The study of differential expressions of MCPH and Seckel syndrome genes and their paralogues

Authors: Mohsin M[1], Riaz S[1], Imran M[1], Cox JJ[2] and Khurshid M[1,3]

Affiliations: 1- Department of Biochemistry & Molecular Biology, the Islamia University of Bahawalpur, Pakistan

2- Wolfson Institute for Biomedical Research, University College London, UK

3- Corresponding author: maryam.khursheed@iub.edu.pk

Abstract

Microcephaly forms which are established are; MCPH, that is non-syndromic and Seckel syndrome which is syndromic and carries microcephaly with dwarfism whereas earlier form carries only microcephaly. The microcephaly genes are parts of essential cellular pathways but are regulated differently in both forms of microcephaly. These genes have diverged from single ancestral genes in to related but different gene paralogues during evolution. It is suggested that the paralogues might be redundant in other body organs and might rescue any abnormality in case of MCPH but not in Seckel syndrome. The important aspect is to study the spatio-temporal dynamics of these genes together with their paralogues at mRNA level where the different tissues of embryos might exhibit different transcript variants or different levels of the various transcript variants to find the complementation in case of the MCPH but divergence in case of the Seckel syndrome. Here, we studied Cdk5Rap2-Pde4dip, Phc1-Phc3 (for MCPH) and Cep63-Ccdc67 (for Seckel) in the embryonic mice tissues via RT-PCR throughout the peak of neurogenesis from E12.5 to E18.5. We found already known and novel transcript variants in the tissues in addition to different levels of expression proving differential regulation. The results of PCR were analysed on the agarose gel and were quantified by giving relative intensity percentage scores to each bands. Our data, shows that Cdk5Rap2-Pde4dip showed up-regulation at mid-late neurogenesis and showed redundancy whereas Cep63-Ccdc67 duo was very different and expression of Cep63 was consistent throughout neurogenesis, but Ccdc67 gradually increased during development and was highest at the late neurogenesis and development suggestive of diverged and essential roles of the two genes which are both essential in development and lastly, Phc1-Phc3 showed up-regulation at early and late neurogenesis and seem to have complementary roles in brain development. The expression profiles of Cep63-Ccdc67 were very different indicative of their diverged roles in development but the MCPH genes and their paralogues followed complementary patterns suggestive of complementation in brain development.

Introduction

Microcephaly both prenatal and postnatal is reported in humans. It might be due to genetic factors or genes that might cause non-syndromic microcephaly(1) where only head size is reduced or syndromic forms where all organs are small together with craniofacial and other abnormalities(2,3). Autosomal recessive primary microcephaly or MCPH is non-syndromic condition in which affected have brain size reduction together with mild to moderate mental

retardation and rest of the organs in the body are of normal size(4). Seckel syndrome on the other hand, is syndromic microcephaly which occurs in utero but is progressive and shows severe microcephaly with dwarfism and other anomalies(2,5). Both MCPH and Seckel syndrome are heterogeneous and hence, indicate the different mechanisms of pathogenesis resulting from the gene mutations which might overlap sometimes such as MCPH6 gene CENPJ which is also involved in the Seckel syndrome(3,6). Both are heterogeneous and single genes are causative agents of both, the agents which are inherited in autosomal recessive manner so there must be some differential regulation done by the genes in each category and must control development in different manner(3,7). These genes are targeted to several core pathways such as centriole biogenesis, centrosome/spindle pole, kinetochore-microtubular pathway, nuclear pathways including the condensation-decondensation and nuclear envelope dynamics and other sub-cellular pathways(1,8–10).

During evolution genes diverge from ancestral genes in to different molecular entities with different acquired functions that may be related as these retain a considerable similarities termed as gene paralogues(11). Primary microcephaly genes have been studied in the past decade for expression and by gene knock out in animal models which have given a large insight in to the disease but the data suggests ubiquitous roles of these proteins in all body cells and mice knockout are not true models of the disease with small to no reduction of organs(12). Furthermore, the comparative expression analysis of genes and their paralogues might answer the question. Some studies are done with double knockout of more than one PM gene and found a severe phenotype indicative of involvement in one master pathway of the cell incorporating all dynamics such as cell cycle, centrosomal/spindle pole, centriole biogenesis, nuclear and others where these genes have established roles(13–15).

The one aspect that was missing was the differential expression patterns of microcephaly genes and their paralogues that can be explored for spatio-temporal expression analysis together with their established paralogues in both forms of microcephaly(16,17). Here we selected *CDK5RAP2-PDE4DIP* and *PHC1-PHC3* genes and their respective paralogues duos for MCPH and *CEP63-CCDC67* genes from Seckel syndrome(8,18–23). These genes were selected because of their targeting was to different pathways that might be differentially important at different embryonic days. We used Swiss albino mice for the study and did mating via timed mating protocol and got staged embryos(24). Later, we employed RT-PCR using gene specific primers to understand if there is some variations in transcripts or/and level of expression of the transcripts.

CDK5RAP2 at MCPH5 is involved in centrosome dynamics whereas *PHC1* at MCP11 is involved in nuclear dynamics and Seckel gene, *CEP63* at SCKL6 is a core centrioler gene and is involved in centriole duplication and their respective paralogues have related functions in the same or different pathways(4). In order to see if genes and their paralogues complement each other or if they replace each other's function being redundant or/and completely diverged and have different roles we set out to understand via simple but informative RT-PCR analysis an established method to uncover differential transcript variants or/and levels of expression.

Materials & Methods

Animal mating and tissue harvesting

Swiss albino mice strain was used and the mice were purchased from the Department of Pharmacy at Islamia University of Bahawalpur. The housing was done as per international standards. Timed mating of mice (follow DOI: 10.1101/2023.12.16.57983) was done at proestrous-estrous stage and staged embryos were collected later at specific required days in RNA later™ (Invitrogen) after dissecting mice anesthetized by xylazine-ketamine in 1:10 combination and were subjected to cervical dislocation. The mice work was approved by institutional ethical committee of IUB via letter no. 3990/ORIC.

RNA Extraction

The total RNA was extracted from 100mg of tissue using TRIzol (Ambion) method. The minced tissue was incubated with 1mL of TRIzol reagent overnight at 37°C. Next day, the DNA and Proteins were separated by phase separation and RNA was pelleted down. The pellet was washed with 70% Ethanol and air dried at room temperature for 5 mins. Pellet was resuspended in 20µl of RNase/DNase free water.

First strand cDNA synthesis

First strand cDNA was synthesized after calculating the RNA as 1µg from nanodrop method (Gen5 software). RevertAid (Thermo fisher scientific) first strand cDNA synthesis kit was used to do the reverse transcription using the described protocol. OligodT primers supplied with the kit were used to amplify the first strand.

Second strand cDNA synthesis

The first strand cDNA synthesized was amplified into double stranded DNA by using gene specific primers and the estimated known sizes were upto 900bp so Taq DNA polymerase system (Thermo fisher scientific/ Synbio) was used as a standard method.

Gene specific primers

The intron spanning gene specific primers targeting all known common transcript variants and sometimes specific transcripts were used. Primer 3plus and primer blast was used for the primer designing. The Insilico- PCR was done to calculate product sizes and targeted transcripts from UCSC genome browser. The melting temperatures and GC content was also calculated. The sequences are attached in the supplementary tables.

PCR program

The PCR was done at initial denaturation at 95°C for 5 mins, 35 cycles of denaturation at 95°C for 30 sec, annealing at 53/55/56°C for 30 sec and extension at 72°C for 1 min. Final extension was done at 72°C for 10 mins. PCR machine was set at 4°C forever.

Gel electrophoresis

The 2% agarose gel was prepared and 3µl ethidium bromide was added for viewing in UV transilluminator. The 100 bp DNA ladder was used with each primer to determine the size of bands.

Relative Intensity scoring

The DNA bands were marked by presence and relative intensity. The highest intensity was scored as 4 (100%) and was further calculated down to 1(25%) as intensity in samples drop. The absence of band was numbered 0 (0%). Later, cumulative scores of all bands for each primer were added and relative percentage was calculated for each sample. Means were taken for more than one sample wherever applicable.

Excel software

Free Microsoft excel was used to plot the relative percentages for each samples. The error bars are also added.

