1-motifs at -70 and -187 are located (Lengyel *et al.* 1996; Okan *et al.* 2001) (**Fig. 2**). The human -69 site is present as a 1-bp mismatch [5′-TCA-3′ to 5′-TCG-3′] in all the species analyzed, only in rodents the "original" [5′-TGA(C/G)TC**A**-3′] site is preserved. In that sequence variation, a guanine (instead of an adenine) in this position is compatible with the binding of, e.g., Fos-Jun heterodimers to the DNA (Ramirez-Carrozzi and Kerppola 2003). A guanine in that position may directly interfere with the binding intensity and the conformation (Ramirez-Carrozzi and Kerppola 2003) of AP-1 subunits and can influence gene expression, so maybe this "strategic" position has a specific task in "large" animals (see below). The upstream human AP-1 site (-190/-171) is found in the species human, cow and dog only, and all three sequences match 100% with the AP-1 consensus sequence. In RKO, HCT116 and GEO human colon cancer cells, this motif is bound by Jun-D, c-Jun, Fra-1 and or c-Fos (Lengyel *et al.* 1996; Allgayer *et al.* 1999c) and also in the intestinal epithelial cell line IEC4-1 by Jun-D and Fra-2 after induction with TGFβ (Yue *et al.* 2004). In CAT-reporter assays of RKO and HCT116 cells, the mutation of the AP-1 site at -190/-171 (TGAGTCA to TatcTCA) results in 80% or 70% reduced promoter activity when compared to the wild-type promoter, respectively. The consensus AP-1 element is not perfectly symmetrical, since the central C/G base pair results in two different overlapping half-sites [5′-TGA**C**-3′ or 5′-TGA**G**-3′]. Interestingly, both AP-1 sites (at -70 and -190/-171) differ in their core cytosine / guanine and can potentially bind different subunits of AP-1-members. Fos-Jun heterodimers and Jun homodimers have similar, however not identical binding preferences for variants of the AP-1 recognition sequence (Kerppola and Curran 1994). It remains unclear whether Fos and Jun have different DNA recognition specificities for the core AP-1 regulatory element of this particular region, and whether such differences may influence the structural and functional characteristics of Fos-Jun heterodimers.

Sp1 (-94 and -103)

Sp1 is known to be a general activator of transcription (Suske 1999). Soravia *et al*. reported that the basal expression of the PLAUR gene is regulated by multiple Sp1 motifs proximal and upstream of the transcriptional start site. In the proximal promoter region of human PLAUR, two Sp1 sites (-94 and -103) and one Sp1-like motif (-152/-135) are located (Soravia *et al.* 1995; Allgayer *et al.* 1999) (**Fig. 2**). These sites (-94 and -103) are necessary for PLAUR expression in human monocyte-like U937 cells and furthermore for the TGFβ stimulation of that gene (Park *et al.* 2000). There are two alternative ways to align the human Sp1-sites at -94 and -103. First, the -94 site matches 100% in cow, mouse, and with one mismatch, in dog [5′-GGGA **G**G-3′ to 5′-GGGA**A**G-3′], but not in rat, and the second Sp1-site at -103 is conserved in cow, dog and rat, but not in mouse (data not shown). In the second scenario (**Fig. 2**), after a new arrangement of the sites around -103 and -94, in contrast to Soravia *et al*. it seems that the second locus (-103) only is conserved across the mammalian phylogeny (of the analyzed species). In this arrangement, the first Sp1-site (-94) remains in cow, but not in mouse and dog. In both possibilities, the two human Sp1-sites are always conserved in the cow sequence.

Another Sp-site is located in the u-PAR promoter at region -402/-350 (Hapke *et al.* 2001a). Within this area a sequence, reaching from -380/-354 (not shown in **Fig. 2**), is co-mediating downregulation of u-PAR promoter activity through the novel tumor supressor Pdcd4 in colorectal cancer cells (Leupold *et al.* 2007a). However, the exact binding positions of the Sp factors are not yet defined. This region is absolutley conserved from -354 to -372 in primates and to a lesser extent in dog and cow, but not in rodents, whereas the last 7 bp (-373/-380) of that part are rarely conserved, in all species analyzed.

AP-2α-like Sp1/Sp3-like-motif (-152/-135)

The complete AP-2α-like Sp1/Sp3-like-motif (-152/-135) as described by Allgayer *et al.* (1999), spanning approximately 15 bp, exists only in cow, dog and *primates*. The same results were revealed by an alternative CLUSTAL W (Thompson *et al.* 1994) and MUSCLE (Edgar 2004) alignment. In that specific region, 13/15 in dog and 10/15 nucleotides in cow are conserved respectively, as compared to human. Segment -147/-141 of *Canis familiaris* exhibits conserved nucleotides which are necessary for the binding of the human Sp1-/3-like-factors [5′-GGCCGCG-3′] (Allgayer *et al.* 1999). This has been reproduced by mutation analysis of the human sequences (Allgayer *et al.* 1999). In contrast, in the sequence of *Bos taurus*, most of these specific nucleotides are substituted with adenine. In human colon cancer cell lines, as shown with mutated CAT-reporter constructs, the Sp1/3 motif is especially important for induction with the *c-src*-oncogene and also for suppression of u-PAR by Pdcd4 (Allgayer *et al.* 1999b; Leupold *et al.* 2007a). Interestingly, the basepairs important for the binding of the AP-2α-like protein being closely related to, however not identical to, authentic AP-2α, are absolutely conserved in dog as well as in cow. Binding of the AP-2α-like protein was found to be important for a constitutively high u-PAR-promoter activity in a highly invasive colon cancer cell lines, and for PMA-stimulated u-PAR expression in a cell line with low constitutive u-PAR expression (Allgayer *et al.* 1999).

