Title

TINCR is not a non-coding RNA but encodes a protein component of cornified epidermal keratinocytes

Journal

Experimental Dermatology, 29(4)

ISSN

0906-6705

Authors

Eckhart, Leopold
Lachner, Julia
Tschachler, Erwin
et al.

Publication Date

2020-04-01

DOI

10.1111/exd.14083

Peer reviewed

TINCR is not a noncoding RNA but encodes a protein component of cornified epidermal keratinocytes

Leopold Eckhart[1]*, Julia Lachner[1], Erwin Tschachler[1], Robert H. Rice[2]

1. Department of Dermatology, Medical University of Vienna, Vienna, Austria; 2. Department of Environmental Toxicology, University of California, Davis, CA, USA.

*Corresponding author: Leopold Eckhart, Department of Dermatology, Medical University of Vienna, Lazarettgasse 14, 1090 Vienna, Austria, Tel. +43-1-40400-73735; E-mail: leopold.eckhart@meduniwien.ac.at

Key words: cornification, epidermis, differentiation, ubiquitin, evolution

Abstract

Long noncoding RNAs have been implicated in the regulation of a plethora of biological processes, yet it has been challenging to verify that they are truly not coding for proteins. Terminal differentiation-induced noncoding RNA (TINCR) is a 3.7-kilobase mRNA that is highly abundant in epidermal keratinocytes prior to cornification. Here, we report the presence of an evolutionarily conserved open reading frame in *TINCR* and the identification of peptides derived from this open reading frame in the proteome of human stratum corneum. Our results demonstrate that TINCR is a protein-coding RNA and suggest that the TINCR-encoded protein is involved in keratinocyte cornification.

Short title: TINCR encodes a protein

Background

Long noncoding RNAs (lncRNAs) are RNAs of at least 200 nucleotides length that are not translated into proteins. They represent a heterogeneous group of RNAs including mRNA-like intergenic transcripts (lincRNAs), antisense transcripts of protein-coding genes and others [1]. The functions of many lncRNAs are not known but some lncRNAs were shown to control nuclear architecture and transcription in the nucleus and to modulate mRNA stability and translation in the cytoplasm [1]. lncRNAs as potential

regulators of many cellular processes have sparked great interest among researchers in dermatology and several important roles of lncRNAs in skin cells have been demonstrated [2-7].

Terminal differentiation-induced noncoding RNA (TINCR) was identified in differentiating epidermal keratinocytes [8]. *TINCR* RNA contains so-called 'TINCR box' motifs, which are 25 nucleotides long and were reported to mediate the interaction with 'TINCR box' motifs in multiple cellular mRNAs. Furthermore, *TINCR* RNA reportedly binds to the staufen1 protein and subsequently stabilizes keratinocyte differentiation-associated mRNAs [8]. Depletion of TINCR and staufen1 impaired differentiation of keratinocytes, suggesting that TINCR is essential for this process. Subsequent studies revealed transcription factor signaling through MAF:MAFB as downstream targets of TINCR [9]. Additional mechanisms of action and various mechanism of regulation of TINCR in skin and other organs have been reported in recent years [4, 10,11].

TINCR was first cloned from human hippocampus in the course of the National Institutes of Health, Mammalian Gene Collection project and was originally designated “Homo sapiens placenta-specific 2 (non-protein coding), mRNA” (GenBank accession number BC036545) [12]. Alternative names such as LINC00036, NCRNA00036, and onco-lncRNA-16 supported the noncoding nature of this RNA. Automated analysis of the DNA sequence of human chromosome 19 [13] led to the identification of an open reading frame in *TINCR* that was deposited in the Uniprot database under the accession number A0A1B0GVN0. The Uniprot database was used as a reference for the proteomic analysis of human stratum corneum and peptides corresponding to TINCR were identified in cornified envelopes [14].

Questions Addressed

Here we address the question as to whether TINCR is a noncoding or a protein-coding RNA.