Results

CDK5RAP2 and *PDE4DIP*

CDK5RAP2 at MCPH3(4) is one of the first few MCPH genes to be identified. CDK5RAP2 and PDE4DIP/Myomegalin (21) are identified and confirmed as paralogues of each other by Position specific iteration BLAST, PSI-BLAST program. The multiple sequence alignment, MSA showed very high similarities between the two proteins throughout most regions of the two proteins. These were further subjected to spatio-temporal investigation (17,25) in mice staged embryos in order to see if they actually can act by replacing each other. Cdk5Rap2, a protein kinase involved in the centrosome dynamics is found be most highly expressed in the head of E14.5 mice and interestingly its paralogue PDE4DIP is also most high in the E14.5 mice. At E14.5 the basal progenitors are abundant and it is the mid-neurogenesis peak and hence represent that both genes are essential most specifically to the basal progenitors. Furthermore, we observed expected transcript variants in the samples and also some unexpected bands for potential new transcript variants. We plotted the cumulative expression as well as only the expected transcript variants' expression by intensity scoring and also the potential new transcript variants were treated the same way. All of these different scenarios gave us the similar profile where the expression of *Cdk5Rap2* gradually up-regulated after E12.5 and E13.5 and became highest in the E14.5 head and stayed at the same level at E15.5 and E18.5 in the mice embryonic head suggesting a very important role of Cdk5Rap2 in these stages when basal progenitors are most abundant and SVZ is the major layer in neuroepithelium. Its paralogue, *PDE4DIP*, the highest expression of all the transcript variant expected and novel in the E13.5 was shown to be on top followed by E14.5. However, in the potentially new variants the E14.5 was highest and the expression down regulated in all scenarios later in neurogenesis. This is very interesting as at E13.5 and E14.5 there are apical as well as basal progenitors with later being abundant and it might be possible that the embryonic development in

mice is dependent on the expression of genes and there paralogues in the brain where both are required for brain development and are complementary but the time window for the paralogue expression is short compared with the MCPH gene, *CDK5RAP2* suggesting prolonged role of this gene is required for normal brain development which is also the longest stage in the human development.

Basal progenitors in humans and in the higher primates are largest in number and contribute neurons to all six layers of the cerebral cortex and thought of as the evolving factor between the highly convoluted gyrencephalic brain in higher mammals and lissencephalic brain as is in the mice. Our data suggest roles of CDK5RAP2 and PDE4DIP more critically in these progenitors. Both *CDK5RAP2* and *PDE4DIP* are shown to be expressed in similar locations such as centrosomes and involved in microtubule and spindle dynamics furthermore, the *PDE4DIP* if mutated cause anomalies of the cardiac system suggested the diverged roles of these proteins in different organ development. We suggest both these protein are essential for development where CDK5RAP2 specifically controls the brain development by prolonged expression through neurogenesis whereas *PDE4DIP* expresses in small time window that might be essential for the other organs development

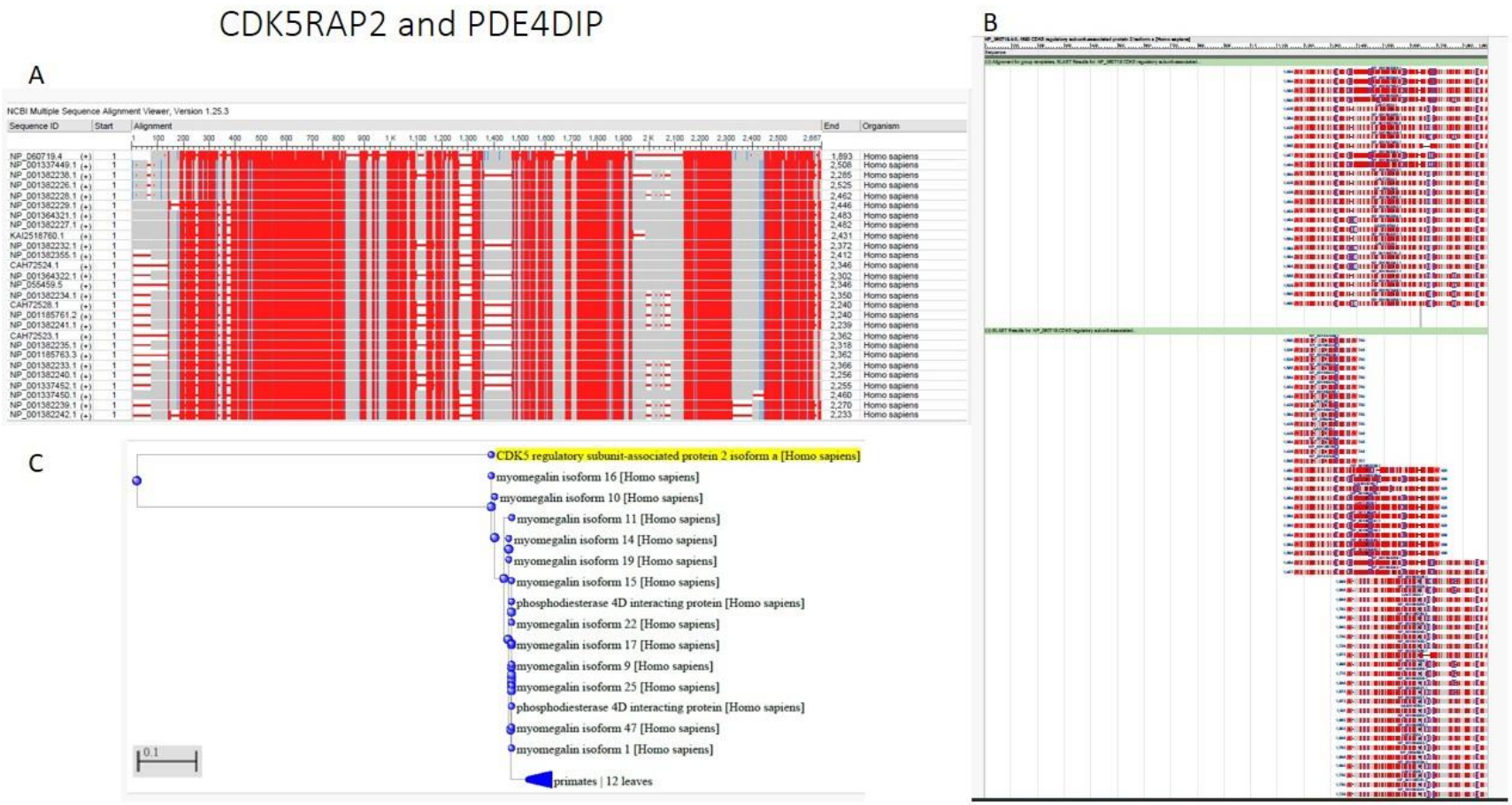


Figure 1: PSI BLAST was run to identify and confirm the ancestry, similarity and overall similarity between CDK5RAP2 and PDE4DIP/Myomegalin in Humans. A) Multiple Sequence Alignment between the CDK5RAP2 and Myomegalin/PDE4DIP is shown and is indicative of very high similarity especially at the N-terminus and other domains, B) BLAST between Isoform 1 of CDK5RAP2 was done and came up with similarities in various regions with almost all isoforms of Myomegalin/PDE4DIP, C) Ancestry of both CDK5RAP2 and PDE4DIP was confirmed by Gene tree where both genes split in a single evolution event and myomegalin showed a lot of further splits indicative of active evolution.