PEA3/ets (-248)

The region around -248 contains a potential silencer element of PLAUR. Experimental data (e.g. CAT reporter assay, EMSA) indicate that a PEA3-element at -248 bp of the sequence acts as a mediator of human integrin-induced suppression in CHO cells (chinese hamster ovary, Hapke *et al.* 2001a). The canonical binding sequence for that site is 5′-AGGAAG-3′ (not included in the Figures). The sequence is absolutely conserved in *primates* and with only one A to G mismatch [5′-AGGA**G**GA-3′] in dog, whereas in other species there is very low (cow only 2 bp match) or no (mouse and rat) conservation. Interestingly, the experiments were done in a hamster cell line with the human promoter sequence, although there is no conservation in the *rodent* promoter sequences at this specific site. However, the authors still detected a biological effect of this site in this system. The influence of this specific site is also shown in ovarian cancer cells (OV-MZ-6) were it plays a role in repres-

sion of the u-PAR gene expression by integrins (Hapke *et al.* 2001b).

Phylogeny spacing of important sites

It has been shown that promoters and other CRM can usually tolerate much more sequence divergence than coding regions and still retain their original functions (Haubold and Wiehe 2004). The assembly of the protein transcription complex can be altered by artificial lengthening of DNA stretches spanning between binding sites, which can lead to unpredictable deregulatory effects (Bonifer 2000). TF and cofactors are expressed differentially in different species, thus the varying TFBSs have a diverse importance for the gene regulation. Spacing and position of the individual TFBSs may be species-specific (some interactions are precisely phased to lie on the same side of nucleosomes (see Lewin 2000)). Seeing that, we investigated the phylogeny sequence space of ECRs mapped in human to TFBSs in reference to the transcriptional start sites in the proximal u-PAR promoter region (**Fig. 2**). Up to now, the transcriptional start sites for men and mouse are reported only (Soravia *et al.* 1995). The authors use primer extension analysis to show three potential transcriptional start sites, the most upstream of which – an A following a C – appeared to be the main transcription initiation site in humans and revealed partial similarity to the consensus initiator sequence of the dihydrofolate reductase (DHFR) gene (Means and Farnham 1990). For mouse, the transcriptional start site, located (Suh *et al.* 1994) upstream as compared to human, also begins with an --A-- which is conserved across the species. The direct flanking area of the transcriptional start is conserved in *primates*, cow and dog as opposed to *rodents*. The relative position (spacing) between the different TFBSs in the proximal promoter region varies between species. In dog only, the distance among TFBSs and the transcriptional start site compared to *primates* is almost the same (three 1 bp gaps upstream the transcriptional start site up to the AP1-site at -190/-171).

EVOLUTIONARY EXAMINATION

Despite 92 Myas (**Table 1**) of evolution separating, the most conserved sites are the NFκB sequence at -52/-23, the AP-1 at -70 and the Sp1 site at -103, which contain up to one nucleotide substitution. This observation suggests that these sites have an evolutionary conserved, superior biological importance in general gene regulation of the *mammalian* homologous PLAUR locus. The absolutely identical alignment between human vs chimpanzee shows that there are no anew (5 Myas) changes in DNA sequence of the basal PLAUR promoter. The latest variances (23 Myas) are substitutions in the rhesus monkey sequence at the NFκB-motif (-52/-23), so this T to C mutation is a highly human specific replacement.

Besides that, a GAC-insertion lies 12 bp 5′-upstream of the human trancriptional start site. When we compared the *laurasiatheria* dog vs. cow (83 Myas) and the *rodentia* rat vs. mouse (41 Mays), the kozak-sequence, the area around the transcriptional start site and the most species-specific TFBS-motifs are conserved in their special super order. In our analysis, the species related closest to men and *primates* are mouse and rat (91 Myas), which are combined in the *euarchontoglires* separated of the *laurasiatheria* (**Table 1**). Interestingly, the basal promoter sequences (-200/+66) from mouse and rat show less than 70% identity compared to men and the TFBS regions only 44% identity to human. In contrast, in the most distantly related *laurasiatheria* (dog and cow) it is above 80% identity for -200/+66, and 86% for the specific TFBSs (**Fig. 2**).

Although rodents have approximately the same evolutionary distance to human (~90 Myas) when compared to dog and cow, the TFBSs of "small" *mammalians* in comparison to "bigger" *mammalians* differ much stronger than is suggested by the genetic background. For example, the human AP-2α-like-motif at -152/-135, the AP-1-site at -190/-171 and the variation at -69 (A to G) of u-PAR remains in *primates*, dog and cow, but not in mouse and rat. These differentially species-specific conservation patterns of different TFBSs are likely to play a role in species-specific regulation. *Primates*, dog and cow are all "larger" organisms in comparison to rodents which have, e.g., another body temperature, heart rate, substantially diverse anatomy, physiology, generation time, wound healing physiology, etc. It is interesting to speculate, that the effect of the previously characterized human TFBSs on the gene expression of *mammalian* PLAUR homologous might have evolved, among other things, in parallel to the animal size. A comparison of 20 well-characterized regulatory regions in *mammals* revealed that approximately one third of binding sites in humans are probably not functional in rodents (Dermitzakis and Clark 2002).