Experimental Design

We obtained amino acid sequences of peptides from a mass spectrometry (MS) analysis of human stratum corneum proteins that was reported in detail previously [14]. In brief, cornified envelopes were collected with adhesive discs from healthy forearm skin, eluted, incubated with SDS-dithioerythritol, and separated into a solubilized and an insoluble (envelope) fraction which were analyzed by liquid chromatography-MS/MS [14,15]. RNA was prepared from the skin of chickens and subjected to reverse-transcription polymerase chain reaction (RT-PCR) with the intron-spanning primers GgTINCRs 5’-GGATGCTCCTCTCTGCCACA-3’ and GgTINCRa 5’-CACGCTGCGTTCCATGGTCA-3’. The PCR product was

sequenced (GenBank accession number MN85754). The open reading frames of human TINCR and TINCR of other species were translated and the resulting amino acid sequences were aligned with the Multalin algorithm [16]. Amino acid sequences were subjected to Conserved Domain search at https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi [17].

Results

Human *TINCR* RNA is produced by transcription of the *TINCR* gene which is located at chromosome 19p13.3 and comprises 3 exons (Figure 1A). Peptides identified in stratum corneum [14] correspond to a protein encoded by an open reading frame (ORF) that spans two exons of *TINCR* (Fig. 1A, Suppl. Fig. S1). The corresponding translated protein has a length of 87 amino acid residues and is predicted to fold into a ubiquitin-like 3-dimensional structure (Suppl. Figure S2).

Semi-quantitative analysis of stratum corneum proteins suggests that the abundance of TINCR is in a similar range as that of established keratinocyte differentiation proteins such as Rnase7, histidase (HAL) and involucrin (IVL) (Fig. 1B). Like two cysteine-rich cornification proteins (CRIP1 and CRCT1) and two substrates of cornification (SPRR5 and IVL), TINCR was detected in the cross-linked fraction of the cornified envelope protein but not in the solubilized protein fraction (Fig. 1B) [14], suggesting that TINCR is efficiently integrated into cornified envelopes.

Amino acid sequence analysis showed that TINCR does not contain cysteine residues whereas glutamine and lysine residues are present as potential sites of transglutamination (Fig. 1C). Comparison of amino acid sequences of TINCR orthologs showed high degrees of sequence conservation among mammals and more than 50% sequence identity to a predicted TINCR protein of the chicken (Fig. 1C; Suppl. Fig. S1), suggesting that the open reading frame of *TINCR* has been conserved since the evolutionary divergence of the lineages leading to mammals and birds more than 300 million years ago [18].

Conclusions

The presence of TINCR peptides in human stratum corneum and the conservation of the TINCR open reading frame through evolution indicate that TINCR encodes a protein. It is intriguing that TINCR is predicted to fold into a ubiquitin-like domain which may facilitate specific interactions with other proteins. TINCR RNA is expressed in the skin where it is confined to differentiating keratinocytes of epidermis [8] and at lower levels in the esophagus and placenta. Proteins of these tissues should be investigated for binding partners of TINCR protein. Importantly, as deduced from the mass spectroscopic analysis TINCR

protein persists throughout cornification of keratinocytes so that it is detectable in the stratum corneum [14]. Whether TINCR acts primarily as component of cornified envelopes or whether it has other functions during cornficiation remains to be determined in future studies.

Previously, TINCR was classified as a long noncoding RNA and interactions of TINCR RNA with protein-coding mRNAs, miRNAs, and proteins were suggested to mediate effects of TINCR on keratinocyte differentiation. These interactions were not investigated in the present study and therefore binding of TINCR RNA to other molecules is not excluded. However, the fact that *TINCR* encodes a protein is definitely not compatible with its current designation as noncoding RNA. Thus, we propose that TINCR should stand for Terminal differentiation-INduced Cornification Regulator.

Acknowledgments

Author contributions: LE, ET, and RHR designed the research study. LE, JL, and RHR performed the research and analyzed the data. LE wrote the manuscript. All authors revised the paper. This work was supported by the Austrian Science Fund (FWF): P28004, P32777.

Conflict of interest

The authors have no conflict of interest.

