1

# **Conserved Sequences and Secondary RNA Structures Confirm That H19 Functions as a Pri-miRNA and as an RNA-binding Protein Platform**

Eric J. Devor<sup>1,\*</sup>, Jillian N. De Mik<sup>2</sup>, and Brandon M. Schickling<sup>3</sup>

<sup>1</sup>Department of Obstetrics and Gynecology, University of Iowa Carver College of Medicine Iowa City, Iowa 52242, USA <sup>2</sup>Department of Pediatric Pulmonary Medicine, Vanderbilt University Medical Center Nashville, TN 37232, USA <sup>3</sup>Department of Internal Medicine, University of Iowa Carver College of Medicine Iowa City, Iowa 52242, USA

**Abstract:** The imprinted non-coding RNA H19 was discovered nearly thirty years ago yet, to date, its function has not been ascertained. Recently, H19 has been shown to be the primary transcript of the microRNA *miR-675*. Both H19 and *miR-675* are known to be highly expressed in early fetal development and in an array of solid tumors in humans. We examined H19 sequence and secondary RNA structure in the context of therian evolution to begin to understand the relationship between H19, *miR-675*, development and oncogenesis. Both H19 sequence and, to a somewhat greater extent, secondary RNA structure is conserved, particularly among eutherians. The pattern of secondary RNA structure conservation in which not only is *pre-miR-675* highly conserved but also a number of other, smaller hairpins and a stable fold, coupled with the presence several conserved poly-pyrimidine tracts that are putative binding sites for the RNA-binding protein IMP1, suggests that H19 has two functions compatible with a non-coding RNA. The first is as *pri-miR-675* and the second is as a stable docking platform for a regulatory RNP composed of the 3' half of the H19 transcript and up to four molecules of IMP1.

Keywords: H19, miR-675, conservation, mammal.

## **INTRODUCTION**

H19 is an imprinted non-coding RNA, first identified nearly three decades ago, that is only found in mammalian genomes [1,2]. Though known to be expressed during fetal development, in particular in mesoderm- and endodermderived tissues [3,4], expression all but ceases at birth only to re-emerge in a number of cancers in humans including colorectal [5,6], hepatocellular [7], bladder [8-10], esophageal [5] as well as breast and ovarian [11,12]. We have recently found that H19 is highly expressed in uterine cancers, particularly the more aggressive serous tumors of the endometrium. In spite of these, and other, observations related to embryonic development and carcinogenesis [13], the functional role of H19 remains unknown. As a starting point in attempting to understand the role of H19 in endometrial and other cancers, we examine H19 RNA structure and processing in the context of mammalian evolution.

H19 is unique to metatherian and eutherian genomes and human H19 is located at chromosome 11p15.5. The locus is reciprocally imprinted with nearby insulin-like growth factor 2 (IGF2) wherein H19 is maternally imprinted and paternally expressed and IGF2 is paternally imprinted and maternally expressed. The primary H19 transcript is 2,666nt in length and is composed of five exons and four small introns. Cai and Cullen [14] showed that a microRNA, miR-675, is located near the 3' end of exon 1. Another seven smaller but stable hairpins, one upstream and six downstream of premiR-675, and two internal folds were proposed by Juan and colleagues [15] as part of the secondary structure of the H19 primary transcript. We show that these structures are conserved throughout the placental mammals. We also show that pre-miR-675 and the four introns are processed in the nucleus which produces a 1.3kb long mature H19 transcript detected in the cytoplasm. Within this mature H19 transcript are six stable, conserved hairpins and a conserved fold that serves to present the hairpins and downstream polypyrimidine tracts as docking sites for the RNA trafficking protein insulin-like growth factor mRNA-binding protein (IMP1). We, therefore, propose that H19 has two functions, one is as the primary miR-675 transcript (pri-miR-675) and the other is as a conserved docking platform used by IMP1 to create a stable ribonucleoprotein (RNP) for its function in posttranscriptional gene regulation in early development and, at least in humans, in cancer.