CONSERVED MOTIFS UPSTREAM OF THE BASAL PROMOTER

To identify additional conserved 5′-upstream sequences, we performed an extended analysis with Mulan (Ovcharenko *et al.* 2005; http://www.dcode.org/) or MUSCLE (Edgar 2004) of approximately 12 kb from human ATG (**Table 1**). We accepted conserved non-coding sequences of >100 bp lengths with at least 70% identity only. These settings provide a high sensitivity for analyzing human/mouse conservation profiles and thus provide appropriate tools to determine potential regulatory elements (Loots *et al.* 2000). With these chosen parameters, three non-coding conserved regions at -123/+53 (basal promoter), -2168/-1870 and at -5583/-5343 were detected. The regions at -123/+53 and -5583/-5343 contain putative conserved binding sites for NFκB, AP-1 and USF (upstream stimulating factor).

IN VIVO GENE REGULATION AND POTENTIAL PROGNOSTIC RELEVANCE

As we mentioned above, u-PAR is expressed by the leading edge of re-epithelializing wounds, and plays a role in physiological processes such as wound healing, angiogenesis and embryogenesis. Moreover, it is overexpressed in several cancers and mediates invasion, progression, and metastasis among other phenomena. It has been characterized quite extensively how an upregulation of u-PAR occurs *in vitro*, and some of the TFBSs we identified in sequence alignments have been shown to be important u-PAR regulators *in vitro*, however, few studies so far addressed transcriptional mechanisms regulating u-PAR *in vivo*. Wang *et al*. (2003) defined u-PAR promoter regions required *in vivo* for the expression of this gene in transgenic mice bearing a LacZ reporter regulated by varying amounts (-400, -1500, and -8500 bp) of upstream sequence. The -400 bp u-PAR promoter directed weak and strong LacZ expression in the placenta and epididymis, respectively, both of which are tissues that express endogenous u-PAR. Conversely, transgene expression in the apical cells of the colon positive for endogenous u-PAR protein required -1500 bp of upstream sequence for optimal expression, containing the basal promoter and certain upstream sequences discussed before (section: conserved motifs upstream of the basal promoter). Placental transgene expression was augmented with the -8500-bp upstream fragment compared to the shorter -1500-bp fragment, indicating additional element(s) between -1500 to -8500 bp for optimal expression. These data suggest new upstream sequences especially in the region spanning -1500 to -8500 bp for tissue-specific u-PAR expression from a transgenic *in vivo* model. Interestingly, with our extended sequence alignment (chapter above) we could detect one conserved region within this specific area at -5583/-5343.

Besides mouse models, the CAM (chorionallantoic membrane) model can give evidence for an *in vivo* relevance of data, by specifically measuring an impact on tumor cell intravasation. Leupold *et al*. (2007b) used this model

and showed that the AP-1 site at -190/-171 is one of the mediators of Src-induced intravasation of colorectal cancer cells. Taken together, some studies suggested evidence for an important function of particular u-PAR promoter- and upstream motifs in *in vivo* models.

In addition, first own u-PAR promoter studies were preformed at resected patient tissues. In a study on 145 patients with resected colorectal or gastric cancers, we investigated the AP-2α-like Sp1/Sp3-like-site (from -152 to -135), which an intensive BLASTN search revealed to be highly specific and unique for the PLAUR gene. Especially, the AP-2α-like binding site is 100% conserved in dog, cow and primates. We found an almost tumor-specific transcription factor binding to this important u-PAR promoter motif of especially AP-2α-like protein and Sp1, in approximately 60% of cases when comparing primary tumors and corresponding normal mucosae, this being significantly associated with high endogenous u-PAR protein amounts in the tumor tissues (Schewe *et al.* 2003).

As indicated above, a further conserved motif, important for PLAUR gene regulation in cancer had been implicated by previous *in vitro* studies, which was a consensus AP-1 region (-190/-171). In addition, our *in vitro* studies have suggested a synergism between this AP-1 motif, and the combined AP-2α-like Sp1/Sp3-like-motif in cell lines (Lengyel *et al*. 1996; Allgayer *et al*. 1999c). For this "second", conserved AP-1 motif we performed an additional translational study on 103 colorectal cancer patients comparing resected tumor and normal tissues, asking for tumor-specificity of this motif and *in vivo* evidence of synergism (Schewe *et al*, 2005). Tumor-specific AP-1-binding to conserved region -190/-171 of the u-PAR-promoter was found in 40% of patients, which is less than for the AP-2α-like Sp1/Sp3-like-motif. The AP-1-binding correlated significantly with u-PAR protein amounts in both normal and tumor tissues ($p<0.001$), this being in contrast to a tumor-specific correlation with u-PAR of the AP-2/Sp1-region. In analysis for both promoter regions, 62% of cancers showed simultaneous binding for AP-1, AP-2 and Sp1, 11% for AP-1 and AP-2, and 16% for AP-2 and Sp1, and a minority of cases for binding of one factor only. The binding of AP-1, AP-2 and Sp1 correlated significantly with each other ($p<0.001$), and the combination of AP-1 and AP-2 showed the highest correlation with u-PAR ($p=0.008$). Preliminary survival analysis indicated a trend for poorer prognosis for binding of all three factors. Therefore, the conserved AP-1-site -190/-171 appeared to be a less tumor-specific regulator than the Sp1/AP-2-motif -152/-135, which is again interesting for future potential clinical consequences. Data furthermore corroborated the hypothesis of a synergism between both elements in resected tumors.

