The mammalian Kalrn locus gives rise to several novel long non-coding RNAs

Muneesh Pal

CSIR-IGIB: CSIR Institute of Genomics & Interactive Biology - South Campus

Divya Chaubey

CSIR-IGIB: CSIR Institute of Genomics & Interactive Biology - South Campus

Mohit Tanwar

CSIR-IGIB: CSIR Institute of Genomics & Interactive Biology - South Campus

Beena Pillai (✉ beenapillai@igib.res.in)

CSIR-IGIB: CSIR Institute of Genomics & Interactive Biology - South Campus <https://orcid.org/0000-0002-9302-9878>

Research Article

Keywords: Kalrn, lncRNA, durga, Neuron, Dendrite, Chromatin

Posted Date: February 15th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1187363/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Abstract

The *KALRN* gene encodes several multi-domain protein isoforms that localise to neuronal synapses, and play dynamic roles in shaping axonal outgrowth, dendrite morphology and dendritic spine re-modelling. The genomic locus is implicated in several neurodevelopmental and neuropsychiatric diseases including autism, schizophrenia and bipolar disease. Mutations in the coding regions, inherited in a classical Mendelian manner, have also been implicated in certain forms of autism and intellectual disability. At the molecular level, the protein isoforms, encoded by reported transcript isoforms, share some core domains arising from the central exons, while other domains, especially towards the C terminal may be selectively incorporated. This heterogeneity seems to confer the ability to grow and retract dendritic spines, thus making Kalirin a critical and dynamic player in dendritogenesis. We have previously shown that in the zebrafish genome, a novel brain specific non-coding RNA arising from the 5' end of the Kalirin gene, *durga* regulates neuronal morphology.

In search of the mammalian equivalent, we characterized the mammalian *KALRN* loci in detail, annotating multiple novel non-coding RNAs, including linear and circular variants, through analysis of transcriptomics data and experimental approaches. By comparing the mouse and human loci and studying the expression of the novel lncRNAs arising from the locus during differentiation of primary cortical neurons in culture, we show that certain non-coding RNAs arising from the locus show a temporal expression profile that coincides with a subset of Kalirin protein coding isoforms. In humans, mouse and zebrafish the 5' end of the *KALRN* locus gives rise to a chromatin associated lncRNA that is present in adult ovaries besides being expressed during brain development and in certain regions of the adult brain. Besides correcting some of the annotations available in public databases, we propose that this lncRNA arising from the 5' end of the *KALRN* locus is the mammalian ortholog of zebrafish lncRNA *durga*.

Introduction

The Kalirin gene is located on chromosome 3 of the human genome and gives rise to the Dbl family Guanine nucleotide exchange factor protein called KALRN. It was originally identified as an interacting partner of Huntingtin-associated protein1 and called HAPIP . It is expressed selectively in the brain and central nervous system while highly similar proteins like DUO and TRIO show expression in other tissues too . Kalirin protein consists of many protein-protein interaction domains that allow it to interact with several partners in the dendritic spine. Most of its functions are dependent on the RAC and Rho GTPase function, which results in the release of GDP from RAC and Rho and activation of GTP signalling in the cell. For instance, the Kalirin-7 isoform localizes to the dendritic spines and interacts with PDZ domain-containing proteins using a specific region in its C-terminal region and regulates dendritic morphogenesis . Kalirin isoforms are known to affect the formation of dendritic spines, dendritic arborization, axonal growth, synapse formation and plasticity ; ; .

Mutations in *Kalrn* that are deleterious to protein function are associated with several neurodevelopmental diseases like autism, intellectual disability as well as neuropsychiatric diseases like

schizophrenia ; ; ; . Intronic variants that do not change in the protein but result in altered expression have been linked to addiction and reward anticipation . *Kalrn* expression was also found to be reduced in post-mortem Alzheimer's brain samples. Overall, the disease phenotypes seen in humans agree well with the loss of expression or function of *Kalrn* protein and resultant behavioural effects seen in mouse models. An interesting complication is the disparate effects of the Rac and Rho GEF domains: mutations in the Rac GEF domain reduce RAC signalling, resulting in reduced neuronal branching, spine density and leads to decrease in cortical thickness while a mutation in the Rho GEF domain had similar effects in spite of activating Rho GTP signalling

The multi-function *Kalrn* protein exerts its diverse effects through its many isoforms. The locus consists of 60 exons in humans, which give rise to 52 transcript isoforms. The regulated expression of *Kalrn* isoforms is thought to be driven by promoter usage, although these promoters are yet to be characterized in detail. The transcript isoforms can be clustered into two major groups: One group of relatively short isoforms, mostly including the exons at the 5' end includes the *Kal7* isoform that has been studied extensively in mice and cell culture models. *Kal7* knockout also resulted in schizophrenia-like effects in adolescent mice . The *Kal7* isoform includes the RacGEF domain but not the RhoGEF domain. Restoring levels of *Kal7* can reverse the Alzheimer's like effects shown by mice with reduced *Kal7* ; . Besides the catalytic RacGEF domain that is critical to its function, it also contains a C-terminal PDZ domain which is thought to target it to post-synaptic regions and this facilitates its interactions with other proteins in this subcellular compartment. The longer isoforms which include the RhoGEF domain form a distinct group comprising *Kal9* and *Kal12*, which play important roles during neuronal development ; . Although it is clear from mutations linked to diseases and phenotypes in mouse models that the correct spatio-temporal expression of *Kalrin* isoforms is as critical as the final protein product itself, there is relatively little information on the regulatory potential of the genomic locus.