# MATERIALS AND METHODS

# H19 Genomics

The human H19 sequence NG\_016165 was used to search for orthologs in placental mammal species *via* BLAST in Ensembl (http://www.ensembl.org/Multi/blast-view). Full length H19 sequences were obtained for eleven

<sup>\*</sup>Address correspondence to this author at the Department of Obstetrics and Gynecology University of Iowa Carver College of Medicine 3234 MERF, Iowa City, Iowa 52242, USA; Tel: 319-321-7695; Fax: 319-356-3901; E-mail: eric-devor@uiowa.edu

species. Among these were three primates (orangutan, *Pongo pygmaeus*, rhesus macaque, *Macaca mulatta*, and marmoset, *Callithrix jacchus*), three rodents (mouse, *Mus musculus*, rat, *Rattus norvegicus*, and guinea pig, *Cavia porcellus*), three ungulates (cow, *Bos taurus*, sheep, *Ovis aires*, and pig, *Sus scrofa*), a carnivore (dog, *Canis familiaris*) and a cetacean (bottle-nose dolphin, *Tursiops truncatus*). Partial H19 sequences were extracted from GenBank including cat (*Felis cattus*, AF190057), horse (*Equus caballus*, NR\_027326) and wallaby (*Macropus eugenii*, AM993150).

Sequences of *miR-675* orthologs were obtained for sixteen mammal species using the same strategy with the addition of BLAST searches in GenBank TRACE Archive (http://www.ncbi.nlm.nih.gov/blast/mmtrace.shtml). All potential pre-miR-675 sequences were validated *via* BLAST in miRBase (http://www.mirbase.org/).

All sequence alignments were performed in CLUSTAL W using default settings [16].

## **RNA** Thermodynamics

Specific regions of H19 sequences were submitted as RNA sequences (U for T substitutions) for RNA folding in mFOLD (http://mfold.rna.albany.edu/?q=mfold) [17]. All foldings were performed at 37°C with linear RNA sequences and default conditions. Fold stability was assessed by the Gibbs' free energy value ( $\Delta G$ ).

## **Cell Culture and RNA Preparation**

In order to examine the H19 transcript itself, we chose to use cultured human cells with which we were familiar. A number of endometrial (uterine) cancer cell lines are maintained in our lab including Ishikawa H, Hec50co, Hec1A, KLE, RL95-2, ECC-1, AN3CA, and SK-UT-1b so these were our starting point. Total RNA was prepared from each cell line at the time of routine passaging using the miRvana RNA Isolation Kit (Ambion, Life Technologies). H19 expression level in each cell line was assayed in triplicate via the Applied Biosystems (Life Technologies) non-coding RNA qPCR expression assay Hs00262142 g1 that targets exon 5. Expression was normalized ( $\Delta$ Ct) within cell lines using 18S rRNA (Applied Biosystems, Life Technologies, Hs99999901 s1). RL95-2 cells displayed significantly higher H19 expression than did the other cell lines. Thus, we selected RL95-2 cells for the RT-PCR studies. These cells were grown and maintained at 37°C in 5% CO<sub>2</sub> using DMEM F12 media supplemented with 10% fetal bovine serum (FBS) and penstrep (p/s) antibiotic (GIBCO, Invitrogen).

# **PCR** Amplifications

Aliquots of 3 X 10<sup>6</sup> RL95-2 cells were collected and the nuclear and cytoplasmic fractions isolated using the Norgen Cytoplasmic and Nuclear RNA Purification Kit (Norgen Biotek) according to manufacturers' recommendations. Cytoplasmic and nuclear RNA fractions were evaluated on an Agilent 2100 Bioanalyzer. Nuclear RNA was accepted for further use if the amount of 18S and 28S rRNA observed was  $\leq$  10% of that in the corresponding cytoplasmic fraction in equal total RNA mass aliquots.

PCR primers were designed for different regions of the H19 RNA using PrimerQuest (Integrated DNA Technologies). Primer sequences are shown in Table **1**. RNAs were reverse transcribed with SuperScript III (Invitrogen) and amplified using the PCR conditions  $95^{\circ}C_{5:00}[95^{\circ}C_{0:30}; 60.0^{\circ}C_{0:30}; 72^{\circ}C_{0:30}]_{35}72^{\circ}C_{5:00}$ .

