Gene 554 (2015) 75–80

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

An evolutionarily conserved non-coding element in casein locus acts as transcriptional repressor

Suneesh Kaimala *[,1], Satish Kumar *[,2]

Centre for Cellular and Molecular Biology, Council of Scientific and Industrial Research (CSIR), Hyderabad, India

ARTICLE INFO

Article history:
Received 7 June 2014
Received in revised form 7 October 2014
Accepted 10 October 2014
Available online 13 October 2014

Keywords:
ECR
Odam
Mammary gland

ABSTRACT

In mammals, the casein locus consists of stretches of non-coding DNA, the functions of most of which are unknown. These regions are believed to harbour elements responsible for spatio-temporally regulated expression of genes in this locus and so far, only a few such elements have been identified. In this study, we report a novel regulatory element in the casein locus. Comparative analysis of genomic DNA sequences of casein loci from different mammals identified a 147 bp long evolutionarily conserved region (ECR) upstream of *Odam*, a gene in this locus. The ECR was found in close proximity of *Odam* gene in all the mammals examined. In-silico analysis predicted the ECR as a potential regulatory element. Functional analysis in different cell lines identified it as a unidirectional repressor element. From our findings we speculate that the ECR may be involved in the repression of the *Odam* expression in the mammary gland during lactation.

1. Introduction

Despite intensive studies, our understanding of the human genome is deficient. Especially with regard to the microRNAs, alternately spliced transcripts and gene regulatory regions, our comprehension of the human genome is far from being complete. Systematic identification and analysis of regulatory regions are essential for understanding the organization and variability of genes and regulatory information across cellular contexts, species and individuals (Bernstein et al., 2012). In mammals, the casein locus harbours 8 to 11 genes depending upon the species, interspersed by a huge chunk of non-coding DNA (Fig. 1a) (Rijnkels, 2002). In mice, the casein genes are *Csn1s1*, *Csn2*, *Csn1s2a*, *Csn1s2b* and *Csn3* and these genes express in mammary gland from mid-pregnancy until weaning of the offspring (Rijnkels, 2002). The locus also harbours other genes namely, *Odam* (odontogenic ameloblast associated) and *FDCSP* (follicular dendritic cell secreted protein) (Rijnkels, 2002) which express in the teeth associated epithelium and activated T-cells (Marshall et al., 2002) respectively. The presence of genes those encode proteins like Histatins and Statherin, which express in salivary glands has also been reported in the human and cow casein loci (Rijnkels, 2002).

The casein genes show spatio-temporally regulated expression (Rosen et al., 1999). The other genes in casein locus are sparsely expressed in the mammary gland. The functions of long stretches of non-coding DNA sequences flanking many of the genes in this locus are least understood. Yet, a few regulatory regions have been identified in these regions. In human, an enhancer of *CSN2* (β-casein enhancer or BCE) has been identified in the non-coding region between the *CSN2* and *CSN1S2A* (Myers et al., 1998). In rabbit, an enhancer of *Csn1s1* was identified about 3.1 kb upstream of its transcription start site (Pantano et al., 2002).

The expression of genes in the casein locus is regulated by hormones like prolactin (PRL), insulin and hydroxycortisone (Eisenstein and Rosen, 1988). The chromatin at this locus is in open conformation during pubertal development of mammary gland (Rijnkels et al., 2013). However, the proximal promoter regions of all the casein genes are in closed conformation until pregnancy. The chromatin conformation at the proximal promoter regions switches to an open state during pregnancy and reverts to closed state towards involution (Rijnkels et al., 2013). Nevertheless, the differential expression of individual genes in casein locus during lactation suggests the independent regulation of each of them.

Odam, a gene in the casein locus, is expressed in the tooth associated cells namely ameloblasts, odontoblasts and osteoblasts (Lee et al., 2012a). During tooth development, this protein is involved in matrix protein production (Lee et al., 2012b). In ameloblasts, nuclear Odam has been reported as a regulator of the expression of matrix metalloproteinase-20 (MMP-20), a gene that is responsible for the mineralization of enamel (Lee et al., 2010a). Also, *Odam* is expressed in malignant ameloblasts and other types of human neoplastic cells, including those of gastric, lung, and breast

Abbreviations: ECR, evolutionary conserved region; Odam, odontogenic ameloblast associated protein; PIP, percentage identity plot; RLU, relative light units; LCR, locus control region; FDCSP, follicular dendritic cell secreted protein; PRL, prolactin.

* Corresponding author at: Transgenic and Gene Knock-out Laboratory, Centre for Cellular and Molecular Biology, Council of Scientific and Industrial Research (CSIR), Hyderabad, India.

E-mail addresses: suneesh@uaeu.ac.ae (S. Kaimala), satishk@ccmb.res.in, satishk@niab.org.in (S. Kumar).

[1] Present address: Department of Microbiology and Immunology, College of Medicine and Health Sciences, UAE University, PO Box: 17666, Al-Ain, United Arab Emirates.

