Available online at www.sciencedirect.com

SCIENCE DIRECT

Gene 362 (2005) 85 – 97

GENE
SECTION
EVOLUTIONARY GENOMICS

www.elsevier.com/locate/gene

Sequence evolution, processing, and posttranslational modification of zonadhesin D domains in primates, as inferred from cDNA data

Holger Herlyn *, Hans Zischler

Institute of Anthropology, University of Mainz, Colonel-Kleinmann-Weg 2 (SB II), D-55099 Mainz, Germany

Received 27 September 2004; received in revised form 3 May 2005; accepted 2 June 2005
Available online 26 September 2005
Received by W. Makalowski

Abstract

Zonadhesin is a mammalian transmembrane sperm ligand. Precursor zonadhesin essentially consists of MAM (meprin/A5 antigen/mu receptor tyrosine phosphatase) domains, a mucin-like repeat, and D domains (homologous to von Willebrand D). Recent immunovisualization and binding assays indicate that zonadhesin D domains 1–3 bind postacrosomally to the zona pellucida. This feature has attracted considerable interest in the evolution of zonadhesin and its possible biological and biomedical implications. Previous molecular evolutionary analyses, however, were confined to cDNA sequences of only few distantly related species. Moreover, except for rabbit and pig, little is known about zonadhesin's processing. To delineate the situation in primates including humans, we analyze here the evolution of zonadhesin on the basis of D domain encoding cDNAs of about 4900 base pairs (bp) lenght from a representative primate sampling (1 Strepsirhini, 3 Cercopithecidae, 3 Platyrrhini, and human; 7 new sequences) plus GenBank data from mouse, rabbit, and pig. Site-specific (CODEML and HyPhy) analysis indicates positive evolution of zonadhesin. Moreover, moving window analysis (CRANN) points to a positive correlation of sequence evolution and sperm-competition. Significant accumulations of positively selected sites across interspecifically variable motifs (identified by PROSITE) suggest that positive selection promotes differences between species by amino acid exchanges and changes in posttranslational modification. In the case of zonadhesin D domains, positive selection might thus contribute to the species-specific binding of zonadhesin and zona pellucida. A high conservation of processing and dimerization motifs of primate zonadhesin in analogy to pig, on the other hand, illustrates that zonadhesin's backbone needs to meet basic requirements in order to retain function.

Keywords: Primates; Positive evolution; Moving window analysis; Motif search; Phylogenetic shadowing; Speciation

1. Introduction

Zonadhesin is a multidomain transmembrane sperm-ligand of mammals located in the acrosomal membrane. In pig, human, and rabbit, the unprocessed precursor essentially consists of two MAM domains, a mucin-like multiple tandem repeat, five D domains, a transmembrane segment, and a short cytoplasmic tail (Hardy and Garbers, 1995; Gao and Garbers, 1998; Lea et al., 2001; Fig. 1). Mouse precursor zonadhesin differs from this general pattern by the additional presence of one MAM domain and 20 partial D domains (Gao and Garbers, 1998). While the binding partners of zonadhesin MAM domains and mucin-like repeat are still to be determined, binding

Abbreviations: bp, base pair(s); cDNA, DNA complementary to RNA; D domains, domains D0-4, domains homologous to von Willebrand D; Δl, log-likelihood difference; EGF, epidermal growth factor; f_b, fraction of the beta distributed site classes in M8; f_0, f_1, fraction of the sites classes with $\omega \leq 1$ in M3; f_p, fraction of the positively selected site class in M3 and M8; f_s, fraction of the single site class in M0 (always=1); l, log likelihood; LRT I+II, likelihood ratio tests for rate heterogeneity and positive selection, respectively; M0, M3, M7, M8, evolutionary models; MAM domain(s), meprin/A5 antigen/mu receptor tyrosine phosphatase domain(s); $\omega = d_n/d_s$, rate ratio of non-synonymous to synonymous substitutions; ω_b, ω_s, ω_0, ω_1, ω_p, mean ω of site classes with fractions f_b, f_s, f_0, f_1, and f_p; p, significance level; p_{beta}, q_{beta}, beta distribution parameters; PCR, polymerase chain reaction; $p_{(\omega>1)}$, posterior probability that a certain codon site is under positive selection.

* Corresponding author. Tel.: +49 6131 3923179; fax: +49 6131 3923799.
E-mail address: herlyn@uni-mainz.de (H. Herlyn).