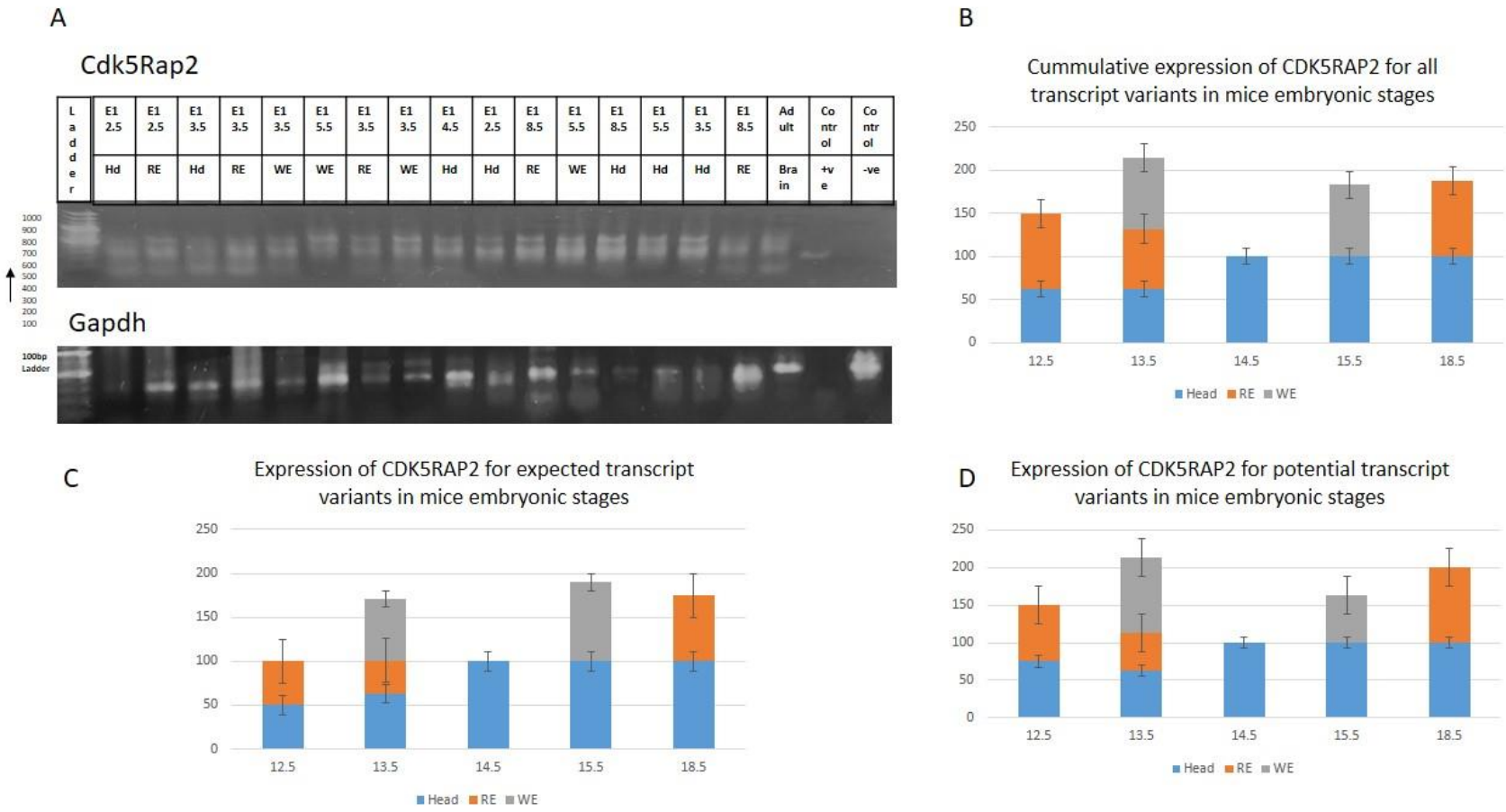


Figure 2: *Cdk5Rap2* an MCPH gene was subjected to transcriptional analysis by RT-PCR using gene specific primers. The *Gapdh* was used as a control. The results have shown expected band sizes and some potential new variants. The analysis was done by relative quantification of level of expression by band density and by given it a score from 4(100%) to 0(0%). The values for all different stages and embryonic regions were plotted in the graphs. A) Gel pictures, B) Cumulative expression at spatio-temporal dynamics C) Expected transcript variants scores D) New potential variants were plotted for each embryonic stage and region.

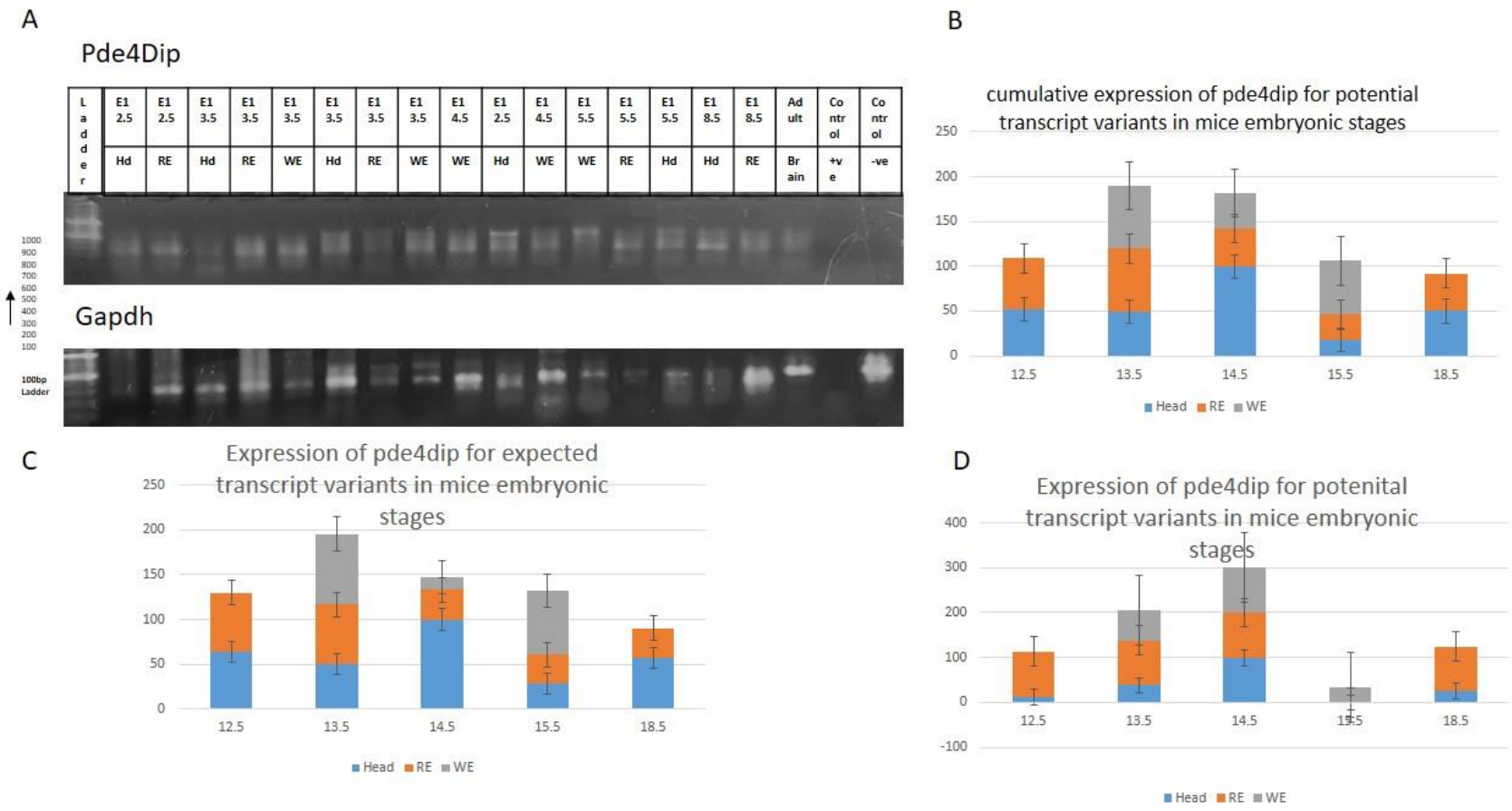


Figure 3: *Pde4dip/Myomegalin* was subjected to RT-PCR analysis with mRNA being by oligodT primers and dsDNA was made by gene specific primers. A) Gel pictures of *Pde4Dip* and *Gapdh* are shown at various embryonic days in mice in various embryonic regions (i.e., Head, rest of the embryo and whole embryos). B) The relative quantifications data for cumulative scores of all transcript variants is shown in various stages of neurogenesis in mice in various regions. C) Expected transcript variants data.

CEP63 and *CCDC67*

CEP63 at SCKL6 on chromosome 3q22(8,13,23) is a gene involved in the broad development involving different organs and organ systems. The autosomal recessive mutation in *CEP63* gene cause overall reduction leading to dwarfism and severe microcephaly. The CEP63 protein as a core centrosomal protein involved in the centriole duplication and recruitment of CEP152 at MCPH9 codes for a protein that also targets to the centrosome. Centriole duplication is an important part of centrioler/centrosomal/spindle pole pathway which is required for faithful mitosis and cell cycle progression. Its paralogue CCDC67/DEUP1(22) is also studied to be involved in the centriole duplication and de novo centriole assembly and hence thought to compliment the roles of CEP63 in development. *DEUP1* mutations reported show the multi-organ affected with the disease including retinopathy, cardiomyopathy, hypogonadism and others however, the reduction in size of the brain is not reported and mental retardation is not common. Hence, although both proteins are important for development CEP63 seem to have brain specific roles. The two proteins were subjected to PSI-BLAST and MSA showed a smaller similar regions which were coiled coiled regions essential to centriole/centrosome targeting. We then started to dig in to spatio-temporal expression analysis of both Seckel gene and the paralogue. We found that cumulative expression of all the transcript variants for *Cep63* in the various embryonic samples at all stages of neurogenesis was more or less consistent from E12.5 to E18.5 with highest expressions in E18.5 head. The novel transcript variants were expressed in embryonic head at similar levels at E12.5 then down-regulated at E13.5 and then again up-regulated and remained at the same highest level as in E12.5 from E14.5 to E18.5 suggesting important functional relevance to the early and late progenitors. Whereas *DEUP1* cumulative expression gradually up-regulated from E12.5 to E18.5 in a consistent way. The potential novel variants also showed the same trend with gradual up-regulation from E12.5 to E18.5. The rest of the embryo where we separated embryonic head from rest of the body showed gradual increase later from E15.5 to E18.5. The *Cep63* in the rest of the embryo samples also showed gradual up-regulation from E15.5 to E18.5 with highest being at E18.5. The data suggests that the role of DEUP1 to be more general and is required in large copies at the end of neurogenesis and development, on the other hand CEP63 seem to be highly essential at early to mid and late neurogenesis with highest cumulative scores at E12.5, E15.5 and E18.5.