This is also congruent with the observation of Maurer *et al.* (2007), who found a decreased survival of colorectal cancer patients with the number of bound TFs to both motifs (-190/-171 and -152/-135). The binding of all transcription factors (AP-1, Sp1 and AP-2α-like) defined novel high-risk groups for disease-specific survival. Furthermore, with TFs bound to the u-PAR promoter, a first molecular staging model in colorectal cancer patients was defined. The authors demonstrated that TF binding to the u-PAR promoter has an earlier prognostic relevance than the u-PAR protein itself (Maurer *et al.* 2007).

These studies demonstrated that certain promoter elements seem to be employed differentially in different human tissues, may be specifically activated by the tumor to promote invasion/metastasis *via* upregulation of this gene, and implicate novel high-risk groups in cancer that can be predicted very early in clinical follow-up. Certainly such tissue-dependent specifications cannot be predicted with theoretical sequence alignment.

However, bioinformatic sequence alignment is a tool to reduce the amount of data to be considered as functional TFBSs for time-consuming *in vivo* studies.

WHICH FACTOR AND/OR *CIS*-REGULATORY-MODULES IS MOST IMPORTANT FOR THE EXPRESSION OF HUMAN U-PAR?

This question can not be answered in a general way. There are different levels of u-PAR gene expression in diverse tissues and organs (Wang *et al.* 2003), and a differential importance of different TFs and binding sites for expression of u-PAR in diverse scenarios. In tissues such as placenta, conserved sites such as -5583/-5343 within -8500 bp upstream are very likely to be important for u-PAR gene expression (Wang *et al.* 2003). In contrast, experiments done in RKO and GEO cancer cells as well as clinical studies in colon and gastric cancer patients indicate that the AP-2α-like Sp1/Sp3-like-motif at -152/-135, and the AP-1 motif at -190/-171, are highly relevant for strong expression of PLAUR in solid cancers (Lengyel *et al.* 1996; Allgayer *et al.* 1999; Schewe *et al.* 2003, 2005; Maurer *et al.* 2007). In addition, it can be speculated that diverse members of the AP-1 family bind differentially to this site (Lengyel *et al.* 1996) in cells derived of various tissues, and thereby finetune the u-PAR gene expression (Schewe *et al.* 2005). Moreover, a synergism of different TFBSs is important for expression *in vivo*. In addition to main promoter motifs mediating several means of u-PAR-control, further promoter motifs serve as co-modulators of u-PAR-regulating pathways. For example, an upstream Sp1-motif at -380/-354 mediates Pdcd4-induced u-PAR-suppressing in addition to Sp3 bound to motif -152/-135.

CONCLUSION AND PERSPECTIVES

At the "pure" sequence level, the human TFBSs previously characterized as functional, and essential for u-PAR gene expression, were identified mostly as interspecies-specific TFBSs (ECRs, evolutionary conserved regions), and the relevance *in vivo* for many of them has not yet been investigated. However, suchlike studies might harbor a potential for diagnosis and even treatment. For example, as mention before, in a recent study we have seen that simultaneous TF-binding to some sites implicate a worse clinical outcome for cancer patients (Schewe *et al.* 2003, 2005; Maurer *et al.* 2007). However, since these sites are still less conserved (mostly in *primates*, *Canis familiaris* and *Bos taurus*) as compared to more conserved 3′-downstream regions, it may be additionally interesting to investigate the functional and clinical/prognostic relevance of these even more conserved sites. On the other hand, it is interesting to speculate that cancer may use less conserved upstream regions to destabilize steady-state gene expression brought about by well-conserved downstream motifs.

We conclude that most of the distal promoter regions, especially the human TFBSs of the PLAUR gene discussed in this paper, are also present in the PLAUR orthologues of different *mammalian* species (cow, dog, rat and mouse). In particular, the NFκB sequence at -52/-23, the AP-1-site at -70 and the Sp1 site at -103 are most conserved, and therefore interspecies and *mammalian* specific motifs are potentially involved in a common *mammalian* transcriptional regulation of PLAUR orthologues. The AP-2-like-binding motif at -152/-135 as well as the AP-1 site at -190/-171 are conserved in *primates*, C*anis familiaris* and in *Bos taurus*, but not in *rodents*. The Sp1 site at -93 (conserved in *Bos taurus*), the Sp1-/3-like-motif at -152/-135 (*Canis familiaris*), the PEA3/ets site at -248 and the Sp1/Sp3 at -380/-354 seem to be are more "species-specific" motifs. However, especially for the AP-2α-like Sp1/Sp3-like-motif and the AP-1 motif at -190/-171, an *in vivo* and also prognostic relevance has already been shown for cancer patients. The biological significance of further entirely conserved sites (TFBS in the basal promoter of different species) still remains to be tested.

Taken together, combining sequence studies with *in vivo* investigations is certainly an interesting tool to elucidate transcriptional regulators of gene expression relevant for

diseases.