0378-1119/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2005.06.009

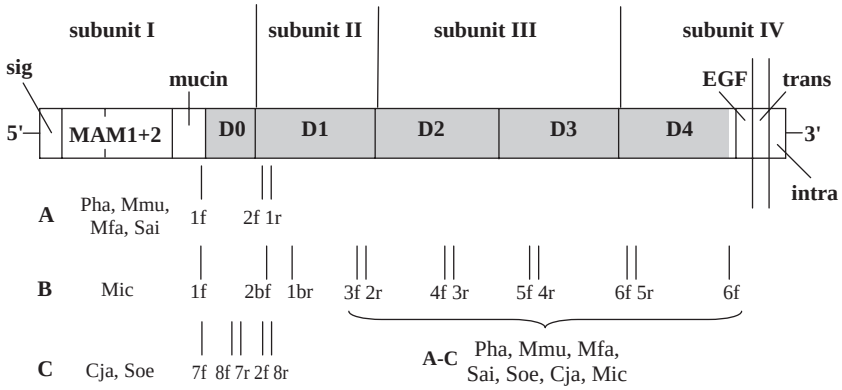

Fig. 1. Schematic domain structure and processing pattern into four subunits in pig and primate zonadhesin (modified after Bi et al., 2003). The mucin-like repeat (mucin) is arbitrarily truncated. The range of the present cDNA dataset is highlighted in gray (about 4900 bp). Below the domain structure, the approximate binding sites of PCR-primers (vertical bars) used to amplify overlapping fragments of (A) *Macaca fascicularis* (Mfa), *Macaca mulatta* (Mmu), *Papio hamadryas* (Pha), and *Saimiri sciureus* (Sai), (B) *Microcebus murinus* (Mic), and (C) *Callithrix jacchus* (Cja) and *Saguinus oedipus* (Soe) are shown. Codes of primers refer to Table 1. From primer 3f on, an identical set of primers was used for all species considered (A–C). Symbols are as follows: D0–D4, domains D0-4; EGF, EGF-like domain; intra, intracellular segment; MAM1+2, MAM domains 1 and 2; muc, mucin-like repeat (truncated); sig, signal peptide; trans, transmembrane segment. The figure is not drawn to scale.

assays and immunovisualization revealed that domains D1-3 bind post-acrosomally and species-specifically to the zona pellucida (Hardy and Garbers, 1995; Hickox et al., 2001; Lea et al., 2001; Bi et al., 2003). In analogy to the sperm ligand lysin of marine gastropods of the genera *Haliotis* and *Tegula*, the binding of zonadhesin might create a hole in the zona pellucida, thus enabling the spermatozoon to reach the egg cell membrane (see Olson et al., 2004).

Like other sex-related proteins of a broad range of taxa encompassing for instance the vagile algae *Chlamydomonas*, marine molluscs, sea urchins as well as Mammalia (reviewed in Swanson and Vacquier, 2002; see also Kamei and Glabe, 2003; Torgerson and Singh, 2003; Kouprina et al., 2004; Dorus et al., 2004), zonadhesin has been shown to be under positive selection (Swanson et al., 2003; for MAM domains see also Herlyn and Zischler, in press). However, possible correlations of amino acid replacements and posttranslational modification had only been studied for MAM domains so far (Herlyn and Zischler, in press). Moreover, previous analysis of the processing of the precursor zonadhesin focused on pig (Bi et al., 2003) and rabbit (Lea et al., 2001). Thus, still little is known about sequence evolution, processing and posttranslational modification of zonadhesin within humans and nonhuman primates. Especially, additional data on proteins involved in human reproduction appear desirable, as they might open up new perspectives in contraception and the treatment of sub- and infertility.

In pig and rabbit, posttranslational proteolytic processing of precursor zonadhesin into four subunits has been shown, whith the cleavage sites being located in the D domains (Lea et al., 2001; Bi et al., 2003). Given this central importance of zonadhesin D domains for processing, we here report the sequence evolution, processing, and posttranslational modification of zonadhesin on the basis of D domain encoding cDNAs of about 4900 bp length. As in evolutionary studies on other proteins with biomedical relevance to humans (see e.g., Sumiyama et al., 2002; Filip and Mundy, 2004; Wang and Su, 2004), the present study focuses on a representative primate sampling comprising new sequences from 1 Strepsirhini, 3 Platyrrhini, and 3 Cercopithecidae, plus the human orthologue from GenBank. To assess sequence evolution on the level of single codon sites, we use CODEML (implemented in PAML version 3.14; see Yang, 1997; Nielsen and Yang, 1998; Yang et al., 2000) and HyPhy (Kosakovsky Pond et al., 2004) estimates of the rate ratio of non-synonymous to synonymous substitutions ($d_n/d_s=\omega$) as a measure. Thereby, ω values of 1.0 indicate neutral evolution, while $\omega<1.0$ indicates negative or purifying selection, and $\omega>1.0$ points at positive or adaptive evolution. To compare sequence evolution between species, moving window analysis (CRANN; Creevey and McInerney, 2002) is performed, using the primate ancestor sequence generated by BASEML (PAML package) as common reference. Subsequent predictions of processing and posttranslational modification are based on motif search carried out per eye and using PROSITE (Hofmann et al., 1999). The functional relevance of indicated motifs is then rated, based on three criteria: 1) High phylogenetic conservation of motifs (principle of "phylogenetic shadowing"; see e.g., Boffelli et al., 2003), 2) convergent evolution of motifs, and 3) overlaps of motifs with positively selected sites. Like in the case of semenogelin II (*SEM2*; Dorus et al., 2004), zonadhesin data suggest a positive correlation of sperm-competition and positive selection. Moreover, there is evidence that positive evolution on the level of single amino acid sites promotes changes in posttranslational modification. On the other hand, processing of the primate precursor zonadhesin is apparently analogous to pig (Bi et al., 2003), thus showing high conservation.