Both genes being highly up-regulated at mid to late gestation points shows their potential co-relation with the SVZ progenitors which are abundant at these stages. There are also small proportion of OSVZ progenitors in mice which are detected in late neurogenesis and are of high relevance in humans and both genes seem to be also contributing to their neurogenesis. The rest of the embryo gave some up and down regulation in *Cep63* but remained stable in Deup1 suggesting a more controlled and consistent expression of *Ccdc67*. These results suggest that at the time of development Cep63 and Ccdc67 essentially control the brain development but the other embryonic tissues require the maintenance of Ccdc67 throughout embryonic development. Hence broad spectrum disorders of organs and organ systems are found in both *Cep63* as well as *Ccdc67* in mutated individuals.

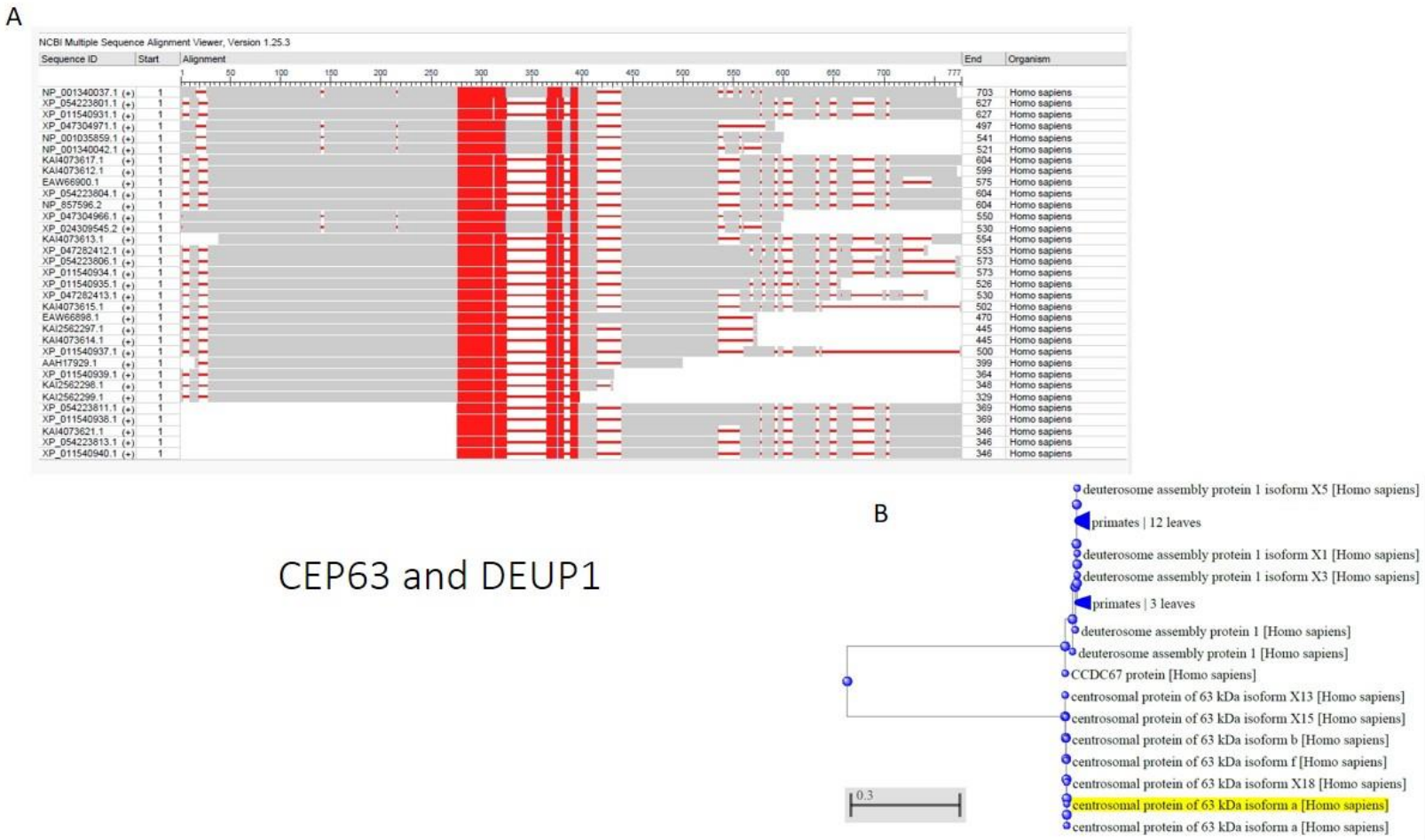


Figure 4: *CEP63* and *DEUP1/CCDC67* are two genes which are both undergoing active evolution and got split from a same ancestral gene. A) The Multiple Sequence Alignment of CEP63 is shown with various isoforms of DEUP1 and shows regions of high similarities scattered throughout but mostly in the center. Both proteins are targeted to the centrioles and the shared regions have coiled coiled domains important to the centrioler targeting. In addition they also harbor a lot of varying regions where the two proteins are clearly showing divergence and new acquired functions B) Gene tree of both gene is shown where they have split in a single evolution event and undergoing active evolution.