ACKNOWLEDGEMENTS

H. A. was supported by the Alfried Krupp von Bohlen und Halbach Foundation, Essen, Wilhelm Sander Stiftung, Munich, the Auguste-Schaedel-Dantscher-Stiftung, Garmisch, Germany, and the Hella-Bühler Foundation, Heidelberg, Germany. We gratefully thank Dr. Beate K. Straub (Insitute of Pathology, University Heidelberg, Heidelberg, Germany) for constructive discussions and critical reading of the manuscript.

REFERENCES

Allgayer H, Wang H, Wang Y, Heiss MM, Bauer R, Nyormoi O, Boyd D (1999a) Transactivation of the urokinase-type plasminogen activator receptor gene through a novel promoter motif bound with an activator protein-2α-related factor. *Journal of Biological Chemistry* **274**, 4702-4714

Allgayer H, Wang H, Gallick GE, Crabtree A, Mazar A, Jones T, Kraker AJ, Boyd DD (1999b) Transcriptional induction of the urokinase receptor gene by a constitutively active Src. Requirement of an upstream motif (-152/-135) bound with Sp1. *Journal of Biological Chemistry* **274**, 18428-18437

Allgayer H, Wang H, Shirasawa S, Sasazuki T, Boyd D (1999c) Targeted disruption of the K-ras oncogene in an invasive colon cancer cell line down-regulates urokinase receptor expression and plasminogen-dependent proteolysis. *British Journal of Cancer* **80**, 1884-1891

Allgayer H (2006) Molecular regulation of an invasion-related molecule – options for tumour staging and clinical strategies. *European Journal of Cancer* **42**, 811-819

Behrendt N, Ronne E, Ploug M, Petri T, Lober D, Nielsen LS, Schleunig WD, Blasi F, Appella E, Dano K (1990) The human receptor for urokinase plasminogen activator. NH_2-terminal amino acid sequence and glycosylation variants. *Journal of Biological Chemistry* **265**, 6453-6460

Berkowitz B, Huang DB, Chen-Park FE, Sigler PB, Ghosh G (2002) The x-ray crystal structure of the NF-κ B p50.p65 heterodimer bound to the interferon beta -kappa B site. *Journal of Biological Chemistry* **277**, 24694-24700

Beyer BC, Heiss MM, Simon EH, Gruetzner KU, Babic R, Jauch KW, Schildberg FW, Allgayer H (2006) Urokinase system expression in gastric carcinoma: prognostic impact in an independent patient series and first evidence of predictive value in preoperative biopsy and intestinal metaplasia specimens. *Cancer* **106**, 1026-1035

Bianchi E, Cohen RL, Thor AT, Todd RF 3rd, Mizukami IF, Lawrence DA, Ljung BM, Shuman MA, Smith HS (1994) The urokinase receptor is expressed in invasive breast cancer but not in normal breast tissue. *Cancer Research* **54**, 861-866

Blasi F (1988) Surface receptors for urokinase plasminogen activator. *Fibrinolysis* **2**, 73-84

Blasi F (1993) Urokinase and urokinase receptor: A paracrine/autocrine system regulating cell migration and invasiveness. *BioEssays* **15**, 105-111

Boguski MS (2002) Comparative genomics: the mouse that roared. *Nature* **420**, 515-516

Boffelli D, McAuliffe J, Ovcharenko D, Lewis KD, Ovcharenko I, Pachter L, Rubin EM (2003) Phylogenetic shadowing of primate sequences to find functional regions of the human genome. *Science* **299**, 1391-1394

Bonifer C (2000) Developmental regulation of eukaryotic gene loci: which *cis*-regulatory information is required? *Trends in Genetics* **16**, 310-315

Borglum AD, Byskov A, Ragno P, Roldan AL, Tripputi P, Cassani G, Dano K, Blasi F, Bolund L, Kruse TA (1992) Assignment of the urokinase-type plasminogen activator receptor gene (PLAUR) to chromosome 19q13.1-q13.2. *American Journal of Human Genetics* **50**, 492-497

Chen FE, Ghosh G (1999) Regulation of DNA binding by Rel/NF-κB transcription factors: structural views. *Oncogene* **18**, 6845-52

Chen-Park FE, Huang DB, Noro B, Thanos D, Ghosh G (2002) The κ B DNA sequence from the HIV long terminal repeat functions as an allosteric regulator of HIV transcription. *Journal of Biological Chemistry* **277**, 24701-24708

Courey AJ (2001) Regulatory transcription factors and *cis*-regulatory regions. In: Locker J (Ed) *Transcription Factors*, Academic Press, San Diego, CA, pp 17-34

Dang J, Boyd D, Wang H, Allgayer H, Doe WF, Wang Y (1999) A region between -141 and -61 bp containing a proximal AP-1 is essential for constitutive expression of urokinase-type plasminogen activator receptor. *European Journal of Biochemistry* **264**, 92-99

Dano K, Andreasen PA, Grondahl-Hansen J, Kristensen P, Nielsen LS, Skriver L (1985) Plasminogen-activators, tissue degradation, and cancer. *Advances in Cancer Research* **44**, 139-266

de Bock CE, Wang Y (2004) Clinical significance of urokinase-type plasminogen activator receptor (uPAR) expression in cancer. *Medicinal Research Reviews* **24**, 13-39

Dermitzakis ET, Clark AG (2002) Evolution of transcription factor binding sites in Mammalian gene regulatory regions: conservation and turnover. *Molecular Biology and Evolution* **19**, 1114-1121