Long non-coding RNAs are RNA transcripts more than 200nt in length that do not code for functional proteins, but may exert a variety of regulatory roles ranging from miRNA sponging and clearance to providing scaffolds for assembling ribo-nucleoprotein complexes and tethers for chromatin modifiers In zebrafish, we have reported that a maternally inherited non-coding RNA, named *durga*, is the transcript expressed earliest in the developmental time frame, from the *Kalrn* locus. The non-coding RNA partially overlaps with the first exon of protein-coding *Kalrn* isoforms and is restricted in expression to an early developmental window and later, to certain regions of the adult zebrafish brain . To identify similar non-coding RNAs from the mammalian *Kalrn* locus we carefully annotated the transcripts reported in transcriptome studies. By combining data from genome and transcriptome data repositories, we built a detailed map of the non-coding transcripts arising from the mouse *Kalrn* locus, including seven linear and one circular RNA, henceforth called *Kalnc1* to *7* in the order of their position from the 5' to 3' end of the mouse *Kalrn* gene. Further, we tracked their expression during the in vitro differentiation of mouse primary cortical neurons, finding that they were down-regulated in mature neurons. We show here that these non-coding transcripts have distinct spatio-temporal expression patterns in the mammalian brain. Some isoforms coincide in expression with *Kalrn* protein coding isoforms while others have unique patterns. Lastly, we used RNA from post-mortem human tissues and a human neuronal cell line to show that the

human *Kalrn* locus also gives rise to a potential non-coding transcript arising from the 5' end of the *Kalrn* locus, which shares short stretches of sequence similarity and functional features like chromatin association, and a selective enrichment in immature neurons.

Materials And Methods

Genome and transcriptome meta-analysis:

All the analysis and visualization for the mouse genome was carried out on (GRCm38/mm10) and similarly the human genome (GRCh38/hg38) was used. The *Kalrn* genomic locus (mouse = chr16: 33,969,073-34,573,532 human = chr3: 124,033,369-124,726,325) was accessed with the tracks from Gencode v.38 (Ensembl 104) or Gencode v.M25 for mouse (Ensembl 100). Treating *Kalrn*/ENST00000682506.1 as the longest transcript, all transcripts within the limits of this were analysed further. Five transcripts that were mentioned as “processed transcripts” and two denoted “intron-retained” were observed. These transcript sequences were downloaded and used for coding potential analysis in CPAT software(). Oligonucleotide primers (Supplementary Data) specific for transcripts and pan-*Kalrn* were designed using Primer3.

Animals:

All experimental procedures were performed on Balb/c mice and approved by the Institutional Animal Ethics Committee of CSIR-Institute of Genomics and Integrative Biology (IGIB) and followed appropriate guidelines for live animal use in research. All mice used in this study were bred in the CSIR-IGIB animal house facility. They were kept at 24±2°C on a 12h light/dark cycle with ad libitum access to food and water. All institutional guidelines were followed for animal handling and experiments.

Primary Cortical neuron culture:

Adult male and female mice were kept for mating at 1:2 ratio, and confirmed by checking a vaginal plug and after a few days of mating, it was re-confirmed by measuring the weight of a female. Pregnant females were anaesthetized with thiopentone (40 mg/kg) at Embryonic day 16 -18 (E16-18). Embryos were dissected out by making a straight incision in the abdomen. Embryos were immediately kept in cold HBBS (14025 Gibco), and brains were removed. Carefully meninges and blood vessels were removed and cortices were dissected out. They were transferred into the 15ml tube containing cold HBSS. Cortices were washed three times with HBSS for 5 min at 500g and 0.25% trypsin-EDTA (25200-0560, Gibco) was added, then triturated by Pasteur Pipette and kept at 37°C for 5 min. Trypsin-EDTA was blocked by (10082-147, Gibco), and tissue was centrifuged at 500g for 5 min, then washed two times with HBSS and single cell suspension was made in complete neurobasal media (Neurobasal; 10888-022, Gibco + B27; A35828-01, Gibco + Glutamax; 35050-061, Gibco + Primocin; ant-pm-1, Invivogen). Cells were counted and seeded on the poly-D lysine (P7405, Sigma) coated (0.1mg/ml solution) plates. First half media change was done after 12hrs of cell seeding, and thereafter twice a week. Cells were collected at different

days in vitro culture (DIV); DIV0 (the day when cells were seeded), DIV3, DIV7, DIV11, DIV15, DIV21 and DIV28.

Neuro2A cell culture:

N2A cells (originally sourced from National Cell Repository) were cultured in Dulbecco's Modified Eagle's Medium (DMEM, Invitrogen) supplemented with 10% fetal bovine serum (FBS) and incubated in a humidified incubator with 5% CO₂.

SH-SY5Y differentiation:

Sh-SY5Y cells (source: European Collection of Authenticated Cell Cultures) were cultured in Dulbecco's Modified Eagle's Medium (DMEM, Invitrogen) supplemented with 10% fetal bovine serum (FBS) and incubated in a humidified incubator with 5% CO₂. For differentiation of SH-SY5Y, cells at passage number 30-33 were treated with 10µm of *trans* retinoic acid (RA) at day 1, cells were then grown for 4 days with media change at every alternative day and replaced with 10ng/ml BDNF and 10µm RA at day5 and grown for 10days. At day 10 DMEM was replaced with neurobasal media and collected at day12.