[2] Present address: National Institute of Animal Biotechnology, Hyderabad, India.

http://dx.doi.org/10.1016/j.gene.2014.10.026
0378-1119/

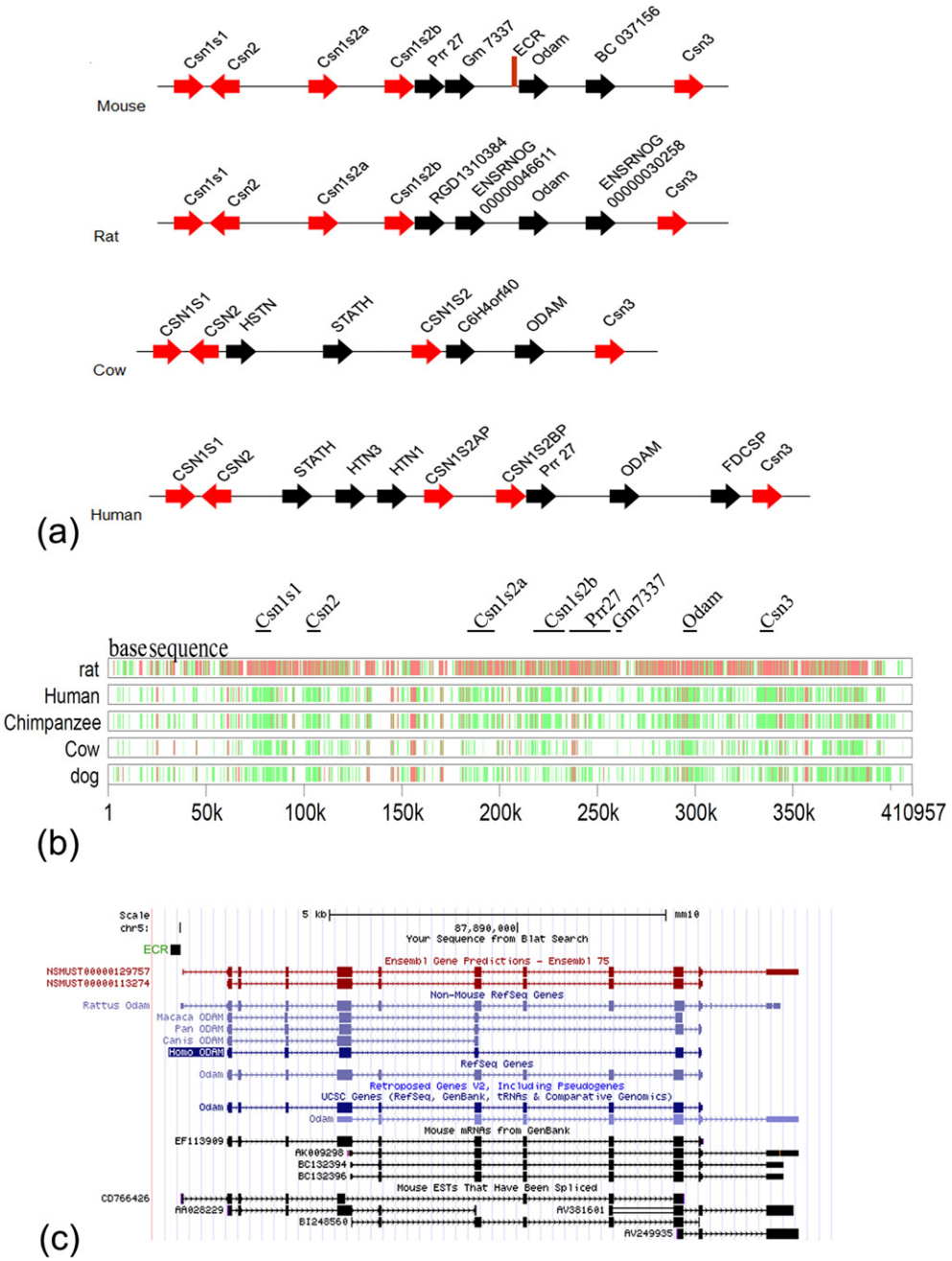

Fig. 1. Comparative analysis of the casein loci in different mammalian species. (a) Graphical representation of casein loci in mouse, rat, cow and human showing the organization of the locus in these mammals (Not drawn to scale). The location of the ECR has been indicated on the mouse casein locus. The genes in the locus are indicated as arrows and the direction of the arrow suggests the direction of transcription of the gene. Casein genes in the loci are coloured in red. ECR refers to evolutionarily conserved region. (b) The percentage identity plot (PIP) obtained by the sequence alignment of casein loci from different mammals is shown. The nucleotide sequence of the mouse casein locus that was used as the reference sequence for the comparative analysis has been indicated in the figure as the 'base sequence'. The regions indicated in red represent evolutionarily conserved regions. The approximate locations of the genes in the locus are marked above the PIP. (c) UCSC genome browser gateway for mouse showing the ECR immediately upstream of the alternate splicing form of *Odam*.

origins but the precise function of *Odam* in these cancers is not known (Kestler et al., 2011). Possibly, it is regulating the MMPs involved in the cancer metastasis.

Recently, a study of the structures of exons and introns of secretory calcium binding phosphoproteins (SCPP) like Odam, FDCSP and caseins indicated that, evolutionarily *Odam* is a more primitive gene than caseins and the caseins might have evolved from *Odam* (Kawasaki et al., 2011). In contrast to the casein genes, which are present only in mammalian vertebrates, *Odam* is present in the genome of all the mammalian vertebrates as well as the non-mammalian vertebrates such as zebrafish, platyfish, spotted gar and *Xenopus* (Ensembl orthologue search results). *Odam* is sparsely expressed in the mammary gland despite its location within the casein locus (Rijnkels et al., 2003) and little is known about its transcriptional regulation. Here we report the identification of a 147 bp long evolutionarily conserved region (ECR) upstream of *Odam* in casein locus that may be responsible for maintaining the expression of *Odam* gene in a repressed state in the lactating mammary gland.

2. Results

2.1. Identification of a conserved genomic region in the casein locus

The comparative analysis of nucleotide sequences of casein loci from different species was carried out using MultiPipMaker. The percentage