Dieterich C, Cusack B, Wang H, Rateitschak K, Krause A, Vingron M (2002) Annotating regulatory DNA based on man-mouse genomic comparison. *Bioinformatics* **18 (Suppl 2)**, 84-90

Dubchak I, Brudno M, Loots GG, Pachter L, Mayor C, Rubin EM, Frazer KA (2000) Active conservation of noncoding sequences revealed by three-way species comparisons. *Genome Research* **10**, 1304-1306

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792-1797

Ellis V, Behrendt N, Dano K (1991) Plasminogen activation by receptor-bound urokinase. *Journal of Biological Chemistry* **266**, 12752-12758

Escalante CR, Shen L, Thanos D, Aggarwal AK (2002) Structure of NF-κB p50/p65 heterodimer bound to the PRDII DNA element from the interferon-β promoter. *Structure* **10**, 383-391

Fernebro E, Madsen RR, Ferno M, Brunner N, Bendahl P, Christensen IJ, Johnson A, Nilbert M (2001) Prognostic importance of the soluble plasminogen activator receptor, suPAR, in plasma from rectal cancer patients. *European Journal of Cancer* **37**, 486-491

Fuchs T, Allgayer H (2003) Transcriptional regulation of the urokinase receptor (u-PAR) – a central molecule of invasion and metastasis. *Biological Chemistry* **384**, 755-761

Hanahan D, Weinberg RA (2000) The hallmarks of cancer. *Cell* **100**, 57-70

Hapke S, Gawaz M, Dehne K, Kohler J, Marshall JF, Graeff H, Schmitt M, Reuning U, Lengyel E (2001a) β(3)A-integrin downregulates the urokinase-type plasminogen activator receptor (u-PAR) through a PEA3/ets transcriptional silencing element in the u-PAR promoter. *Molecular and Cellular Biology* **21**, 2118-2132

Hapke S, Kessler H, Arroyo de Prada N, Benge A, Schmitt M, Lengyel E, Reuning U (2001b) Integrin α(v)β(3)/vitronectin interaction affects expression of the urokinase system in human ovarian cancer cells. *Journal of Biological Chemistry* **276**, 26340-26348

Hardison RC (2000) Conserved noncoding sequences are reliable guides to regulatory elements. *Trends in Genetics* **16**, 369-372

Haubold B, Wiehe T (2004) Comparative genomics: methods and applications. *Naturwissenschaften* **91**, 405-421

Hollas W, Blasi F, Boyd D (1991) Role of the urokinase receptor in facilitating extracellular matrix invasion by cultured colon cancer. *Cancer Research* **51**, 3690-3695

Huai Q, Mazar AP, Kuo A, Parry GC, Shaw DE, Callahan J, Li Y, Yuan C, Bian C, Chen L, Furie B, Furie BC, Cines DB, Huang M (2006) Structure of human urokinase plasminogen activator in complex with its receptor. *Science* **311**, 656-659

Jankun J, Merrick HW, Goldblatt PJ (1993) Expression and localization of elements of the plasminogen activation system in benign breast disease and breast cancers. *Journal of Cellular Biochemistry* **53**, 135-144

Jukes T H, Cantor C R (1969) Evolution of protein molecules. In: Munro HN (Ed) *Mammalian Protein Metabolism*, Academic Press, New York, pp 21-123

Karin M (1995) The regulation of AP-1 activity by mitogen-activated protein kinases. *Journal of Biological Chemistry* **270**, 16483-16486

Kerppola TK, Curran T (1994) A conserved region adjacent to the basic domain is required for recognition of an extended DNA binding site by Maf/Nrl family proteins. *Oncogene* **9**, 3149-3158

Kozak M (2005) Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene* **21**, 13-37

Lengyel E, Wang H, Stepp E, Juarez J, Doe W, Pfarr CM, Boyd D (1996) Requirement of an upstream AP-1 motif for the constitutive and phorbol ester-inducible expression of the urokinase-type plasminogen activator receptor gene. *Journal of Biological Chemistry* **271**, 23176-23184

Leupold JH, Yang HS, Colburn NH, Asangani I, Post S, Allgayer H (2007a) Tumor suppressor Pdcd4 inhibits invasion/intravasation and regulates urokinase receptor (u-PAR) gene expression via Sp-transcription factors. *Oncogene* in press

Leupold JH, Asangani I, Lengyel E, Post S, Allgayer H (2007b) Src induces urokinase-receptor (u-PAR) gene expression and invasion/intravasation via AP-1/p-c-Jun in colorectal cancer. *Molecular Cancer Research* **5**, 485-496

Lewin B (2000) *Genes VII*, Oxford University Press, Oxford, 990 pp

Li L, He S, Sun JM, Davie JR (2004) Gene regulation by Sp1 and Sp3. *Biochemistry and Cell Biology* **82**, 460-471

Liotta LA, Steeg PS, Stetler-Stevenson WG (1991) Cancer metastasis and angiogenesis: An imbalance of positive and negative regulation. *Cell* **64**, 327-336

Llinas P, Le Du MH, Gardsvol H, Dano K, Ploug M, Gilquin B, Stura EA, Menez A (2005) Crystal structure of the human urokinase plasminogen activator receptor bound to an antagonist peptide. *The EMBO Journal* **24**, 1655-1663