***Kalnc2* over-expression:**

The lncRNA *Kalnc2* was cloned in mammalian plasmid pCDNA3.1+ and transfected by lipofectamine 2000 (116668-019, Invitrogen) in primary cortical neurons at DIV7 and in Neuro2A cell line. After transfection, cells were collected at different time points and the expression of *Kalnc2* along with Kalrn isoforms was measured by quantitative RT-PCR.

RNA isolation and RT-PCR:

Mice were euthanized with Sodium thiopental (40 mg/kg), and brain and other body parts were removed. RNA was isolated using trizol (15596018, Ambion) from the whole brain, cortex, hippocampus, cerebellum, ovary, liver, heart and primary cortical neurons. cDNA was prepared by 1µg of total RNA, and then RT-PCR was done using SYBR Green master mix (PKG025-A Genetix). *Gapdh* was used as a normalization control in RT-PCR data analysis. The RT-PCR products were visualized by agarose (2%) gel electrophoresis.

Subcellular fractionation:

Cells were washed with 1X PBS, and then collected by 0.25% trypsin-EDTA, pellet down (100-200g for 5 mins), and counted. 175µl cytoplasmic lysis buffer (tris HCl pH 8.0 50mM, NaCl 140mM, MgCl₂ 1.5mM, NP-40 0.5%, DTT 1mM, EDTA 5mM) was added to 1X10⁶ cells, suspended and incubated at ice for 5 min, and then cell suspension was centrifuged at 300g for mins at 4°C. Supernatant was transferred in the fresh tube and it is a cytoplasmic fraction, and pellet was washed twice with 200µl of cytoplasmic lysis buffer at 300g for 2 mins at 4°C. 175µl nuclear lysis buffer (tris HCl pH 8.0 50mM, NaCl 500mM, MgCl₂ 1.5mM, NP-40 0.5%, DTT 1mM) added per 1X10⁶ cells pellet, suspended and incubated at ice for 5 min,

and then cell suspension was centrifuged at 16,400g for 2 mins at 4°C. Supernatant was collected and this constitute soluble nuclear fraction and pellet is a chromatin fraction, and then RNA isolation, cDNA preparation and RT-PCR was done for all three fractions. The RT-PCR products were visualized by agarose (2%) gel electrophoresis.

Results

The full length mRNA transcript (15.4 kb) arising from the mouse *Kalrn* gene consists of 60 exons spread over 604.4 kb of the genome (Fig. 1) and matches the complexity of the human *Kalrn* gene. The start codon occurs within the first exon and the stop codon of the largest isoform occurs in the 60th exon. Alternative splicing of these protein coding transcripts gives rise to about 14 coding variants, of which *Kal7*, *Kal9* and *Kal12* have been characterized in relative detail. Besides these coding transcripts we found seven presumably non-coding transcripts, referred to here as *Kalnc1* to *Kalnc7* all expressed in the same direction as the protein coding transcripts. Some of these transcripts comprise many exons and are several kilobases in length. For instance, *Kalnc1* consists of 37 exons that are together 6.1 kb in length after removal of introns. Although it overlaps significantly with the protein coding transcripts (35/37 exons), the use of distinct exons at the termini avoiding the canonical start codon results in a low coding potential. Some other ncRNA from the locus were as short as 373nt, comprising only 2 exons. We next checked their coding potential and reliability of the annotation (Fig. 1A). For coding potential analysis, we used the CPAT algorithm which recommends a score of less than 0.44 for non-coding transcripts. *Kalnc2* and *Kalnc5* showed low coding potential scores of 0.15 and 0.09 respectively, whereas the other ncRNAs arising from the *Kalrn* locus have scores ranging from 0.75 to 1.00 perhaps attributable to the overlap with coding exons of the *KALRN* gene. Four of the linear ncRNAs *Kalnc1*, *Kalnc2*, *Kalnc3* and *Kalnc6* are bioinformatically assembled transcripts, while *Kalnc4* has a relatively better level of confidence, with every splice junction being supported by at least one transcript. Interestingly, we found that a 294nt circular RNA, apparently formed by the back splicing of exon 53 and 54 has also been detected in high throughput sequencing of ribosomal RNA depleted RNA but not been verified further. Taken together, bioinformatics predictions suggested that the *Kalrn* gene locus gives rise to seven linear and one circular RNA (Fig. 1B) calling for thorough experimental validation that are essential for studying the regulatory potential of this region.

Using total RNA isolated from the brain of adult mice, and primers designed to differentiate between isoforms, we carried out RT-PCR to check the presence of these novel transcripts of unknown function. We also included a pair of pan-*Kalrn* primers that would detect the coding transcripts. As expected, the pan-*Kalrn* primers support the notion that *Kalrn* is expressed across tissues, and *Kalnc1*, the largest non-coding isoform was also similar (Fig. 2A). Even with this single pair of primers, it was obvious that *Kalnc1* shows further splicing heterogeneity. *Kalnc6*, a relatively small ncRNA, was also expressed in all the tissues tested. The circular RNA predicted by was detected using divergent PCR primers (Fig. 2B). The size of the observed band was 217nt, slightly smaller than the predicted 294nt. It was readily detected in the cerebrum, cerebellum, spinal cord and ovary of adult mice. During development it was detected only in the hippocampus at postnatal day1.

Our group has previously shown that in zebrafish, a lncRNA arising from the *Kailrn* gene locus is maternally inherited. To test if the mammalian locus also had potentially inherited lncRNAs, we included RNA from the mouse ovary in our RT-PCR experiments. All the ncRNAs barring *Kalnc7* showed expression in the ovary besides being expressed in one or more parts of the central nervous system. *Kalnc3* and *Kalnc4* were particularly interesting because they were specifically expressed in the cerebrum and cerebellum respectively, even though this end-point PCR was not quantitative.