References

1. Yao RW, Wang Y, Chen LL. Cellular functions of long noncoding RNAs. *Nat. Cell. Biol.* 2019, 21, 542.

2. Sonkoly E, Bata-Csorgo Z, Pivarcsi A, Polyanka H, Kenderessy-Szabo A, Molnar G, Szentpali K, Bari L, Megyeri K, Mandi Y, Dobozy A, Kemeny L, Szell M. Identification and characterization of a novel, psoriasis susceptibility-related noncoding RNA gene, PRINS. *J. Biol. Chem.* 2005, 280, 24159.

3. Antonini D, Mollo MR, Missero C. Research techniques made simple: Identification and characterization of long noncoding RNA in dermatological research. *J. Invest. Dermatol.* 2017, 137, e21.

4. Szlavicz E, Olah P, Szabo K, Pagani F, Bata-Csorgo Z, Kemeny L, Szell M. Analysis of psoriasis-relevant gene expression and exon usage alterations after silencing of SR-rich splicing regulators. *Exp. Dermatol.* 2018, 27, 656.

5. Ji K, Zhang J, Fan R, Yang S, Dong C. Differential expression of lncRNAs and predicted target genes in normal mouse melanocytes and B16 cells. *Exp. Dermatol.* 2018, 27, 1230.

6. Cai P, Otten AB, Cheng B, Ishii MA, Zhang W, Huang B, Qu K, Sun BK. A genome-wide long noncoding RNA CRISPRi screen identifies PRANCR as a novel regulator of epidermal homeostasis. *Genome Res.* 2019, 30, 22.

7. Li D, Kular L, Vij M, Herter EK, Li X, Wang A, Chu T, Toma MA, Zhang L, Liapi E, Mota A, Blomqvist L, Gallais Sérézal I, Rollman O, Wikstrom JD, Bienko M, Berglund D, Ståhle M, Sommar P, Jagodic M, Landén NX. Human skin long noncoding RNA WAKMAR1 regulates wound healing by enhancing keratinocyte migration. *Proc. Natl. Acad. Sci. U S A.* 2019, 116, 9443.

8. Kretz M, Siprashvili Z, Chu C, Webster DE, Zehnder A, Qu K, Lee CS, Flockhart RJ, Groff AF, Chow J, Johnston D, Kim GE, Spitale RC, Flynn RA, Zheng GX, Aiyer S, Raj A, Rinn JL, Chang HY, Khavari PA. Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature* 2013, 493, 231.

9. Lopez-Pajares V, Qu K, Zhang J, Webster DE, Barajas BC, Siprashvili Z, Zarnegar BJ, Boxer LD, Rios EJ, Tao S, Kretz M, Khavari PA. A LncRNA-MAF:MAFB transcription factor network regulates epidermal differentiation. *Dev. Cell.* 2015, 32, 693.

10. Li B, Tsoi LC, Swindell WR, Gudjonsson JE, Tejasvi T, Johnston A, Ding J, Stuart PE, Xing X, Kochkodan JJ, Voorhees JJ, Kang HM, Nair RP, Abecasis GR, Elder JT. Transcriptome analysis of psoriasis in a large case-control sample: RNA-seq provides insights into disease mechanisms. *J. Invest. Dermatol.* 2014, 134, 1828.

11. Yu S, Wang D, Shao Y, Zhang T, Xie H, Jiang X, Deng Q, Jiao Y, Yang J, Cai C, Sun L. SP1-induced lncRNA TINCR overexpression contributes to colorectal cancer progression by sponging miR-7-5p. *Aging (Albany NY)* 2019, 11, 1389.

12. Strausberg RL, Feingold EA, Grouse LH, et al. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc. Natl. Acad. Sci U S A.* 2002, 99, 16899.

13. Grimwood J, Gordon LA, Olsen A, et al. The DNA sequence and biology of human chromosome 19. *Nature* 2004, 428, 529.

14. Karim N, Phinney BS, Salemi M, Wu PW, Naeem M, Rice RH. Human stratum corneum proteomics reveals cross-linking of a broad spectrum of proteins in cornified envelopes. *Exp. Dermatol.* 2019, 28, 618.

15. Rice RH, Durbin-Johnson BP, Mann SM, Salemi M, Urayama S, Rocke DM, Phinney BS, Sundberg JP. Corneocyte proteomics: Applications to skin biology and dermatology. *Exp. Dermatol.* 2018, 27, 931.