# RESULTS

# H19 Sequence Conservation in the Mammals

Clustal W alignment of complete H19 sequences among twelve placental mammal species plus the partial F. cattus sequence indicates that there are small, scattered patches of high sequence conservation (Supplemental File 1). Two of these patches are associated with the transcription start site (TSS) and the polyadenylation site. At the former, a 22nt sequence lying just upstream of the AGC TSS sequence is nearly perfectly conserved in all species. The function of this sequence is unknown but the level of conservation in such close proximity to the TSS suggests that it could be a binding site. Similar conservation is seen in the sequences comprising hairpin 8, the poly-adenylation signal and polyadenylation site. Overall, however, high sequence conservation is confined within orders with the lowest level of sequence similarity between the subclasses Eutheria (placental mammals) and Metatheria (marsupials) (Supplemental File 2).

All of the exon/intron boundaries are conserved as well but there are differences in exon size, particularly in exons 1 and 5, the two largest of the exons (Supplemental File **3**). Again, the overall conservation among placental mammals

Table 1.PCR Primer Sequences Used in this Study. Sequences were Chosen Using the PrimerQuest Tool (Integrated DNA Technologies) and Analyzed for Secondary Structures and Dimers Using the OligoAnalyzer Tool (Integrated DNA Technologies). Amplicon Sizes for HsaH19F0/HsaH19R0 and HsaH19F3/HsaH19R0 are Presented with and Without Introns

Primer ID	Sequence	Tm	Amplicon Size		
HsaH19F0	TGCTGCACTTTACAACCACTGCAC	60.1°C			
HsaH19R0	CTCCACAACTCCAACCAGTGCAAA	59.5°C	480(399)		
HsaH19F3	GTGACAAGCAGGACATGACATGGT	59.5°C	907(651)		
HsaH19F5	AGGCATTCATCCCGGTCACTTT	59.0°C			
HsaH19R5	TCCTGAGAGCTCATTCACT	59.5°C	805		

#### Conserved Sequences and Secondary RNA Structures Confirm

does not extend to the two marsupial species shown in Supplemental File **3**. In addition, the highly conserved sequence at the TSS is absent in the wallaby but the poly-A site sequence is conserved in both the wallaby and the opossum (Supplemental File **4**).

#### **Conservation of H19 Secondary RNA Structure**

Juan and colleagues [15] predicted that the primary H19 transcript would have eight hairpin structures and two folds. Hairpin 2 is, in fact, pre-miR-675. We designate the two folds as pre-hairpin 1 and pre-hairpin 8. Using the locations of the proposed hairpins and annotated in the thirteen species Clustal W alignment (Supplemental File 1), RNA conversions of putative H19 hairpins for 17 mammals were submitted to Mfold analyses. Results of the RNA thermodynamics analyses (Table 2) suggest that hairpin stability, measured by Gibbs' free energy ( $\Delta G$ ), is conserved among the placental mammals but not as much among the metatherians represented by the wallaby (Macropus eugenii) and the gray South American opossum (Monodelphis domestica). The one obvious exception to this is pre-miR-675 which is the largest and most stable H19 hairpin in all species. Predicted hairpins 1, 3, 4, 5, 6, 7 and 8 are shown for numerous placental mammal species in Supplemental File 5a-5g. Also shown are comparative sequences in ten placental mammal species for the pre-hairpin 8 fold that preserves the predicted secondary alignment (Supplemental File 5h).

Finally, the relationship between primary H19 sequence and predicted secondary RNA structures is shown in a series of VISTA plots [18,19] showing that regions of highest sequence similarity among placental mammal species correspond to the locations of the eight hairpins, the TSS and the poly-adenylation site (Supplemental File 6).

## miR-675 Alignments

The conservation pattern seen overall in H19 is also found when considering miR-675 on its own. Alignment of *pre-miR-675* in a total of nineteen mammalian species that includes the two metatherian species shows that opossum and wallaby are much more similar to each other than either is to the placental species (Fig. 1). Also seen in Fig. (1) is the customary "camel" pattern of nucleotide differences wherein the majority of such differences reside in the stem and the loop of the pre-miR [20, 21]. However, there are a number of nucleotide differences in the mature miRNA sequence including in the "seed" region, a pattern that is not characteristic of the vast majority of miRNAs [21]. Such seed region differences suggest the possibility that *miR-675* may have shifted mRNA target specificities in a lineage-specific manner over time.