Since these transcripts overlap with coding transcripts and each other, in situ hybridization using larger probes cannot be used to resolve the spatial expression pattern in greater detail. To look at the brain expression pattern in greater detail, we carried out quantitative RT-PCR on RNA from the cortex and hippocampus in mice of postnatal day 1, 3, 5 and 18 followed by 3 weeks, 1 month and 4 months. As shown in Fig 2B, *Kalnc4* was readily detectable (Ct value <10 with respect to *Gapdh*) at all developmental stages in the cortex and hippocampus. This level of abundance is higher than expected for lncRNAs since it roughly correlates to one or two orders of magnitude less than the abundant *Gapdh* mRNA. All the other lncRNAs were far less abundant, but a general trend in the expression pattern was that all the lncRNAs from this locus, were least abundant at P3 and P5 (postnatal days 3 and 5), a period that coincides with the switch from neurogenesis to astrogenesis (Fig. 2C). During the later stages when synapses mature, the expression of the *Kalrn* locus lncRNAs are again relatively high.

Kalnc2 was of particular interest because of its location of origin and its temporal expression pattern. As shown in Fig. 1B, *Kalnc2* is a relatively small 523nt lncRNA comprising 2 exons, with a very low coding potential. Its location coincides with the 5' end of the *Kalrn* protein coding isoforms, similar to the zebrafish *durga* lncRNA which we previously identified. To understand the role of this lncRNA in greater detail, we studied its sub-cellular localization in primary cortical neurons in culture at the time of plating and at 15 days of culture (Fig. 3A,B). As is the case with zebrafish *durga*, it was largely present in the nucleus. To resolve the intranuclear localization in further detail, we used the rodent neuroblastoma cell line Neuro2A. Here too, *Kalnc2* was localized to the nucleus and further, it was enriched in the chromatin fraction (Fig. 3C,D). *Malat1* and *Gapdh* showed the expected nuclear and cytoplasmic localization respectively, ruling out the mixing of fractions during the experiment.

Next, we characterized the expression profiles of *Kalnc2* and the major *Kalrn* mRNA isoforms, *Kal7*, *Kal9* and *Kal12* during in vitro differentiation of cortical neurons (Fig. 4). In agreement with previous reports, *Kal9* and *Kal12* are expressed during the early proliferative stages but as neurons start forming synapses and these synapses mature, their expression is reduced. The expression of the *Kal7* isoform, in sharp contrast, rises only in the later stages. In vivo, it has been shown that *Kal9* and *Kal12* dominate during embryonic development while *Kal7* is largely found in the adult brain. *Kalnc2* showed a strong similarity in expression pattern to *Kal12*, which is essential for axonal outgrowth. Like *Kal12* it was expressed most strongly in the early states of culture, and decreased gradually with the formation of the neuronal network. The circular RNA arising from a back splicing of exon 54 and exon 53 of the *Kalrn* gene corresponding to *Kal9/12* isoforms was also expressed highly in the initial stages, but decreased during differentiation until it was undetectable at 21 days in vitro (Fig. 5).

Next, we explored the role of a lncRNA arising from the comparable location in the *KALRN* locus of the human genome. The human *KALRN* locus is comparable to the mouse locus in complexity with 29 ncRNAs arising from the 692 kb locus. We found a transcript (ENST00000684441.1/ENST00000488825.5) that was annotated in the direction of the coding mRNAs at the 5' end - a position corresponding to the *durga* lncRNA in zebrafish and *Kalnc2* in mouse. The first exon of this transcript showed a sequence conservation across vertebrates and is henceforth referred to as *hsKALNC2* (supplementary data Fig. 2). We used directional reverse transcriptase PCR to validate the expression and confirm the orientation of this lncRNA. In agreement with the observation in mice, this RNA is also expressed in the cerebrum and ovary, but is excluded from tissues like cerebellum, liver and heart where *Kalrn* is readily detected (Fig. 6A). In human tissues and in the cell line SH-SY5Y, the lncRNA was of much lower abundance when compared to *KALRN* mRNA transcripts collectively (Fig. 6B). Taken together, we propose that mammalian genomes express a syntenically conserved lncRNA from the 5' end of the *Kalrn* protein coding gene which is nuclear localized and selectively enriched in sub-regions of the brain and ovary. The *mmKalnc2* and *hsKALNC2* share an identical 200nt stretch at the 5' end suggesting that mutational studies in the future should focus on this region. Owing to their conservation in mammals and relative synteny with respect to zebrafish *durga* lncRNA, we reasoned that they may be involved in similar functions.

As shown in Fig. 4, *Kalnc2* was down-regulated during differentiation of mouse primary cortical neurons in vitro. To test if the human lncRNA, *KALNC2*, may show a similar temporal regulation, we used the in vitro differentiation of SH-SY5Y under retinoic acid treatment (Fig. 7A). The human transcript was approximately 50% reduced in expression in 12 day old cultures of SH-SY5Y under the previously established retinoic acid+BDNF differentiation protocol. During this period, as expected from the development of long processes (data not shown), *KALRN* expression was increased (Fig. 7B).