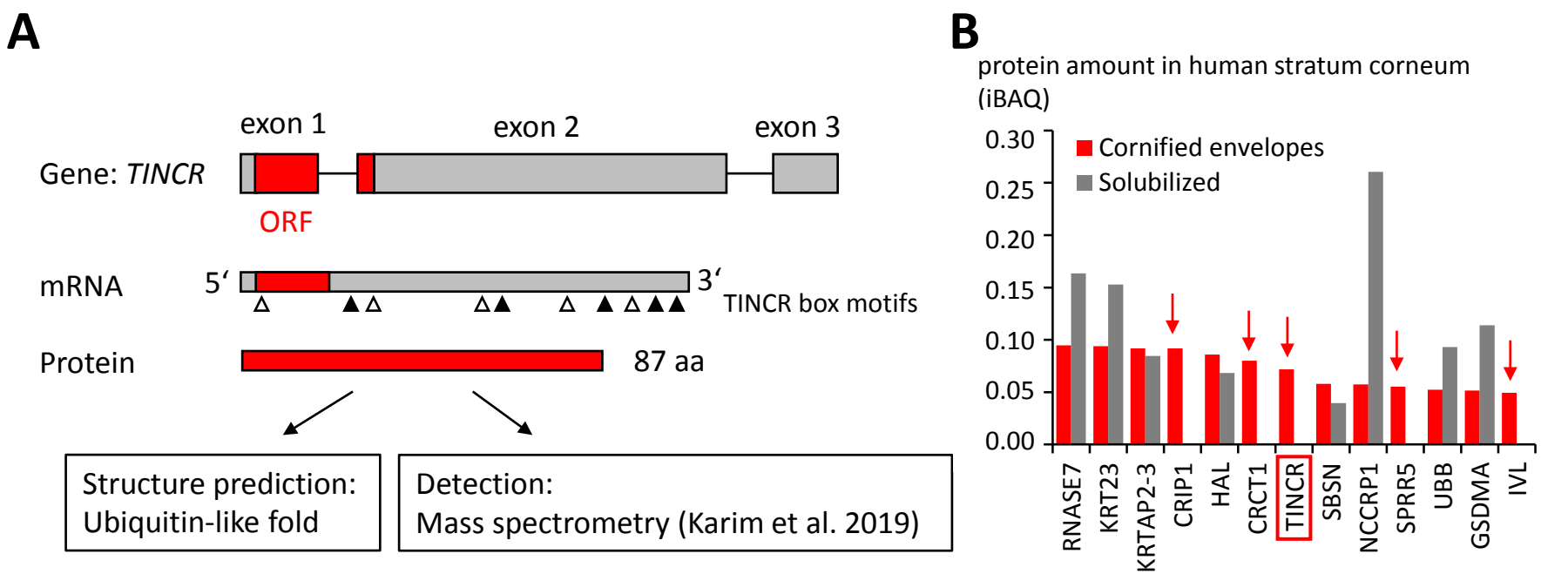
16. Corpet F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* 1988, 16, 10881.