## **RNA Processing**

The primary H19 transcript contains *pre-miR-675* and four small introns that are processed by DROSHA/DGCR8 and the normal spliceosome machinery respectively. We determined the size and composition of the mature, processed H19 RNA by selecting PCR primers that amplify informative parts of the primary transcript (Table 1). Location

						miR-6	575														
Human	TGCGG	CCCI	GGG	IC <u>TG</u>	GTGC GGA	GAGG	GCCCAC	AGTG	GAC-TT	GGTGA	CGCI	'G	-TAT	GCCC	TCACCG	CTC-A	GCCC	CTO	GGGGG	CTGG	CT
Orangut an	880	::::	:::					m	0.50			:	-:::	::::		())-3		::			::
Rhesus	нш				mm	m	mm	m		nт	:A::	:			mm	())-(				m	н
Marmoset		::::	:::				:T::::				:A::	;	:::	::::		:C:-:		:::			:G
Mouse		::::		A:::		A:::				:::-:	:A::	;	:::		:A::::	:::-:	:T::		:::T		:A
Rat		::::	::::	A:::		A:::			:	:::-G	:A::	;	::	::::	:A::::	:::-:	:T::	:::	:::T:		:A
Deer Mouse			A: ::	A:::	: C : : : : :	A:::			:::-GC	TTGTG	:A::	;	:::	::::	:A::::	:::-:		:::	:::T:		:A
Guinea Pig		::::	::A	C:::	: C : : : : :						:A::	:	:::		:A::::	:	::::	:::	:::T	:C::	:A
Gopher	:::A:			-:::	: C : : : : :	A::::	:::		::-::	:T:C:		:	:::		: <b>A</b> ::::	:::-:			:::T		:A
Horse				G:::	: C : : : : :				C::-:C	A::::	:A::	:	:::	:::1		:::-:			::: <b>A</b> :		:A
Pig	000		:::	G:::	: C : : : : -	C:::	::::C:	G:::1	::-::	::::G	:A::	:	·:G(	huu	::G:::				:::A	uu	:A
Elephant	0811	::::	:::	G:::	::::::G	A:::	1000	::::	(G:-::		:A::	:	:::	:::1	:A::::	:C:-:	:T::		000		:A
Cow	нш	:::6	::A	G: CA	-C::::-	m			0.00	::G::	:A::	: GGGACI	AC : T (	A:::	GA::::	::G		T :	:::A)	(G::	:A
Sheep		::::	:::	G: CA	-C::::-		::::-:	::::(	:-::	::G::	:A::	: GGGACI	AC:TO	A:::	GA::::	::G			::: <b>A</b> 1	[G::	:A
Dolphin		::::	:::	G:::	:C:::-		::::C:					: GTGAC	AC:TO	) 	:T::::	:C:			:::A		:A
Dog		::::	:::	G:G:	: C : : : : :			::::(	::-:C		:A::	;	:::			:::-:		:::	:::T		::
Cat			A: :	G:C:	: C : : : : :	A:A:	:::-::	:A:T	:C:-:C	:::T:	:CT:	;	:::	::::		:::-:		:::	:::T:		::
Opossum	:::AC	:::6	:::A	G:C:	:A:::::	::AA	::AA::		:GG-C:	CCGAG		:	:C:	::TI	: AT : : :	:::CT		:::		:::A	.:-
Wallaby	:::A:	:::6	:::A	G:C:	:A:::::	::AA	::AA::		::66C:	CCGAG	:A::	:	:C:	::TI	: AT : : :	:::CT	::T:			:::A	d-
-	***	***	*	*	****	*	*	*			* *	*	*	*	***	*	* *	* *	***	*	*

**Fig. (1).** Alignment of *pre-miR-675* in seventeen eutherian and two metatherian species. Nucleotide identity is indicated by colons, mature *miR-675* is underlined and the "seed," positions 2-9, is double underlined. All nineteen pre-miRs were validated as *miR-675* by BLAST in miRBase.