To test for a direct regulatory role of the lncRNA in expression of *Kalrn* mRNAs, we modulated the expression level of *mmKalnc2*. Primary cortical neurons were cultured (as described in the methods) for 15 days. As shown in Fig. 4, at this time endogenous *mmKalnc2* is at a low abundance. The *mmKalnc2* was cloned under the control of CMV promoter in pcDNA3.1(+), this clone was transfected using lipofectamine (see materials and methods for details) and total RNA was collected at 24, 48 and 72hrs post-transfection. We measured the steady state expression of mouse *Kal7*, 9 and 12 using isoform specific primers in qRT-PCR. A similar experiment was also carried out in Neuro2A cells but at 12, 24 and 48hrs. In both cases, we could change the expression level of the lncRNA by >100 fold (data not shown), the steady state expression of the *Kalrn* isoforms was not altered in primary neurons (Fig. 8A) and Neuro2A (Fig. 8B). Therefore, according to the current evidence it appears that ectopic expression of the lncRNA cannot alter the expression of endogenous *Kalrn* mRNA isoforms in Neuro2A cells.

Discussion

Our results reveal that the *KALRN* locus gives rise to several linear and one circular ncRNA that are selectively expressed in the adult mouse nervous system. The circular RNA, *mmu_circ_0000686*,

mmKalnc2 and *hsKALNC2* from the human *Kalrn* gene locus are also expressed during the early stages of development. The same pattern is recapitulated during in vitro differentiation showing a gradual decrease with time.

The circular RNA, *mmu_circ_0000686* and the linear RNAs, *Kalnc2* and *Kalnc5* were clearly non-coding, but the other RNAs described here show some coding potential in our bioinformatics analysis, but were exclusively localized to the nucleus. In our opinion, the ambiguity in coding potential of *Kalnc1, 3, 4 and 7* is due to the overlap with coding RNAs. The inclusion of introns which usually contain stop codons may convert a coding transcript to a non-coding variant. *Kalnc6* showed some coding potential, and is at least partially localized to the cytoplasm (data not shown). Further studies are required to establish the non-coding nature of this transcript. The ultimate evidence for the coding status of a transcript can only be provided by the demonstration of the protein product. Association with ribosomes and the presence of short ORFs are both now known to be only indicators of coding potential. Further, certain transcripts may perform non-coding RNA functions and also give rise to small peptides or proteins.

The mammalian *Kalrn* locus has been extensively mined for mutations associated with several neurodevelopmental and neuropsychiatric diseases; . The protein coding isoforms arising from the locus have several interaction domains that provide spatio-temporal heterogeneity to the core Rho GTPase role of *Kalrn*. The large number of isoforms seem to depend on different promoters for their expression. Certain isoforms promote axonal outgrowth while others seem to determine dendrite numbers, dendritic arborization and maturation of dendritic spines.

We explored the possibility that *Kalnc2* drives the switch from longer to shorter *Kalrn* protein coding isoforms during neuron differentiation and maturation. However, ectopic overexpression did not change the expression level of the mouse *Kalrn* mRNA isoforms in primary neurons or Neuro2A cells. However, we cannot rule out the possibility that the lncRNA may be involved in establishing the expression of *Kalrn* isoforms during development. More refined experiments addressing this issue in specific neuronal subtypes or in specific regions of the brain may reveal a regulatory role for these lncRNAs during a short developmental window. Loss of function studies through knock-down and mutations are needed to establish the function of this lncRNA. We cannot rule out the possibility that localization and RNA modification of endogenous *Kalnc2* may not be modulated correctly through ectopic expression.

In spite of the apparent lack of sequence conservation, the lncRNAs *durga* (zebrafish), *Kalnc2* and *KALNC2* from the human *KALRN* gene locus show striking similarities: they share the same protein coding neighbour, associate with chromatin and are expressed in immature neurons. In future, chromatin immunoprecipitation can reveal the genomic region to which these lncRNAs bind, perhaps in association with chromatin remodeling complexes. They may also be involved in defining 3D chromatin organization. Neuronal differentiation and maturation are accompanied by large scale changes in 3D chromatin structure. It will be interesting to study the role of lncRNAs, especially from the *Kalrn* locus that could contribute to such spatial re-organization of chromatin in the nuclei of neurons during differentiation.

Regulatory RNAs inherited through the maternal ooplasm are a form of epigenetic inheritance . Recent studies have shown that gametes are a rich source of regulatory RNAs, including miRNA, tRNA fragments and lncRNA . A lncRNA arising from the locus is inherited through the maternal cytoplasm of oocytes in zebrafish . The expression pattern of the mammalian lncRNAs raise the possibility that they are also inherited although it is premature to assume ovarian expression corresponds to expression in oocytes. although we cannot rule out the possibility that the RNA is localized to the non-gametic tissue of the ovary. Besides the nervous system, the lncRNAs arising from the mammalian Kalrn locus are also expressed in the ovary (Fig. 2,6). Our study paves the way for further experiments into the localization of these lncRNAs in the ovary.

In summary, we report the dynamic expression of non-coding RNAs from the mammalian Kalrn gene locus highlighting the regulatory potential of this genomic region that is implicated in several neuropsychiatric and neurodevelopmental disease conditions.

Declarations

Ethics approval and consent to participate:

This study was performed in line with the Institutional Animal Ethics Committee of CSIR-Institute of Genomics and Integrative Biology (IGIB). No human subjects were part of this study.

Consent for publication:

Not applicable.

Availability of data and materials:

All data generated or analysed during this study are included in this published article [and its supplementary information files].

Competing interests:

The authors have no relevant financial or non-financial interests to disclose.