Logan SK, Garabedian MJ, Campbell CE, Werb Z (1996) Synergistic transcriptional activation of the tissue inhibitor of metalloproteinases-1 promoter via functional interaction of AP-1 and Ets-1 transcription factors. *Journal of Biological Chemistry* **271**, 774-782

Loots GG, Locksley RM, Blankespoor CM, Wang ZE, Miller W, Rubin EM, Frazer KA (2000) Identification of a coordinate regulator of interleukins 4, 13, and 5 by cross-species sequence comparisons. *Science* **288**, 136-140

Lund LR, Ellis V, Ronne E, Pyke C, Dano K (1995) Transcriptional and post-

transcriptional regulation of the receptor for urokinase-type plasminogen activator by cytokines and tumour promoters in the human lung carcinoma cell line A549. *Biochemical Journal* **310 (Pt 1)**, 345-352

Maurer GD, Leupold JH, Schewe DM, Biller T, Kates RE, Hornung H-M, Lau-Werner U, Post S, Allgayer H (2007) Analysis of specific transcriptional regulators as early predictors of independent prognostic relevance in resected colorectal cancer. *Clinical Cancer Research* **13**, 1123-1132

Means AL, Farnham PJ (1990) Transcription initiation from the dihydrofolate reductase promoter is positioned by HIP1 binding at the initiation site. *Molecular and Cellular Biology* **10**, 653-661

Min HY, Semnani R, Mizukami IF, Watt K, Todd RF 3rd, Liu DY (1992) cDNA for Mo3, a monocyte activation antigen, encodes the human receptor for urokinase plasminogen activator. *Journal of Immunology* **148**, 3636-3642

Miyake H, Hara I, Yamanaka K, Gohji K, Arakawa S, Kamidono S (1999) Elevation of serum levels of urokinase-type plasminogen activator and its receptor is associated with disease progression and prognosis in patients with prostate cancer. *Prostate* **39**, 123-129

Morita Y, Hayashi Y, Wang Y, Kanamaru T, Suzuki S, Kawasaki K, Ohta K, Yamamoto M, Saitoh Y, Itoh H, Doe WF (1997) Expression of urokinase-type plasminogen activator receptor in hepatocellular carcinoma. *Hepatology* **25**, 856-861

Morita S, Sato A, Hayakawa H, Ihara H, Urano T, Takada Y, Takada A (1998) Cancer cells overexpress mRNA of urokinase-type plasminogen activator, its receptor and inhibitors in human non-small-cell lung cancer tissue: analysis by Northern blotting and in situ hybridization. *International Journal of Cancer* **78**, 286-292

Nardone J, Lee DU, Ansel KM, Rao A (2004) Bioinformatics for the 'bench biologist': how to find regulatory regions in genomic DNA. *Nature Immunology* **5**, 768-774

Nobrega MA, Ovcharenko I, Afzal V, Rubin EM (2003) Scanning human gene deserts for long-range enhancers. *Sciences* **302**, 413

Park IK, Lyu MA, Yeo SJ, Han TH, Kook YH (2000) Sp1 mediates constitutive and transforming growth factor β-inducible expression of urokinase type plasminogen activator receptor gene in human monocyte-like U937 cells. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1490**, 302-310

Okan E, Drewett V, Shaw PE, Jones P (2001) The small-GTPase Ra1A activates transcription of the urokinase plasminogen activator receptor (u-PAR) gene via an AP1-dependent mechanism. *Oncogene* **15**, 1816-1824

Ovcharenko I, Loots GG, Giardine BM, Hou M, Ma J, Hardison RC, Stubbs L, Miller W (2005) Mulan: Multiple-sequence local alignment and visualization for studying function and evolution. *Genome Research* **15**, 184-194

Pappot H, Hoyer-Hansen G, Ronne E, Hansen HH, Brunner N, Dano K, Grondahl-Hansen J (1997) Elevated plasma levels of urokinase plasminogen activator receptor in non-small cell lung cancer patients. *European Journal of Cancer* **33**, 867-872

Parry GC, Mackman N (1994) A set of inducible genes expressed by activated human monocytic and endothelial cells contain κ B-like sites that specifically bind c-Rel-p65 heterodimers. *Journal of Biological Chemistry* **269**, 20823-20825

Ploug M, Behrendt N, Lober D, Dano K (1991) Protein structure and membrane anchorage of the cellular receptor for urokinase-type plasminogen activator. *Seminars in Thrombosis and Hemostasis* **17**, 183-193

Pyke C, Kristensen P, Ralfkiaer E, Grondahl-Hansen J, Eriksen J, Blasi F, Dano K (1991) Urokinase-type plasminogen activator is expressed in stromal cells and its receptor in cancer cells at invasive foci in human colon adenocarcinomas. *American Journal of Pathology* **138**, 1059-1067

Pyke C, Graem N, Ralfkiaer E, Ronne E, Hoyer-Hansen G, Brunner N, Dano K (1993) Receptor for urokinase is present in tumor-associated macrophages in ductal breast carcinoma. *Cancer Research* **53**, 1911-1915

Ramirez-Carrozzi V, Kerppola T (2003) Asymmetric recognition of nonconsensus AP-1 sites by Fos-Jun and Jun-Jun influences transcriptional cooperativity with NFAT1. *Molecular and Cellular Biology* **23**, 1737-1749