of the primers in the H19 transcript is shown in Fig. (2A) and the results of PCR amplifications using cultured human cell RNA are shown in Fig. (2B). Amplicons produced by HsaH19F0/HsaH19R0 (399bp w/o intron IV versus 480 w/ intron IV) and HsaH19F3/HsaH19R0 (651bp w/o introns II, III, and IVversus 907bp w/ introns II, III, and IV) in both the nuclear and cytoplasmic fractions show that the introns have been spliced out but that no additional processing of the RNA such as excision of any of the smaller hairpins has occurred. The presence of an HsaH19F5/HsaH19R5 amplicon in the nuclear fraction but not in the cytoplasmic fraction indicates that miR-675 is still part of the transcript in the nucleus but has been processed prior to transport from the nucleus. These results suggest that the mature H19 RNA transcript detected in the cytoplasm begins just at the 3' side of miR-675 following DROSHA processing and intron splicing and retains the six stable hairpins predicted to lie between that point and the poly-A site as well as the sequences involved in the pre-hairpin 8 fold.

## DISCUSSION

We have shown that the H19 non-coding RNA is well conserved in placental mammal species with respect to sequence but more so with respect to RNA secondary structure. This conservation does not fully extend to marsupial mammals as represented by the opossum and the wallaby. An overall impression from H19 comparative genomics is that selection may have been more stringent among the eutherians than among the metatherians though substantially more metatherian data should be compiled to validate this view. To date, no evidence has been found to suggest that the H19 locus is present at all in the third mammalian reproductive lineage- the monotremes as represented by the platypus. Unfortunately, only three prototherian genera accounting for but five species remain with which to verify the absence of H19 in that reproductive lineage. Viewing *miR*-675, whose precursor lies near the 3' end of Exon 1 of H19, separately does not help since the absence of H19 means the absence of *miR*-675 as well. However, comparative *miR*-675 data, which are a bit more extensive at least among the eutherians, does tend to support a conclusion that H19 may have been under somewhat more intense purifying selection after the metatherian/eutherian divergence regardless of whether the prototherians had it. Further, alterations in the mature *miR*-675, especially in the seed region, hint at lineage-specific changes in mRNA targeting within the eutherians that would imply that *miR*-675 function is at least somewhet different in gau artiodeatule then in reimeter.

what different in, say, artiodactyls than in primates. Again, these data do not shed any light on the actual function of H19 or miR-675 but they do suggest a function that is fundamental in fetal development and that is later co-opted in some human cancers. This latter suggestion must be view with some caution since there is abundant evidence of increased H19 expression in human solid tumors but none, so far as we know, in tumors arising in non-human mammals.

We believe that the H19 transcript is processed in the nucleus such that *pre-miR-675*, cleaved by DROSHA, and an H19 fragment 3' of the pre-miR and roughly 1.3kb long following intron excision, are separately transported out of the nucleus and may have different functions.

Cai and Cullen [14] first recognized that H19 was *pre-miR-675*. Smits and colleagues [2] first demonstrated that *miR-675* likely pre-dated metatherian-eutherian divergence and demonstrated that the H19-IGF2 imprinting mechanism



Fig. (2). A. A map of the H19 transcript indicating the location of the RT-PCR primers used to determine the size and composition of the processed RNA. B. A 1.4% agarose gel containing RT-PCR products using the primers shown in 2A and Table 1. The marker used is the New England Biolabs 100bp ladder and the 18S rRNA primers are Ambion (Life Technologies) Universal 18S primers.

Table 2.Gibbs' Free Energy (ΔG, kcal) for the Eight Putative Hairpin Structures Present in the H19 Transcript of Nineteen Placental and Two Marsupial Mammal Species. A Number of Species are Represented Only by Partial Sequences in Gen-Bank (nd). In the Marsupial Species, Hairpin Structures Could Not be Formed Either in or Near the Equivalent Region of the RNA Transcript in Several Instances (nh)