Funding:

The study was supported by the Council of Scientific & Industrial Research (CSIR) (MLP2102). Author MP has received Fellowships from Indian Council of Medical Research (ICMR) and DC and MT from CSIR.

Authors' contributions:

Conception, design & manuscript writing: Beena Pillai; Experimental work: Muneesh Pal, Divya Chaubey, Mohit Tanwar; Final approval of manuscript; All authors read and approved the final manuscript.

Acknowledgements:

The technical support provided by Arpita Konar. The animal support was provided by the CSIR-Institute of Genomics and Integrative Biology's mouse and zebrafish house facilities.

Affiliations

CSIR-Institute of Genomics and Integrative Biology, Delhi 110025, India

Muneesh Pal, Divya Chaubey, Mohit Tanwar & Beena Pillai

Academy of Scientific and Innovative Research, Ghaziabad 201002, India

Muneesh Pal, Divya Chaubey, Mohit Tanwar & Beena Pillai

Data availability statement

All data generated or analysed during this study are included in this published article [and its supplementary information files].

References

1. Cissé, M., E. Duplan, T. Lorivel, J. Dunys, C. Bauer, X. Meckler, Y. Gerakis, I. Lauritzen, and F. Checler. 2017. "The Transcription Factor XBP1s Restores Hippocampal Synaptic Plasticity and Memory by Control of the Kalirin-7 Pathway in Alzheimer Model." *Molecular Psychiatry* 22 (11): 1562–75.
2. Colomer, V., S. Engelender, A. H. Sharp, K. Duan, J. K. Cooper, A. Lanahan, G. Lyford, P. Worley, and C. A. Ross. 1997. "Huntingtin-Associated Protein 1 (HAP1) Binds to a Trio-like Polypeptide, with a rac1 Guanine Nucleotide Exchange Factor Domain." *Human Molecular Genetics* 6 (9): 1519–25.
3. Grubisha, Melanie J., Chien-Wei Lin, George C. Tseng, Peter Penzes, Etienne Sibille, and Robert A. Sweet. 2016. "Age-Dependent Increase in Kalirin-9 and Kalirin-12 Transcripts in Human Orbitofrontal Cortex." *European Journal of Neuroscience*. <https://doi.org/10.1111/ejn.13351>.
4. Grubisha, Melanie J., Tao Sun, Leanna Eisenman, Susan L. Erickson, Shinny-Yi Chou, Cassandra D. Helmer, Melody T. Trudgen, et al. 2021. "A Kalirin Missense Mutation Enhances Dendritic RhoA Signaling and Leads to Regression of Cortical Dendritic Arbors across Development." *Proceedings of the National Academy of Sciences of the United States of America* 118 (49). <https://doi.org/10.1073/pnas.2022546118>.
5. Guerra-Almeida, Diego, and Rodrigo Nunes-da-Fonseca. 2020. "Small Open Reading Frames: How Important Are They for Molecular Evolution?" *Frontiers in Genetics* 11 (October): 574737.
6. Katrancha, Sara M., Yi Wu, Minsheng Zhu, Betty A. Eipper, Anthony J. Koleske, and Richard E. Mains. 2017. "Neurodevelopmental Disease-Associated de Novo Mutations and Rare Sequence Variants Affect TRIO GDP/GTP Exchange Factor Activity." *Human Molecular Genetics* 26 (23): 4728–40.

7. Kushima, Itaru, Yukako Nakamura, Branko Aleksic, Masashi Ikeda, Yoshihito Ito, Tomoko Shiino, Tomo Okochi, et al. 2012. "Resequencing and Association Analysis of the KALRN and EPHB1 Genes and Their Contribution to Schizophrenia Susceptibility." *Schizophrenia Bulletin* 38 (3): 552–60.
8. LaRese, Taylor P., Yan Yan, Betty A. Eipper, and Richard E. Mains. 2017. "Using Kalirin Conditional Knockout Mice to Distinguish Its Role in Dopamine Receptor Mediated Behaviors." *BMC Neuroscience* 18 (1): 45.
9. Leblond, Claire S., Freddy Cliquet, Coralie Carton, Guillaume Huguet, Alexandre Mathieu, Thomas Kergrohen, Julien Buratti, et al. 2019. "Both Rare and Common Genetic Variants Contribute to Autism in the Faroe Islands." *NPJ Genomic Medicine* 4 (January): 1.
10. Mains, Richard E., Drew D. Kiraly, Jodi E. Eipper-Mains, Xin-Ming Ma, and Betty A. Eipper. 2011. "Kalrn Promoter Usage and Isoform Expression Respond to Chronic Cocaine Exposure." *BMC Neuroscience* 12 (February): 20.
11. Mandela, Prashant, and Xin-Ming Ma. 2012. "Kalirin, a Key Player in Synapse Formation, Is Implicated in Human Diseases." *Neural Plasticity* 2012 (April): 728161.
12. Memczak, Sebastian, Marvin Jens, Antigoni Elefsinioti, Francesca Torti, Janna Krueger, Agnieszka Rybak, Luisa Maier, et al. 2013. "Circular RNAs Are a Large Class of Animal RNAs with Regulatory Potency." *Nature* 495 (7441): 333–38.
13. Miller, Megan B., Yan Yan, Yi Wu, Bing Hao, Richard E. Mains, and Betty A. Eipper. 2017. "Alternate Promoter Usage Generates Two Subpopulations of the Neuronal RhoGEF Kalirin-7." *Journal of Neurochemistry* 140 (6): 889–902.
14. Mulligan, Megan K., Timothy Abreo, Sarah M. Neuner, Cory Parks, Christine E. Watkins, M. Trevor Houseal, Thomas M. Shapaker, et al. 2019. "Identification of a Functional Non-Coding Variant in the GABAA Receptor $\alpha 2$ Subunit of the C57BL/6J Mouse Reference Genome: Major Implications for Neuroscience Research." *Frontiers in Genetics*. <https://doi.org/10.3389/fgene.2019.00188>.
15. Paskus, Jeremiah D., Chen Tian, Erin Fingleton, Christine Shen, Xiaobing Chen, Yan Li, Samuel A. Myers, et al. 2019. "Synaptic Kalirin-7 and Trio Interactomes Reveal a GEF Protein-Dependent Neuroligin-1 Mechanism of Action." *Cell Reports*. <https://doi.org/10.1016/j.celrep.2019.10.115>.
16. Penzes, Peter, Richard C. Johnson, Rita Sattler, Xiaoqun Zhang, Richard L. Huganir, Vikram Kambampati, Richard E. Mains, and Betty A. Eipper. 2001. "The Neuronal Rho-GEF Kalirin-7 Interacts with PDZ Domain-Containing Proteins and Regulates Dendritic Morphogenesis." *Neuron*. [https://doi.org/10.1016/s0896-6273\(01\)00193-3](https://doi.org/10.1016/s0896-6273(01)00193-3).
17. Penzes, Peter, and Kelly A. Jones. 2008. "Dendritic Spine Dynamics—a Key Role for Kalirin-7." *Trends in Neurosciences* 31 (8): 419–27.
18. Quan, Zhenzhen, Da Zheng, and Hong Qing. 2017. "Regulatory Roles of Long Non-Coding RNAs in the Central Nervous System and Associated Neurodegenerative Diseases." *Frontiers in Cellular Neuroscience* 11 (June): 175.
19. Ruiz-Orera, Jorge, Xavier Messeguer, Juan Antonio Subirana, and M. Mar Alba. 2014. "Long Non-Coding RNAs as a Source of New Peptides." *eLife* 3 (September): e03523.