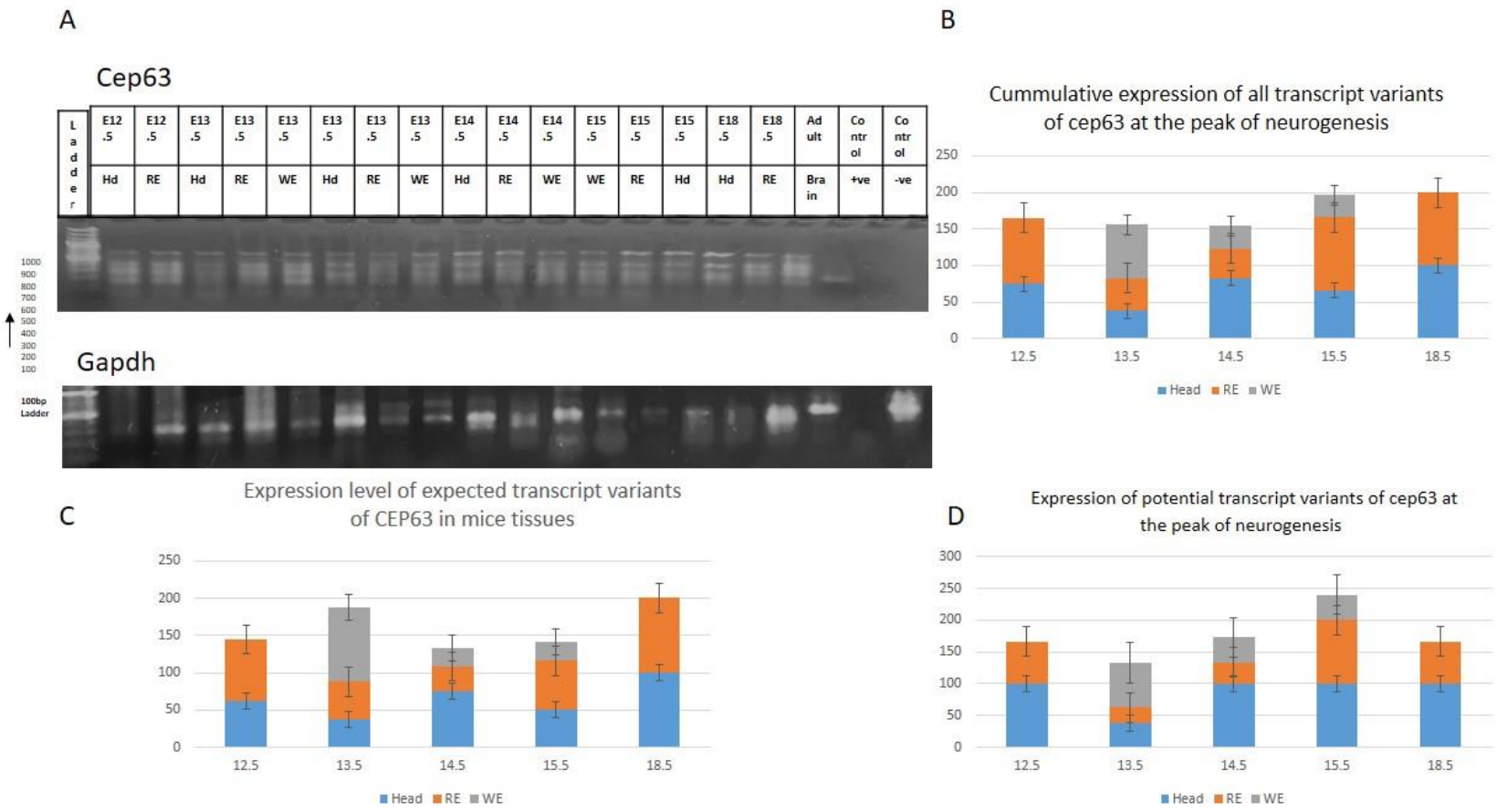


Figure 5: RT-PCR was done for *Cep63* throughout the peak of neurogenesis in mice i.e, E12.5 to E15.5 and E18.5. At each stage the mice embryos were divided in to the head regions, rest of the embryo and some whole embryos were also processed. Gene specific primers were utilized to do RT-PCR. A) Gel picture of RT-PCR in *Cep63* and Control *Gapdh* B) The cumulative expression of all the transcript variants showing highest expression at E12.5 but is showing quite consistent expression whereas head of E18.5 shows highest cumulative expression the time of OSVZ cell abundance. B) Expected transcript variants shows highest expression at E18.5 C) the potential new transcript variants shown expression at earlier and later stages.

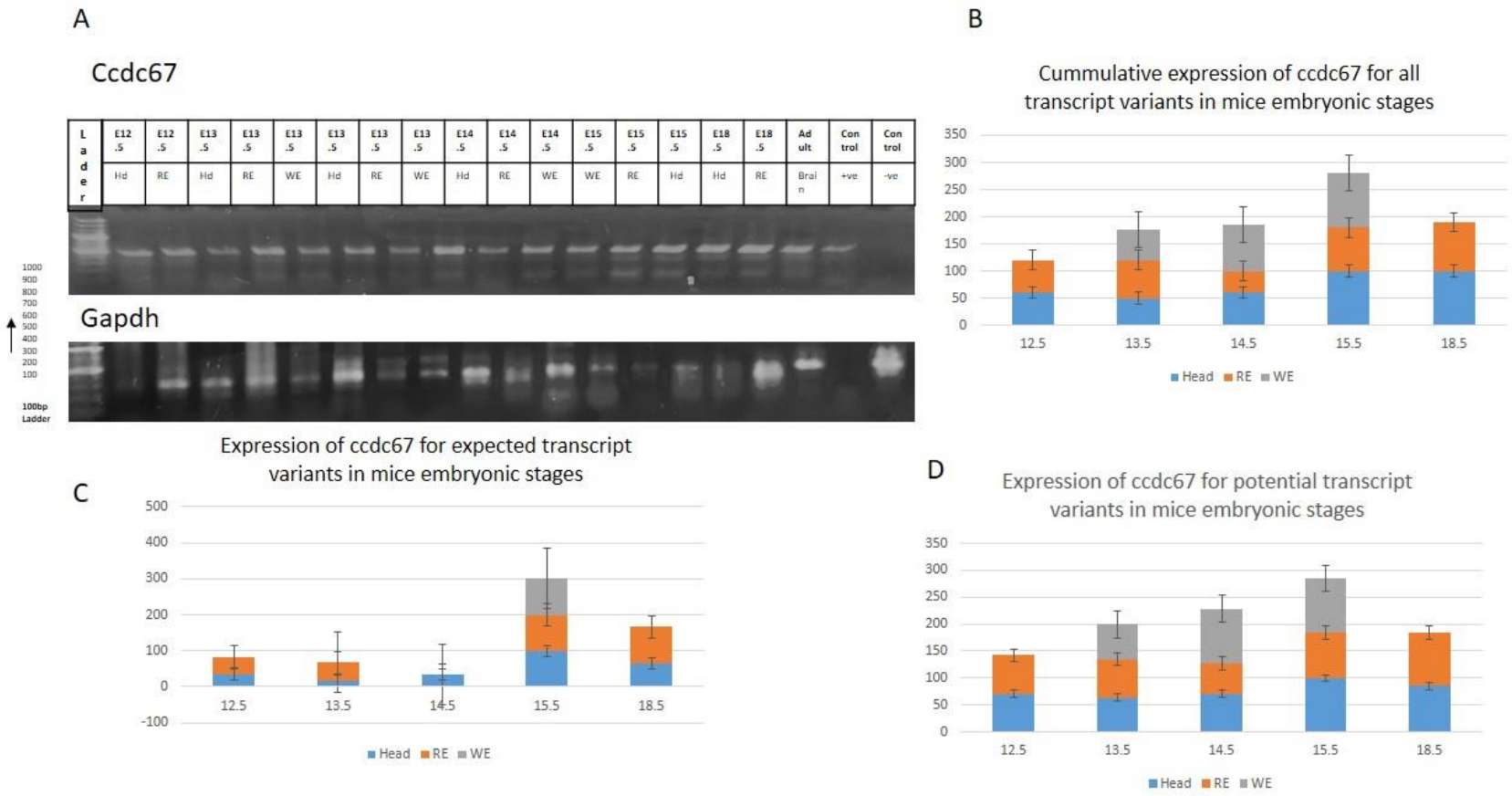


Figure 6: DEUP1 /CCDC67 a centriole targeted protein also a paralogue of Cep63 which also targets to centrioles and is a core component is subjected to transcriptional analysis by RT-PCR. A) Gels of *Ccdc67* and control *Gapdh* B) *Ccdc67* gradually increased from E12.5 to E18.5 reaching highest level gradually at later stages in cumulative scores B) The expected transcript variant also gradually increased C) The potential new variants showed same trend with quite consistent expression in the head region but collectively up-regulated from E12.5 to E18.5

PHC1 and *PHC3*

PHC1 and *PHC3*(19) are the members of Polyhomeotic-Like Proteins family and are targeted to the PcG- PRC1, Polycomb group of repressor complexes, and is involved in maintenance of pluripotency during development by repressing the transcription of genes involved in differentiation especially *HOX* genes in addition to other important roles in the nuclear dynamics(18). This might be of significance in relation to MCPH where a single change in PHC1 can lead to considerable brain size reduction while leaving other organ systems intact. The *PHC3* is one of the paralogues of *PHC1* and was selected for this study because they split from the ancestral genes as two diverged genes in evolution as a single event and hence, could indicate separate acquired functions. The PSI-BLAST showed that two paralogues are very similar and MSA showed C-terminus conservation which are essential domains for chromatin remodeling by PRC1 complex suggestive of complimentary roles. We set out to study their spatio-temporal dynamics by RT-PCR in mice embryonic tissues at the times of neurogenesis from E12.5 to E18.5. RT-PCR results show that the total expression of *Phc1* are highest at E12.5 and E13.5 with gradual down regulation at E14.5 and E15.5 whereas it again up-regulated at E18.5. The expression in the embryonic head was highest at E12.5 and later at E18.5. The expression of rest of embryonic tissues showed consistent expression from E12.5 to E18.5. Hence indicating Phc1 to be essential in development in general whereas a controlled expression patterns in the brain indicating its control in early and late neurogenesis. Its paralogue, PHC3 which is thought to complement its function in the development is also investigated for mRNA expression. Our results show that *Phc3* was highly expressed E12.5 and slightly down regulated at E13.5 to E14.5 and enhanced at E15.5 to E18.5 being highest at these later stages. To dissect it further the embryonic heads and rest of embryos of mice showed consistent *Phc3* expression which was highest at E12.5 and later at E15.5 to E18.5. The data indicate complementary roles of Phc1 and Phc3 in the brain whereas early and late neurogenesis is controlled by these proteins. Interestingly, RT-PCR has showed a novel variant of Phc3 in embryonic heads of early E12.5 and E13.5 in addition to common variant which was expressed equally in all tissues suggesting its relevance to the brain only. This suggest that in the brain Phc1 function is complemented by *Phc3* whereas rest of the embryos the Phc3 and Phc1 play essential diverged roles which might not be sensitive to single gene mutations.