Species Hairpin	1	2	3	4	5	6	7	8
Human	-23.1	-52.9	-18.7	-22.3	-7.4	-15.2	-9.9	-19.9
Orangutan	-23.1	-52.9	-18.7	-20.5	-7.4	-18.2	-11.2	-19.9
Rhesus	-23.1	-53.3	-18.7	-20.5	-7.4	-18.2	-11.2	-19.9
Marmoset	-17.7	-46.8	-18.5	-16.1	-7.1	-14.0	-7.4	-18.3
Mouse	-17.3	-56.0	-14.7	-10.8	-7.7	-13.0	-5.8	-20.3
Rat	-24.6	-57.5	-12.4	-13.8	-7.7	-15.8	-5.8	-20.3
Guinea Pig	-20.2	-48.6	-12.2	-13.4	-8.4	-13.5	-6.0	-17.4
Cow	-20.6	-42.9	-13.1	-15.6	-8.4	-6.6	-13.1	-15.5
Sheep	-26.7	-42.9	-12.9	-17.4	-8.4	-6.6	-13.1	-15.5
Horse	nd	-60.0	-22.0	-17.0	-7.7	-10.5	-6.0	-19.6
Dolphin	-29.7	-56.0	-14.0	-19.7	-7.1	-8.9	-9.9	-14.2
Pig	-26.2	-55.2	-13.9	-16.9	-7.7	-10.9	-9.8	-13.4
Dog	-24.2	-59.0	-17.2	-15.5	-7.7	-10.5	-4.9	-19.9
Cat	-23.5	-34.6	-19.2	-15.5	-7.7	-11.9	-4.6	-18.0
Gopher	-21.3	-45.9	-15.0	nd	nd	nd	nd	nd
Lynx	-8.5	-37.3	-15.1	nd	nd	nd	nd	nd
Elephant	-17.3	-57.3	-13.2	nd	nd	nd	nd	nd
Deer Mouse	nd	-47.2 nd	nd	nd	nd	nd	nd	
Wallaby	-16.0	-56.8	-8.2	-4.3 nh	-4.3	nh	-21.6	
Opossum	nd	-51.7	nh	-8.8	nh	-9.0	nh	-13.3

was established at or near the origin of H19 itself. To date, only two *miR*-675 mRNA targets have been validated but both targets, the tumor suppressor retinoblastoma (RB) [22] and the cartilage matrix component COL2A1 [23], are far older than the miRNA that targets them. Though the role of RB in a wide range of human tumors is well known, its role in normal development is only now being uncovered [24] and its antiquity appreciated [25]. Similarly, COL2A1 became the most abundant and functionally important cartilage matrix protein more than 500 million years ago [26,27]. Thus, regulation of both RB and COL2A1 expression by *miR*-675 is a comparatively recent event whose significance is not yet understood.

Unfortunately, even less is understood about the *miR-675* parent H19 RNA. One study [28] has shown that H19 RNA binds the RNA trafficking protein IMP1 (IGF-II mRNA-binding protein 1; IGF2BP1) in a 1:4 stoichiometry. The putative IMP1 recognition sites are four pyrimidine-rich regions near four conserved hairpins in exons 4 and 5 (hairpins 5, 6, 7 and 8) which are part of the processed RNA that is transported to the cytoplasm [28 and Supplemental File **2**]. The poly-pyrimidine tracts downstream from hairpins 6 and

7 are more extensive than those near hairpins 5 and 8. IMP1 molecules bind to H19 in a manner designed to enhance IMP1-H19 RNP stability [29]. There is a fast monomolecular binding step that creates a low stability intermediate and a fast dimerization step that creates the final stable RNP composed of the RNA dock and IMP1 dimers. The presence of four poly-pyrimidine tracts and the 1:4 H19:IMP1 stoichiometry shown by Runge and colleagues [28] indicates that their *in vitro* experiments detected fast step binding.

IMP1 is a member of the ancient, conserved VICKZ family of RNA binding proteins and shows strong homology (>95%) with the mouse coding region determinant-binding protein (CRD-BP) and the chicken zipcode-binding protein (ZBP1). IMP1 has been shown to localize mRNAs to the leading edge of proliferating cells [30] and ectopic IMP1 expression in transgenic mice results in cells with increased proliferative capacity [31]. Consistent with this, it has recently been shown that H19 induction promotes trophoblast lineage commitment in mouse embryonic stem cells [32]. Further, IMP1 is implicated in posttranscriptional regulation of several mRNAs involved in cell adhesion and increased invasive capacity [31,33,34]. It is possible that H19 plays a role in IMP1 function by providing an evolutionarily conserved stable platform for IMP1 binding in a manner necessary for its regulatory function.