20. Russell, Theron A., Katherine D. Blizinsky, Derin J. Cobia, Michael E. Cahill, Zhong Xie, Robert A. Sweet, Jubao Duan, et al. 2014. "A Sequence Variant in Human KALRN Impairs Protein Function and Coincides with Reduced Cortical Thickness." *Nature Communications* 5 (September): 4858.
21. Russell, Theron A., Melanie J. Grubisha, Christine L. Remmers, Seok Kyu Kang, Marc P. Forrest, Katharine R. Smith, Katherine J. Kopeikina, Ruoqi Gao, Robert A. Sweet, and Peter Penzes. 2018. "A Schizophrenia-Linked KALRN Coding Variant Alters Neuron Morphology, Protein Function, and Transcript Stability." *Biological Psychiatry* 83 (6): 499–508.
22. Rybak-Wolf, Agnieszka, Christin Stottmeister, Petar Glažar, Marvin Jens, Natalia Pino, Sebastian Giusti, Mor Hanan, et al. 2015. "Circular RNAs in the Mammalian Brain Are Highly Abundant, Conserved, and Dynamically Expressed." *Molecular Cell* 58 (5): 870–85.
23. Sarangdhar, Mayuresh A., Divya Chaubey, Abhishek Bhatt, Monisha Km, Manish Kumar, Shashi Ranjan, and Beena Pillai. 2017. "A Novel Long Non-Coding RNA, Modulates Dendrite Density and Expression of in Zebrafish." *Frontiers in Molecular Neuroscience* 10 (April): 95.
24. Sarangdhar, Mayuresh Anant, Divya Chaubey, Nagesh Srikakulam, and Beena Pillai. 2018. "Parentally Inherited Long Non-Coding RNA Cyrano Is Involved in Zebrafish Neurodevelopment." *Nucleic Acids Research* 46 (18): 9726–35.
25. Sharma, Upasna, Colin C. Conine, Jeremy M. Shea, Ana Boskovic, Alan G. Derr, Xin Y. Bing, Clemence Belleannee, et al. 2016. "Biogenesis and Function of tRNA Fragments during Sperm Maturation and Fertilization in Mammals." *Science* 351 (6271): 391–96.
26. Soni, Kartik, Ashwani Choudhary, Ashok Patowary, Angom Ramcharan Singh, Shipra Bhatia, Sridhar Sivasubbu, Shanti Chandrasekaran, and Beena Pillai. 2013. "miR-34 Is Maternally Inherited in *Drosophila Melanogaster* and *Danio Rerio*." *Nucleic Acids Research*.
<https://doi.org/10.1093/nar/gkt139>.
27. Statello, Luisa, Chun-Jie Guo, Ling-Ling Chen, and Maite Huarte. 2021. "Author Correction: Gene Regulation by Long Non-Coding RNAs and Its Biological Functions." *Nature Reviews Molecular Cell Biology*. <https://doi.org/10.1038/s41580-021-00330-4>.
28. Wang, Liguu, Hyun Jung Park, Surendra Dasari, Shengqin Wang, Jean-Pierre Kocher, and Wei Li. 2013. "CPAT: Coding-Potential Assessment Tool Using an Alignment-Free Logistic Regression Model." *Nucleic Acids Research* 41 (6): e74.
29. Xie, Zhong, Michael E. Cahill, and Peter Penzes. 2010. "Kalirin Loss Results in Cortical Morphological Alterations." *Molecular and Cellular Neurosciences* 43 (1): 81–89.
30. Yan, Yan, Betty A. Eipper, and Richard E. Mains. 2015. "Kalirin-9 and Kalirin-12 Play Essential Roles in Dendritic Outgrowth and Branching." *Cerebral Cortex* 25 (10): 3487–3501.