At early and late neurogenesis, the radial glial cells and OSVZ cells are abundant and both have high importance in human brain development during neurogenesis. The radial glial cells being the initial progenitor cells and give rise to various progenitors and neuron lineages might be controlled by these two genes in a way that PHC1 comes first and then PHC3 complements and control the final brain size. At MCP11 *PHC1*, if mutated leads to MCPH and indicate only brain reduction and other body organs are normal in size and might involve these spatio-temporal pattern.

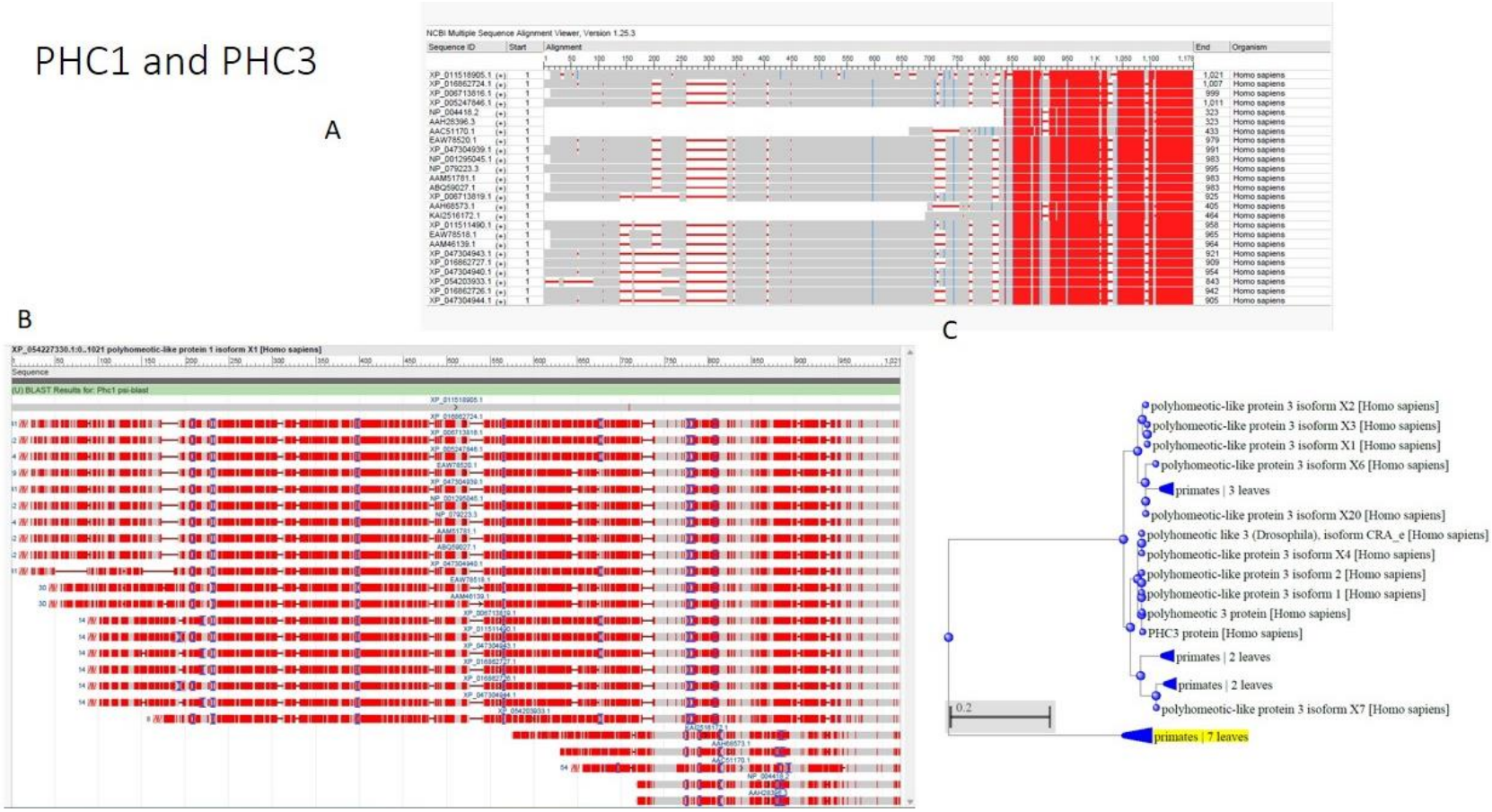


Figure 7: PSI-BLAST was used to identify, calculate similarity and ancestry of MCPH gene, *PHC1*'s paralogue, *PHC3*. Both genes are parts of PcG-PRC1 complex components and control pluripotency by chromatin remodeling. They seem to be targeted to the same pathway and were thought to be complementary to each other. A) PHC1 and PHC3 were aligned by Multiple Sequence Alignment and were shown to be highly similar in the C terminus regions and several other regions. The C terminus of both the proteins has FCS-Zinc finger domains and SAM domains essential for the chromatin remodeling B) The BLAST of Isoform 1 of PHC1 with all isoforms of PHC3 showing regions of similarity throughout in both proteins but very consistent C terminus C) The gene tree of *PHC1* and *PHC3* showing same ancestral gene and active evolution in PHC3.

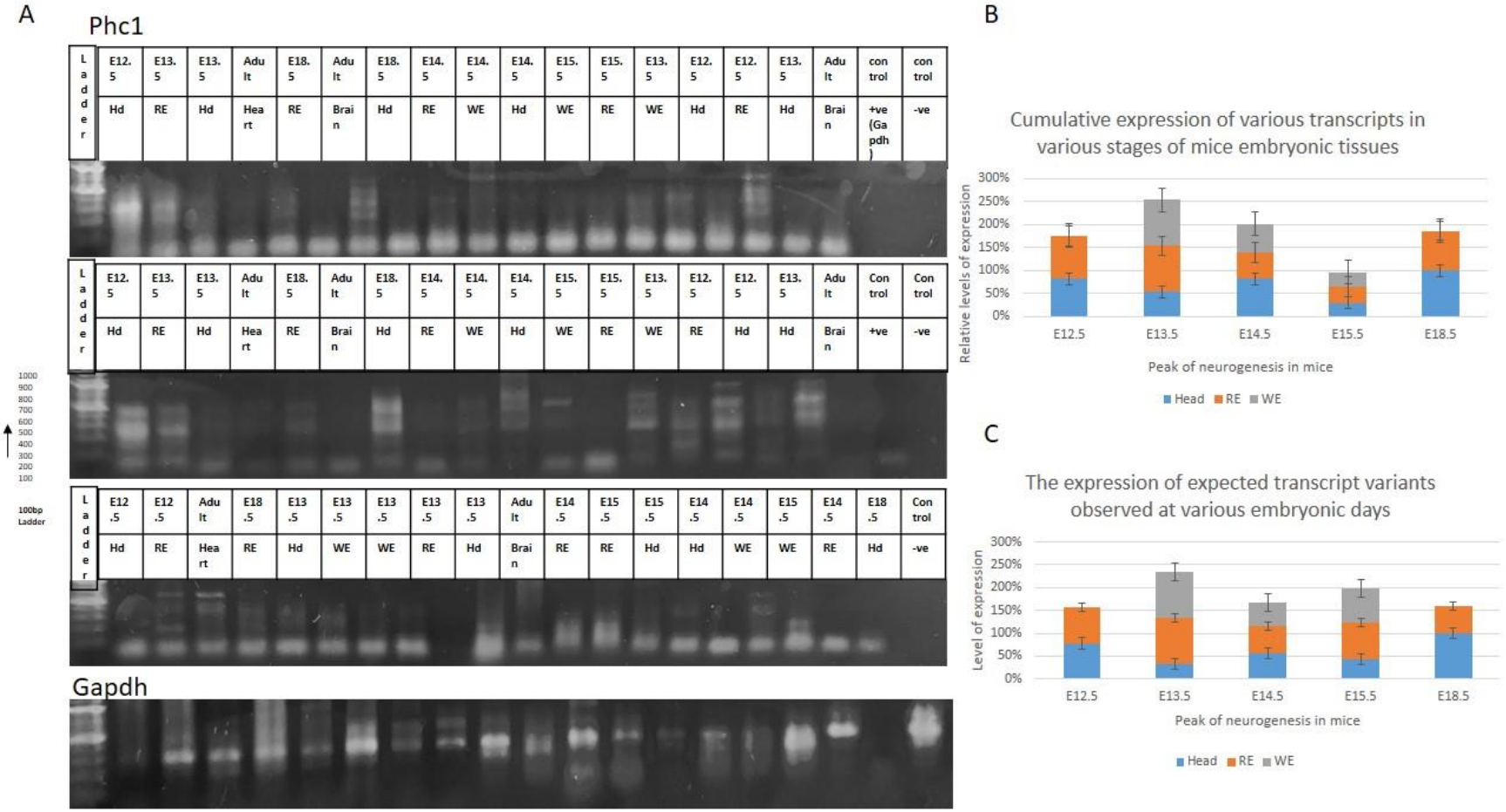


Figure 8: The RT-PCR data of Phc1 in various samples of mice embryos from peak of neurogenesis by gene specific primers. Three pairs of primers were used. A) Gel pictures of *Phc1* and Gapdh primers amplified products by ethidium bromide. B) Relative quantified band intensity for cumulative expression for all the transcript variants showing highest expression in the start, mid and end of neurogenesis but overall showing consistent expression which dips from E14.5 to E15.5 but again rises at E18.5 C) The potential new transcript show similar trend but it relatively consistent at start and end.