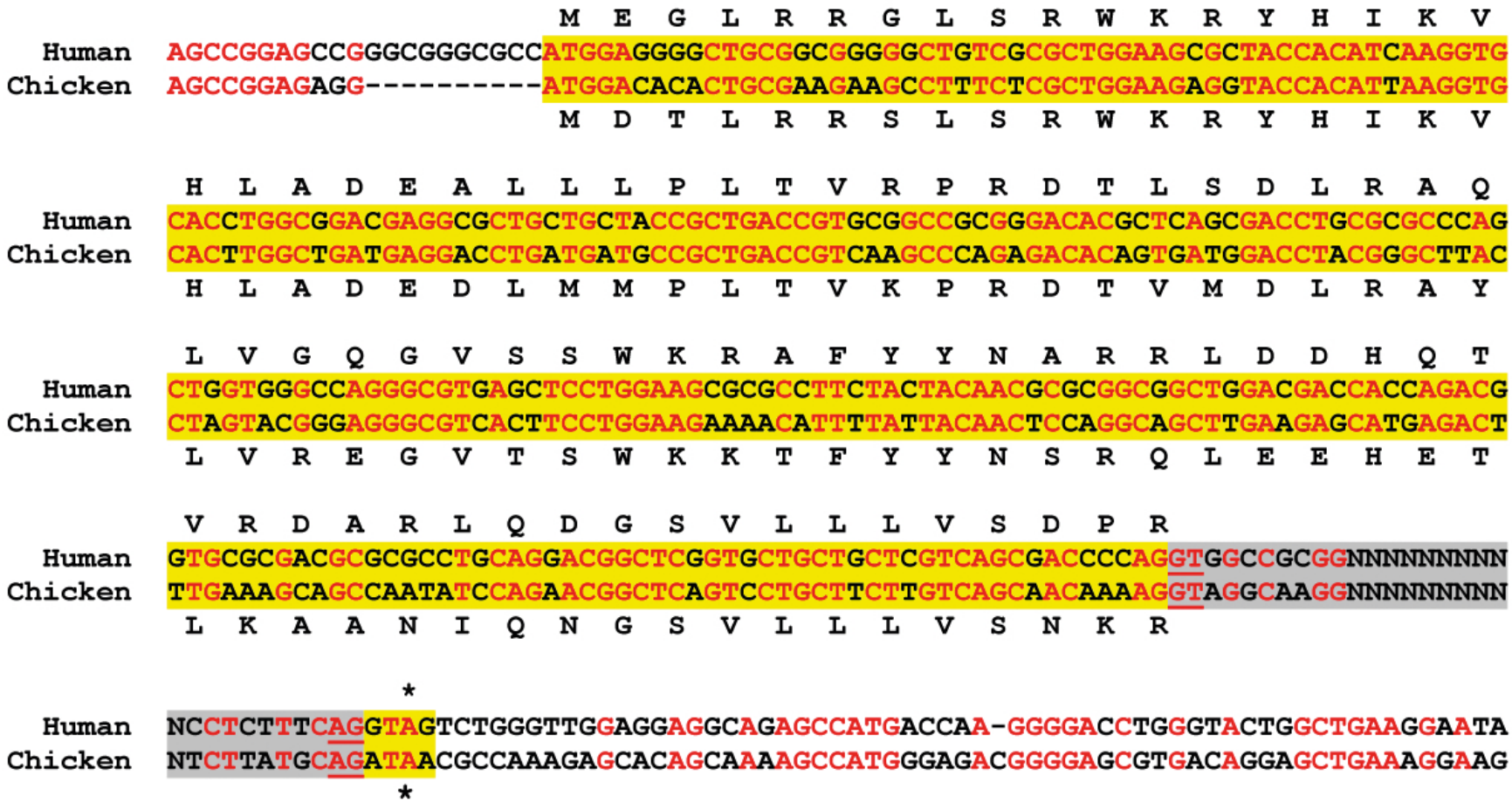
17. Marchler-Bauer A, Bryant S. CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* 2004, 32, 327.

18. Hedges SB, Marin J, Suleski M, Paymer M, Kumar S. Tree of life reveals clock-like speciation and diversification. *Mol. Biol. Evol.* 2015, 32, 835.

Figures

Figure 1. TINCR encodes an evolutionarily conserved protein. (**A**) Schematic depiction of the protein coding role of TINCR. Exons are depicted as boxes with the open reading frame (ORF) shown in red. The coding region and the relative positions of so-called "TINCR box" sequence motifs (white and black triangles) are indicated on the mRNA. (**B**) Protein amounts in cornified envelopes and in the solubilized fraction of human stratum corneum are shown in units of intensity-based absolute quantification (iBAQ). The values were obtained from supplemental tables S2 and S3 of Reference Karim et al. 2019 [14]. Proteins detected only in the cornified envelope fraction are marked by red arrows. (**C**) Amino acid sequence alignment of TINCR proteins from phylogenetically diverse mammals and chicken. Exclusive unique peptides detected by mass spectrometry (MS) [14] are indicated by red boxes in the human TINCR sequence. Numbers (n) of peptide hits are shown above the boxes. Grey shading indicates amino acid residues different from the human counterpart. Residues identical in all TINCR orthologs are indicated by asterisks below the alignment.