Figures 7-8

Figures 7-8 are not available with this version.

Figures

Figure 1

Schematic illustration and coding potential prediction of mouse *Kalrn* gene locus noncoding transcripts. (A) A schematic illustration of mouse *Kalrn* long noncoding transcripts (lncRNAs) was made according to the Ensembl database; blue: coding transcripts, green: processed transcripts, red: retained

intron, black: circular RNA. Noncoding (NC) transcripts were named in the 5'end to 3'end direction from *Kalnc1* to *Kalnc7*. **(B)** Coding potential was measured by CPAT software([Wang et al. 2013](#)). Transcript flags or transcript support level (TSL) is a method to highlight the well-supported and poorly-supported transcript models. The method relies on the primary data that can support full-length transcript structure: mRNA and EST alignment supplied by UCSC and Ensembl. **TSL1** – all splice junctions of the transcript are supported by at least one non-suspect mRNA. **TSL2** – the best supporting mRNA is flagged as suspect or the support is from multiple ESTs, **TSL3** – the only support is from a single EST, **TSL4** – the best supporting EST is flagged as suspect and **TSL5** – no single transcript supports the model structure **(C)** Nuclear/cytoplasmic localization of KALNC1,3,4,6 and 7. NRT; No reverse transcriptase control.

Figure 2

Mostly *Kalrn* locus lncRNAs are brain-enriched. (A) Representative images of endpoint RT-PCR showing expression of *Kalrn* locus lncRNAs; *Kalnc1*(155bp), *Kalnc2*(131bp), *Kalnc3*(112bp), *Kalnc4*(376bp), *Kalnc5*(142bp), *Kalnc6*(138bp), *Kalnc7*(295bp); pan *Kalrn*(157bp), *mmu_circ_0000686*(217bp) and *Gapdh*(75bp) in (a-j) respectively in whole brain, cerebrum, cerebellum, spinal cord, ovary, liver and heart. (B) Schematic illustration of circular RNA *mmu_circ_0000686* at mouse *Kalrn* locus. (C) Relative

expression was measured of *Kalrn* locus lncRNAs in mouse cortex and hippocampus at different time points (P1, P3, P5, P18, P21, 1M and 4M) by qRT-PCR. (Data are shown as mean (SD), N=3 biological replicates; n=3 technical replicates).

Figure 3

***Kalnc2* is enriched in the nuclear fraction.** (A-C) *Kalnc2* expression was measured along with *Gapdh* (cytoplasmic marker) and *Malat1* (nuclear marker) in cellular fractions at (A) DIV0 (days in vitro culture), (B) DIV15 primary cortical neurons and in (C) N2A cells by qRT-PCR. (D) Endpoint PCR of *Kalnc2* in cellular fractions of N2A (Neuro2A) cells and primary cortical neurons. NRT (No reverse transcriptase control), NTC (No template control), DIV (Days of in vitro culture). (Data are shown as mean $\pm$ SD; N=3 biological replicates; n=3 technical replicates; t-test, *p<0.05; **p<0.01; ***p<0.001)

Figure 4

***Kalnc2* expression is similar to *Kal9* and *Kal12*, but opposite to *Kal7*.** (A-G) Bright field images of mouse neural stem cell differentiation and maturation in neurons. (A) DIV0, (B) DIV3, (C) DIV7, (D) DIV11, (E) DIV15, (F) DIV21, (G) DIV28; scale bar 100µm. (H) Endpoint PCR for *Kalnc2*(110bp) and *Kalrn* major coding transcripts (*Kal7*(127bp), *Kal9*(125bp) and *Kal12*(140bp) at different DIVs (DIV0, DIV3, DIV11, DIV15, DIV21 and DIV28). (I-L) *Kal7*, *Kal9*, *Kal12* and *Kalnc2* expressions were measured at DIV0, DIV3, DIV7, DIV11, DIV15 and DIV21.

DIV11, DIV15 and DIV21. DIV(Days of in vitro culture). (Data are shown as mean $\pm$ SD; N=3 biological replicates; n=3 technical replicates; t-test, *p<0.05; **p<0.01; ***p<0.001).

Figure 5

The expression of circular RNA decreases, as cortical neurons mature. (A and B) Circular RNA expression was measured at DIV0, DIV3, DIV7, DIV11, DIV15 and DIV21 with respect to *Gapdh* by quantitative RT-PCR. DIV(Days of in vitro culture).(Data are shown as mean $\pm$ SD; N=3 biological replicates; n=3 technical replicates; t-test, *p<0.05)

Figure 6

***hsKALNC2* is exclusively expressed in brain tissues.** (A) Expression of *KALNC2* (119bp) and *KALRN* (238bp) in human tissue RNA panel available commercially (B) Subcellular fractionation of SH-SY5Y cells *18SrRNA*, *MALAT1* and *45SrRNA* used as cytoplasmic, nucleoplasmic and chromatin markers respectively. (C) Expression of *KALNC2* and *KALRN* in SH-SY5Y cell normalized to *GAPDH* expressions (Data are shown as mean $\pm$ SD; N=3 biological replicates; n=3 technical replicates; t-test, * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterial2.docx](#)