The alignments and RNA folding/thermodynamics data presented here suggest that the H19 non-coding RNA has persisted in therians under moderate purifying selection maintaining secondary RNA structure for at least 180 million years [35]. We propose that H19 has two functions compatible with a non-coding RNA (Supplemental File 7). The first function is as pri-miR-675, a microRNA that definitely targets two mRNAs involved in development that are much older than itself. In the vast majority of miRNAs, the primary transcript, whether intergenic or intronic, serves to present the miRNA hairpin structure for DROSHA processing and is subsequently reprocessed. In the case of H19 part of the primary transcript is not reprocessed but, rather, has acquired a second function. That second function, involving the 3' half of the RNA transcript left over after intron and miR-675 processing, is to provide a stable platform for the RNA-binding protein IMP1 to carry out its regulatory function both in development and carcinogenesis. Importantly, these two roles require that H19 maintains the conserved RNA secondary structures shown here since it is these secondary RNA structures that guarantee proper miRNA processing on the one hand and IMP1 binding sites on the other.

# **CONFLICT OF INTEREST**

The authors confirm that this article content has no conflicts of interest.

### ACKNOWLEDGEMENTS

None declared.

### SUPPLEMENTARY MATERIAL

Supplementary material is available on the publisher's web site along with the published article.

# REFERENCES

- Pachnis V, Belayew A, Tilghman S. Genes linked to alphafetoprotein under the control of the murine raf and Rif genes. Proc Natl Acad Sci USA 1984; 81: 5523-7.
- [2] Smits G, Mungall AJ, Griffiths-Jones S, *et al.* Conservation of the H19 noncoding RNA and H19-IGF2 imprinting mechanism in therians. Nat Genet 2008; 40: 971-6.
- [3] Poirier F, Chan CT, Timmons PM, Robertson EJ, Evans MJ, Rigby PW. The murine H19 gene is activated during embryonic stem cell differentiation in vitro and at the time of implantation in the developing embryo. Development 1991; 113: 1105-14.
- [4] Gabory A, Jammes H, Dandolo L. The H19 locus: Role of an imprinted non-coding RNA in growth and development. Bioessays 2010; 32: 473-80.
- [5] Hibi K, Nakamura H, Hirai A, et al. Loss of H19 imprinting in esophageal cancer. Cancer Res 1996; 56: 480-2.
- [6] Cui H, Onyango P, Brandenburg S, Wu Y, Hsieh C-L, Feinberg AP. Loss of imprinting in colorectal cancer linked to hypomethylation of H19 and IGF2. Cancer Res 2002; 62: 6442-6.
- [7] Ariel I, Miao H-Q, Schneider T, et al. Imprinted H19 oncofetal RNA is a candidate tumour marker for hepatocellular carcinoma. J Clin Pathol 1998; 51: 21-5.
- [8] Elkin M, Shevelev A, Schulze E, et al. The expression of the imprinted H19 and IGF-2 genes in human bladder carcinoma. FEBS Lett 1995; 374: 57-61.