A
exon 1
exon 2
exon 3
Gene: TINCR
ORF
mRNA
5‘
3‘
TINCR box motifs
Protein
87 aa
Structure prediction:
Ubiquitin-like fold
Detection:
Mass spectrometry (Karim et al. 2019)
B
protein amount in human stratum corneum
(iBAQ)
0.30
0.25
0.20
0.15
0.10
0.05
0.00
Cornified envelopes
Solubilized
RNASE7
KRT23
KRTAP2-3
CRIP1
HAL
CRCT1
TINCR
SBSN
NCCRP1
SPRR5
UBB
GSDMA
IVL
C
Exclusive unique peptides (MS):
n=2
n=3
n=3
Human MEGLRRGLSRWKRYHIKVHLADEALLLPLTVRPRDTLSDLRAQLVGQGVSSWKRAFYYNARRLDDHQTVRDARLQDGSVLLLVSDPR
Macaque MEGLRRGLSRWKRYHIKVHLADEALLLPLTVRPRDTLSDLRAQLVGQGVSSWKRAFYYNARRLDDHQTVRDARLQDGSVLLLVSDPR
Mouse MEELRRGLSRWKRYHIKVHLADEALLLPLTVRPRDTLSDLRAQLVGQGVSSWRRTFYYNSRPLPDHQTVREARLQDGSVLLLLSDTR
Cattle MEGLRRGLSRWKRYHIKVHLADEALLLPLTVRPRDTLSDLRAQLVGQGVSSWKRTFYYNARRLDDHQTVRDVRLQDGSVLLLVSDPR
Platypus METLRRSLSRWKRYHIEVHLQEEDRLLPLTVRPTDTVSDLRAQLVRQGVTSWKKTFYYNAKQLADHETVRDVNIQNGSVLLLVGDPR
Chicken MDTLRRSLSRWKRYHIKVHLADEDLMMPLTVKPRDTVMDLRAYLVREGVTSWKKTFYYNSRQLEEHETLKAANIQNGSVLLLVSNKR



Supplementary Figure S1. Nucleotide sequence alignment of human and chicken *TINCR* genes. Nucleotide sequences of the coding segments and their flanking regions within human and chicken *TINCR* were aligned with the Multalin algorithm. Coding sequence is indicated by yellow shading. Amino acid sequences of the encoded proteins are shown above and below the nucleotide seqeuences. An asterisk indicates the end of a protein. The sequence of the intron between the two coding exons is shaded grey. Only the first and the last 10 nucleotides of the intron are shown whereas the main portion of the intron is replaced by N's. Intronic splicing signals GT and AG are underlined. Red fonts indicate identical nucleotides both sequences. Dashes were introduced to optimize the alignment in the non-coding regions. GenBank accession numbers of the nucleotide sequences shown in the figure: Human TINCR (partial sequence), NC_000019.10, nucleotides 5567945-5567655 and 5562221-5562149 (reverse complement); Chicken TINCR (partial sequence), NC_006115.5, nucleotides 4549790-4550070 and 4556982-4557055.

List of domain hits

Name	Accession	Description	Interval	E-value
Ubl_Ddi1_like	cd01796	ubiquitin-like (Ubl) domain found in the eukaryotic Ddi1 family	16-83	4.76e-07

ubiquitin-like (Ubl) domain found in the eukaryotic Ddi1 family; The eukaryotic Ddi1 family, including yeast aspartyl protease DNA-damage inducible 1 (Ddi1) and Ddi1-like proteins from vertebrates and other eukaryotes, has been characterized by containing an N-terminal ubiquitin-like (Ubl) domain and a conserved retroviral aspartyl-protease-like domain (RVP) that is important in cell-cycle control. Yeast Ddi1 and many family members also contain a C-terminal ubiquitin-association (UBA) domain, however, Ddi1-like proteins from all vertebrates lack the UBA domain. Ddi1, also termed v-SNARE-master 1 (Vsm1), is an ubiquitin receptor involved in the cell cycle and late secretory pathway in Saccharomyces cerevisiae. It functions as an UBA-Ubl shuttle protein that is required for the proteasome to enable ubiquitin-dependent degradation of its ligands. For instance, Ddi1 plays an essential role in the final stages of proteasomal degradation of Ho endonuclease and of its cognate FBP, Ufo1. Moreover, Ddi1 and its associated protein Rad23p play a cooperative role as negative regulators in yeast PHO pathway. This family also includes mammalian regulatory solute carrier protein family 1 member 1 (RSC1A1), also termed transporter regulator RS1 (RS1), which mediates transcriptional and post-transcriptional regulation of Na(+)-D-glucose cotransporter SGLT1. Ddi1-like proteins play a significant role in cell cycle control, growth control, and trafficking in yeast and may play a crucial role in embryogenesis in higher eukaryotes.