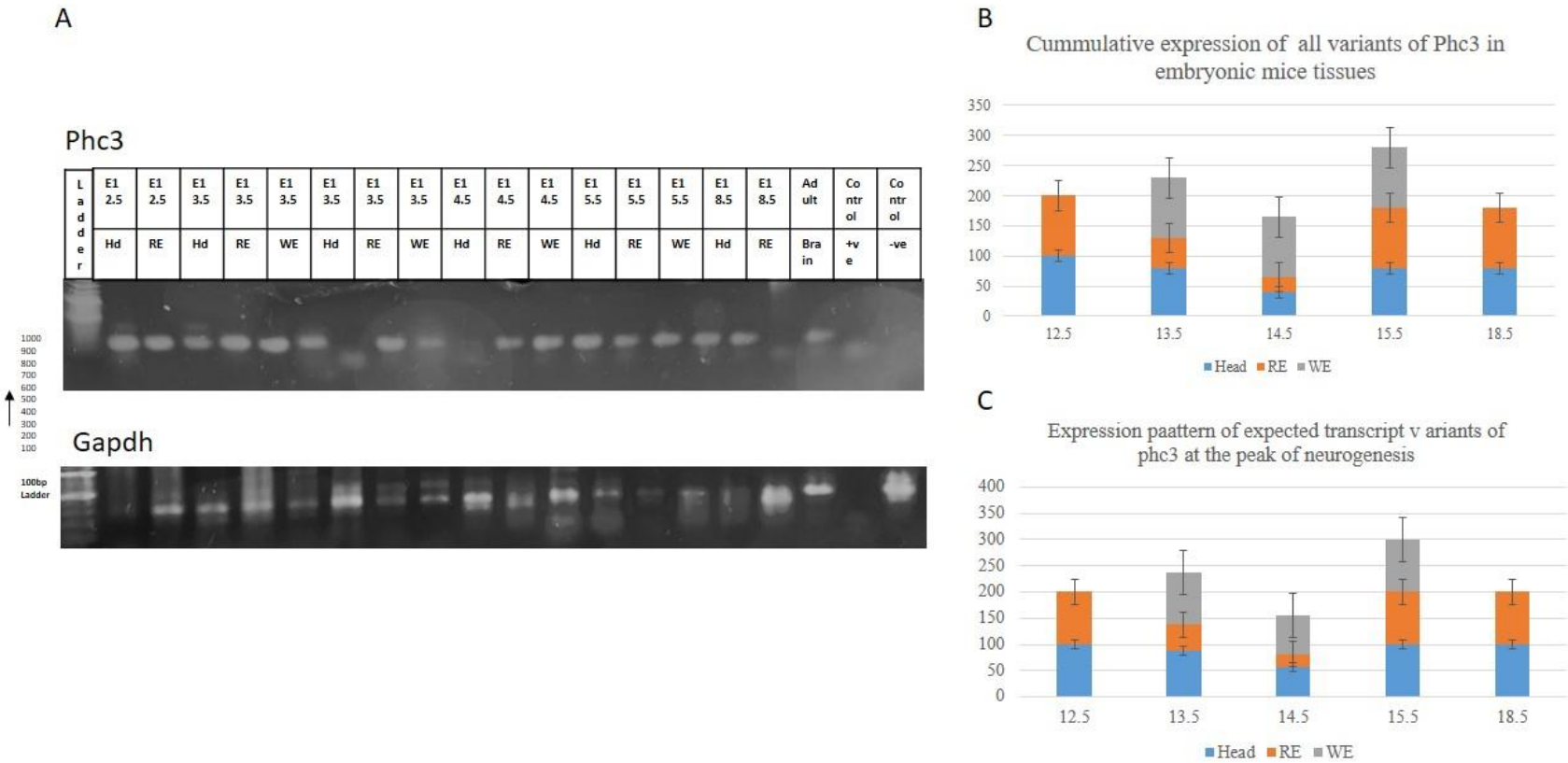


Figure 9: The RT-PCR data of PHC3 by gene specific primers. A) DNA Gel electrophoresis pictures of *Phc3* and control *Gapdh* showing the presence or absence and level of intensity of each sample bands. The peak of neurogenesis in mice was targeted by getting staged embryos from E12.5 to E15.5 and then E18.5 for late neurogenesis. The samples were prepared as head, Hd, rest of embryo, RE and Whole embryo, WE and are showing above each band B) The intensity level of each band was relatively quantified and given scores from 4(100%) to 0(0%) and were plotted for cumulative expression scores and found to be highest throughout neurogenesis with highest at earlier stage of E12.5 but otherwise consistent C) The potential new transcript variant was found to be present in E12.5 and E13.5 head and may be significant but all expected bands showed high trend in start and end of neurogenesis.

Discussion

The MCPH and Seckel are disorders involving a very tragic microcephaly condition which is further added upon by dwarfism in the syndromic later form. Scientists are also fascinated by this condition because the microcephaly is similar to the head sizes in our closest relative higher primates. Not only the size of head, which is most notably small in the cerebral cortex region that is exceptionally developed in *Homo Sapiens,* but it sadly limits the children to learn language or do very complex tasks and moderately to severe mental retardation which might just reflect their inability to mimic human specific brain functions but are quite decent. It is a disorder but opens up new avenues in to studying evolution and human acquired traits. The most work has mostly focused on the uncovering of pathogenic genes which was really challenging and took a long time and was speedup by high throughput sequencing and computational analysis. Furthermore, some really important functional analysis was done on these genes and found to be highly informative but end at a dead point. All these genes are part of essential core pathways that control cells and neural progenitors as the type of cell also have it in all cells and seem to be ubiquitous more or less furthermore, their up-regulation or down-regulation is also reported in the many cancers and tumor cells indicating their regulated roles. Patient derived cell lines show reduced or absent expression but it is also not true reflection when we think of it as developmental disorders. Animal models have shown some promise but had varying results. We set out to the basic developmental principal where genes are differentially regulated with respect to stage and time and controls different specific roles that they play in the development. We set out to study three pairs of genes and their paralogues which are targeted to different pathways and both MCPH and Seckel syndromes. We used RT-PCR to check if there was any variation in their expression levels or presence of absence of transcripts at various stages throughout peak of neurogenesis and we numerically calculated the levels of presence or absence and sometimes used many primers and added up all scores and interestingly these have actually given us positive insight in to the disorders and it is the first report to actually do so. This will further open new chapters in to the exciting research on DNA/RNA and protein levels in future and combing high tech machines could give important insights in to this field of spatio-temporal dynamics during development.

Conclusions

Our data suggests that neural progenitor cell divisions in to neurons through symmetric and asymmetric divisions are controlled by various cellular pathways and MCPH genes indicate series of up-regulation at different embryonic days suggesting their overall hierarchy in the control of neurogenesis. In addition their respective paralogues seem to complement their roles in neural development which might indicate their divergence in the other organ development apart from the brain. *Cdk5Rap2-Pde4dip* showed up-regulation at mid-late neurogenesis and seem to have some redundancy whereas *Cep63-Ccdc67* duo was very different and expression of *Cep63* was consistent throughout neurogenesis, but *Ccdc67* gradually increased during development and was highest at the late neurogenesis and development suggestive of diverged and essential roles of the two genes which are both essential in development and lastly, *Phc1-Phc3* showed up-regulation at

early and late neurogenesis and might be complimentary for brain development. Hence, suggesting the roles of different pathways are important for cell divisions at different times and stages and might be possibly part of a higher single master pathway incorporating all the pathways. All the genes studied showed spatio-temporal variation in expression that must be validated by protein expression analysis and work may be replicated in human based model or single cells from mice.

Acknowledgements

We would like to thank Higher Education Commission, Pakistan for funding this work which was supported by NRPU grant no. 15544. Furthermore, we want to thank all members of our lab for their input in form of work support and team work.