- [9] Ariel I, Sughayer M, Fellig Y, et al. The imprinted H19 gene is a marker of early recurrence in human bladder carcinoma. Mol Pathol 2000; 53: 320- 3.
- [10] Byun H-M, Wong H-L, Aaron EA, Wolff EM, Ling G, Yang AS. Examination of IGF2 and H19 loss of imprinting in bladder cancer. Cancer Res 2007; 67: 10753- 8.
- [11] Lottin S, Adriaenssens E, Dupressoir T, et al. Overexpression of an ectopic H19 gene enhances the tumorigenic properties of breast cancer cells. Carcinogenesis 2002; 23: 1885-95.
- [12] Tanos V, Prus D, Ayesh S, et al. Expression of the imprinted H19 oncofetal RNA in epithelial ovarian cancer. Eur J Obstet Gynecol Reprod Biol 1999; 85: 7-11.
- [13] Matouk IJ, Mezan S, Mizrahi A, et al. The oncofetal H19 RNA connection: Hypoxia, p53 and cancer. Biochim Biophys Acta 2010; 1803: 443-51.
- [14] Cai X, Cullen BR. The imprinted H19 noncoding RNA is a primary microRNA precursor. RNA 2007; 13: 313-6.
- [15] Juan V, Crain C, Wilson C. Evidence for evolutionarily conserved secondary structure in the H19 tumor suppressor RNA. Nucleic Acids Res 2000; 28: 1221-7.
- [16] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994; 22: 4673- 80.
- [17] Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 2003; 31: 3406-15.
- [18] Mayor C, Brudno M, Schwartz JR, et al. VISTA: visualizing global DNA sequence alignments of arbitrary length. Bioinformatics 2000; 16: 1046-7.
- [19] Frazer KA, Pachter L, Poliakov A, Rubin Em, Dubchak I. VISTA: computational tools for comparative genomics. Nucleic Acids Res 2004; 32: W273-9.
- [20] Berezikov E, Guryev V, van de Belt J, Weinholds E, Plasterk RH, Cuppen E. Phylogenetic shadowing and computational identification of human microRNA genes. Cell 2005; 120: 21-4.
- [21] Devor EJ, Peek AS. miRNA profile of a Triassic common mammalian ancestor and pre-miRNA evolution in the three mammalian reproductive lineages. Open Genomics J 2008; 1: 22-32.
- [22] Tsang WP, Ng EKO, Ng SSM, et al. Oncofetal protein H19derived miR-675 regulates tumor suppressor RB in human colorectal cancer. Carcinogenesis 2010; 31: 350-8.
- [23] Dudek KA, Lafont JE, Martinez-Sanchez A, Murphy CL. Type II collagen expression is regulated by tissue-specific miR-675 in human articular chondrocytes. J Biol Chem 2010; 285: 24381-7.
- [24] Chinnam M, Goodrich DW. RB1, development and cancer. Curr Top Dev Biol 2011; 94: 129-69.
- [25] Cao L, Peng B, Yao L, et al. The ancient function of RB-E2F pathway: insights from its evolutionary history. Biol Direct 2010; 5: 55.
- [26] Zhang G, Miyomoto MM, Cohn MJ. Lamprey type II collagen and Sox9 reveal an ancient origin of the vertebrate collagenous skeleton. Proc Natl Acad Sci USA 2006; 103: 3180-5.
- [27] McCauley DW, Sox E, Type II collagen, and evolution of the chondrogenic neural crest. Zool Sci 2008; 25: 982-9.
- [28] Runge S, Nielsen FC, Nielsen J, et al. H19 RNA binds four molecules of insulin-like growth factor II mRNA-binding protein. J Biol Chem 2000; 275: 29562-9.
- [29] Nielsen J, Kristensen MA, Willemoës M, Nielsen FC, Christiansen J. Sequential dimerization of human zipcode-binding protein IMP1 on RNA: a cooperative mechanism providing RNP stability. Nucleic Acids Res 2004; 32: 4368-76.
- [30] Nielsen FC, Nielsen J, Christiansen J. A family of IGF2 mRNA binding proteins IMP involved in RNA trafficking. Scand J Clin Lab Invest 2001; 61(Suppl 234): 93-9.
- [31] Dimitriadis E, Trangas T, Milatos S, et al. Expression of Oncofetal RNA-binding protein CRD-BP/IMP1 predicts clinical outcome in colon cancer. Int J Cancer 2007; 121: 486-94.
- [32] Fujimori H, Mukai H, Murakami Y, Hemberger M, Hippo Y, Masutani M. The H19 induction triggers trophoblast lineage commitment in mouse ES cells. Biochem Biophys Res Commun 2013; 436: 313-8.
- [33] Vikessa J, Hansen TVO, Jonson L, et al. RNA-binding IMPs promote cell adhesion and invadapodia formation. EMBO J 2006; 25: 1456-68.

#### Conserved Sequences and Secondary RNA Structures Confirm

- [34] Köbel M, Weidensdorfer D, Reinke C, *et al.* Expression of the RNA-binding protein IMP1 correlates with poor prognosis in ovarian carcinoma. Oncogene 2007; 26: 7584-9.
- [35] Wakefield MJ, Graves JAM. Marsupials and monotremes sort genome treasures from junk. Genome Biol 2005; 6: 218.

Received: February 27, 2013

Revised: June 21, 2013

Accepted: June 30, 2013

© Devor et al.; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/ by-nc/3.0/) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.