Pssm-ID: 340494 **Cd Length:** 73 **Bit Score:** 42.54 **E-value:** 4.76e-07

```
                 10        20        30        40        50        60        70
          ....*....|....*....|....*....|....*....|....*....|....*....|....*....|.
Query_22360 16 IKVHLADEALLLPLTVRPRDTLSDLRAqLVGQ--GVSSWKRAFYYNARRL-DDHQTVRDARLQDGSVLLLV 83
Cdd:cd01796  3 LTVTTEDDDRLFSLEVSPDMTLEDLKA-LCEAetGIPAAEQVLLHNGQPLtDDKKTLEALGLKDGDLLLLR 72
```

UBQ	smart00213	Ubiquitin homologues	28-83	7.61e-06

Ubiquitin homologues; Ubiquitin-mediated proteolysis is involved in the regulated turnover of proteins required for controlling cell cycle progression

Pssm-ID: 214563 **Cd Length:** 72 **Bit Score:** 39.55 **E-value:** 7.61e-06

```
                    10        20        30        40        50
             ....*....|....*....|....*....|....*....|....*....|....*..
Query_22360    28 PLTVRPRDTLSDLRAQLVG-QGVSSWKRAFYYNARRLDDHQTVRDARLQDGSVLLLV 83
Cdd:smart00213 14 TLEVKPSDTVSELKEKIAElTGIPPEQQRLIYKGKVLEDDRTLADYGIQDGSTIHLV 70
```

ubiquitin	pfam00240	Ubiquitin family	28-83	1.47e-05

Ubiquitin family; This family contains a number of ubiquitin-like proteins: SUMO (smt3 homolog), Nedd8, Elongin B, Rub1, and Parkin. A number of them are thought to carry a distinctive five-residue motif termed the proteasome-interacting motif (PIM), which may have a biologically significant role in protein delivery to proteasomes and recruitment of proteasomes to transcription sites.

Pssm-ID: 333953 **Cd Length:** 71 **Bit Score:** 38.72 **E-value:** 1.47e-05

```
                   10        20        30        40        50
            ....*....|....*....|....*....|....*....|....*....|....*..
Query_22360    28 PLTVRPRDTLSDLRAQL-VGQGVSSWKRAFYYNARRLDDHQTVRDARLQDGSVLLLV 83
Cdd:pfam00240  11 TLEVKPTDTVLQLKEKIaDREGVPPDQQRLIYKGKVLEDDQTLGEYGIEDGSTIHLV 67
```

Supplementary Figure S2. TINCR protein contains a ubiquitin-like fold. The amino acid sequence of human TINCR protein was used as a query. The search for conserved domains was performed at the website (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi?) using default parameters (Search against database: CDD v3.17 - 52910 PSSMs; Expect value threshold: 0.01; Apply low-complexity filter: no; Composition based statistics adjustment: yes). The three top domain hits are shown. References: Marchler-Bauer A, Bryant SH (2004), "CD-Search: protein domain annotations on the fly.", Nucleic Acids Res.32(W)327-331.Marchler-Bauer A et al. (2011), "CDD: a Conserved Domain Database for the functional annotation of proteins.", Nucleic Acids Res.39(D)225-9. Marchler-Bauer A et al. (2015), "CDD: NCBI's conserved domain database.", Nucleic Acids Res.43(D)222-6. Marchler-Bauer A et al. (2017), "CDD/SPARCLE: functional classification of proteins via subfamily domain architectures.", Nucleic Acids Res.45(D)200-3.