Roldan AL, Cubellis MV, Masucci MT, Behrendt N, Lund LR, Dano K, Appella E, Blasi F (1990) Cloning and expression of the receptor for human urokinase plasminogen activator, a central molecule in cell surface, plasmin dependent proteolysis. *The EMBO Journal* **9**, 467-474

Romer J, Lund LR, Eriksen J, Pyke C, Kristensen P, Dano KAD (1994) The receptor for urokinase-type plasminogen activator is expressed by keratinocytes at the leading edge during re-epithelialization of mouse skin wounds. *Journal of Investigative Dermatology* **102**, 519-522

Schewe DM, Leupold JH, Boyd DD, Lengyel ER, Wang H, Gruetzner KU, Schildberg FW, Jauch KW, Allgayer H (2003) Tumor-specific transcription factor binding to an activator protein-2/Sp1 element of the urokinase-type plasminogen activator receptor promoter in a first large series of resected gastrointestinal cancers. *Clinical Cancer Research* **9**, 2267-2276

Schewe DM, Biller T, Maurer G, Asangani IA, Leupold JH, Lengyel ER, Post S, Allgayer H (2005) Combination analysis of activator protein-1 family members, Sp1 and an activator protein-2α-related factor binding to different regions of the urokinase receptor gene in resected colorectal cancers. *Clinical Cancer Research* **11 (24 Pt 1)**, 8538-8548

Shetty S, Kumar A, Idell S (1997) Posttranscriptional regulation of urokinase receptor mRNA: identification of a novel urokinase receptor mRNA binding protein in human mesothelioma cells. *Molecular and Cellular Biology* **17**, 1075-1083

Signori E, Bagni C, Papa S, Primerano B, Rinaldi M, Amaldi F, Fazio VM (2001) A somatic mutation in the 5'UTR of BRCA1 gene in sporadic breast cancer causes down-modulation of translation efficiency. *Oncogene* **20**, 4596-4600

Sperl S, Jacob U, Arroyo de Prada N, Sturzebecher J, Wilhelm OG, Bode W, Magdolen V, Huber R, Moroder L (2000) (4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly selective inhibitors of human urokinase. *Proceedings of the National Academy of Sciences USA* **97**, 5113-5118

Soravia E, Grebe A, de Luca P, Helin K, Suh TT, Degen JL (1995) A conserved TATA-less proximal promoter drives basal transcription from the urokinase-plasminogen activator receptor gene. *Blood* **86**, 624-635

Stoppelli MP, Tacchetti C, Cubellis M, Corti A, Hearing V, Cassani G, Appella E, Blasi F (1986) Autocrine saturation of pro-urokinase receptors on human A431 cells. *Cell* **45**, 675-684

Suh T, Nerlov K, Dano K, Degen JL (1994) The murine urokinase-type plasminogen activator receptor gene. *Journal of Biological Chemistry* **269**, 25992-25998

Suske G (1999) The Sp-family of transcription factors. *Gene* **238**, 291-300

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673-4680

Ureta-Vidal A, Ettwiller L, Birney E (2003) Comparative genomics: genome-wide analysis in metazoan eukaryotes. *Nature Reviews Genetics* **4**, 251-262

Venkatesh B, Yap WH (2005) Comparative genomics using fugu: a tool for the identification of conserved vertebrate *cis*-regulatory elements. *Bioessays* **27**, 100-117

Wagner SN, Atkinson MJ, Thanner S, Wagner C, Schmitt M, Wilhelm O, Rotter M, Hofler H (1995) Modulation of urokinase and urokinase receptor gene expression in human renal cell carcinoma. *American Journal of Pathology* **147**, 183-192

Wang H, Skibber J, Juarez J, Boyd, D (1994) Transcriptional activation of the urokinase receptor gene in invasive colon cancer. *International Journal of Cancer* **58**, 650-657

Wang X, Terzyan S, Tang J, Loy JA, Lin X, Zhang XC (2000a) Human plasminogen catalytic domain undergoes an unusual conformational change upon activation. *Journal of Molecular Biology* **295**, 903-914

Wang Y, Dang J, Wang H, Allgayer H, Murrell GAC, Boyd D (2000b) Identification of a novel nuclear factor-κB sequence involved in expression of urokinase-type plasminogen activator receptor. *European Journal of Biochemistry* **267**, 3248-3254

Wang H, Hicks J, Khanbolooki P, Kim SJ, Yan C, Wang Y, Boyd D (2003) Transgenic mice demonstrate novel promoter regions for tissue-specific expression of the urokinase receptor gene. *American Journal of Pathology* **163**, 453-464

Wang H, Yan C, Asangani I, Allgayer H, Boyd DD (2006) Identification of an histone H3 acetylated/K4-methylated-bound intragenic enhancer regulatory for urokinase receptor expression. *Oncogene* **26**, 2058-2070

Wray GA, Hahn MW, Abouheif E, Balhoff JP, Pizer M, Rockman MV, Romano LA (2003) The evolution of transcriptional regulation in eukaryotes. *Molecular Biology and Evolution* **20**, 1377-1419

Yue J, Sun B, Liu G, Mulder KM (2004) Requirement of TGF-β receptor-dependent activation of c-Jun N-terminal kinases (JNKs)/stress-activated protein kinases (Sapks) for TGF-β up-regulation of the urokinase-type plasminogen activator receptor. *Journal of Cellular Physiology* **199**, 284-292