Complementary interests

Authors declare no competing interests.

References

1. Thornton GK, Woods CG. Primary microcephaly : do all roads lead to Rome ? 2009;501–10.

2. Khetarpal P, Das S, Panigrahi I, Munshi A. Primordial dwarfism: overview of clinical and genetic aspects. Mol Genet Genomics. 2016;291(1):1–15.

3. Asif M, Abdullah U, Nürnberg P, Tinschert S, Hussain MS. Congenital Microcephaly: A Debate on Diagnostic Challenges and Etiological Paradigm of the Shift from Isolated/Non-Syndromic to Syndromic Microcephaly. Cells. 2023;12(4).

4. Zaqout S, Kaindl AM. Autosomal Recessive Primary Microcephaly: Not Just a Small Brain. Front Cell Dev Biol. 2022;9(January).

5. Nano M, Basto R. Consequences of centrosome dysfunction during brain development. Adv Exp Med Biol. 2017;1002:19–45.

6. Cuccurullo C, Miele G, Piccolo G, Bilo L, Accogli A, D'Amico A, et al. Hydranencephaly in CENPJ-related Seckel syndrome. Eur J Med Genet [Internet]. 2022;65(12):104659. Available from: https://doi.org/10.1016/j.ejmg.2022.104659

7. Kaindl AM, Passemard S, Kumar P, Kraemer N, Issa L, Zwirner A, et al. Progress in Neurobiology Many roads lead to primary autosomal recessive microcephaly. Prog Neurobiol [Internet]. 2010;90(3):363–83. Available from: http://dx.doi.org/10.1016/j.pneurobio.2009.11.002

8. Sir JH, Barr AR, Nicholas AK, Carvalho OP, Khurshid M, Sossick A, et al. A primary microcephaly protein complex forms a ring around parental centrioles. Nat Genet. 2011;

9. Nicholas AK, Khurshid M, Désir J, Carvalho OP, Cox JJ, Thornton G, et al. WDR62 is associated with the spindle pole and is mutated in human microcephaly. Nat Genet. 2010;

10. Bakircioglu M, Carvalho OP, Khurshid M, Cox JJ, Tuysuz B, Barak T, et al. The Essential Role of Centrosomal NDE1 in Human Cerebral Cortex Neurogenesis. Am J Hum Genet [Internet]. 2011 May 13 [cited 2017 Dec 15];88(5):523–35. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21529752

11. Castellanos MDP, Wickramasinghe CD, Betrán E. The roles of gene duplications in the dynamics of evolutionary conflicts. Proc R Soc B Biol Sci. 2024;291(2024).

12. Perlman RL. Mouse Models of Human Disease: An Evolutionary Perspective. Evol Med Public Heal. 2016;eow014.

13. Morris-Rosendahl DJ, Kaindl AM. What next-generation sequencing (NGS) technology has enabled us to learn about primary autosomal recessive microcephaly (MCPH). Mol Cell Probes [Internet]. 2015;29(5):271–81. Available from: http://dx.doi.org/10.1016/j.mcp.2015.05.015

14. Farcy S, Hachour H, Bahi-Buisson N, Passemard S. Genetic Primary Microcephalies: When Centrosome Dysfunction Dictates Brain and Body Size. Cells. 2023;12(13):1–34.

15. Jayaraman D, Kodani A, Gonzalez DM, Mancias JD, Mochida GH, Vagnoni C, et al. Microcephaly Proteins Wdr62 and Aspm Define a Mother Centriole Complex Regulating Centriole Biogenesis, Apical Complex, and Cell Fate. Neuron [Internet]. 2016;92(4):813–28. Available from: http://dx.doi.org/10.1016/j.neuron.2016.09.056

16. Hettige NC, Ernst C. FOXG1 Dose in Brain Development. Front Pediatr. 2019;7(November):1–12.

17. Duan L, Yin H, Liu J, Wang W, Li J, Shen J, et al. Fetal neurodevelopmental spatio-temporal dynamic transcriptional landscape of maternal insult-induce autism spectrum disorder risk. Biochem Biophys Res Commun [Internet]. 2022;614:183–90. Available from: https://doi.org/10.1016/j.bbrc.2022.05.040

18. Awad S, Al-Dosari MS, Al-Yacoub N, Colak D, Salih MA, Alkuraya FS, et al. Mutation in PHC1 implicates chromatin remodeling in primary microcephaly pathogenesis. Hum Mol Genet. 2013;22(11):2200–13.

19. Tonkin E, Hagan DM, Li W, Strachan T. Identification and characterisation of novel mammalian homologues of Drosophila polyhomeotic permits new insights into relationships between members of the polyhomeotic family. Hum Genet. 2002;111(4–5):435–42.

20. Kraemer N, Issa-Jahns L, Neubert G, Ravindran E, Mani S, Ninnemann O, et al. Novel alternative splice variants of mouse CDK5RAP2. PLoS One. 2015;

21. Mani A. PDE4DIP in health and diseases. Cell Signal. 2022;94:1–13.

22. Zhao H, Zhu L, Zhu Y, Cao J, Li S, Huang Q, et al. The cep63 paralogue deup1 enables massive de novo centriole biogenesis for vertebrate multiciliogenesis. Nat Cell Biol [Internet]. 2013;15(12):1434–44. Available from: http://dx.doi.org/10.1038/ncb2880

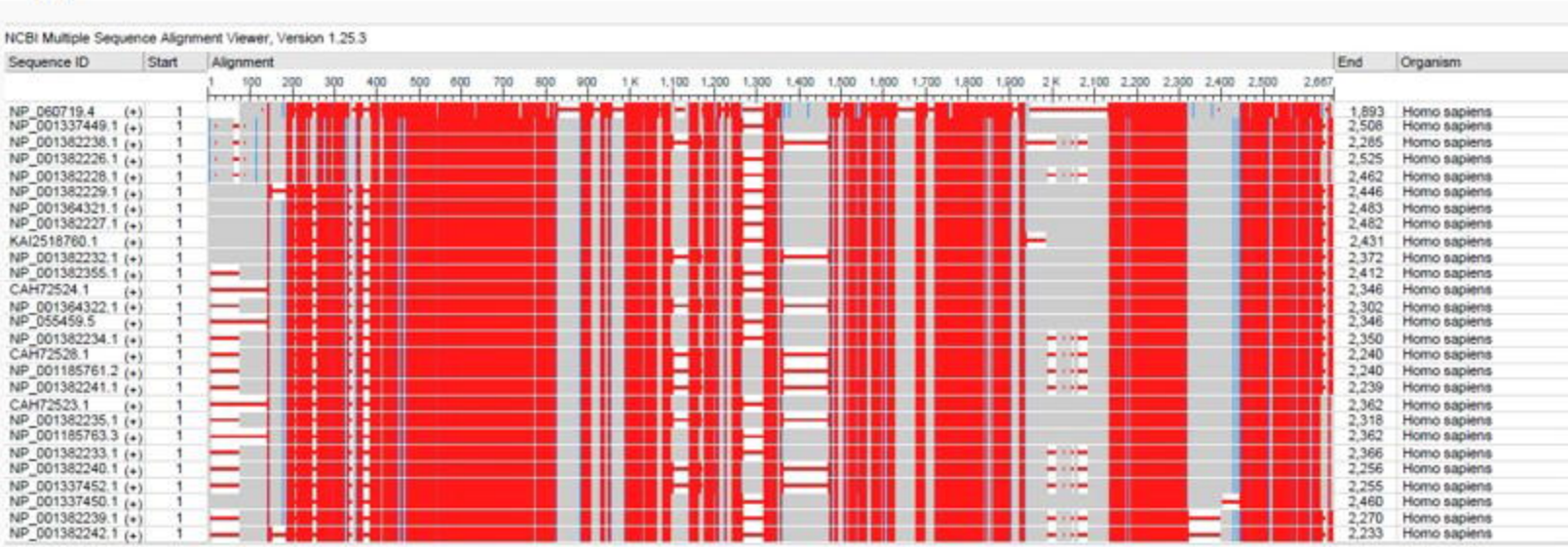
23. Ma L, Quigley I, Omran H, Kintner C. Multicilin drives centriole biogenesis via E2f proteins. Genes Dev. 2014;28(13):1461–71.

24. Byers SL, Wiles M V., Dunn SL, Taft RA. Mouse estrous cycle identification tool and images. PLoS One. 2012;7(4):1–5.

25. Vaid S, Huttner WB. Transcriptional regulators and human-specific/primate-specific genes in neocortical neurogenesis. Int J Mol Sci. 2020;21(13):1–19.

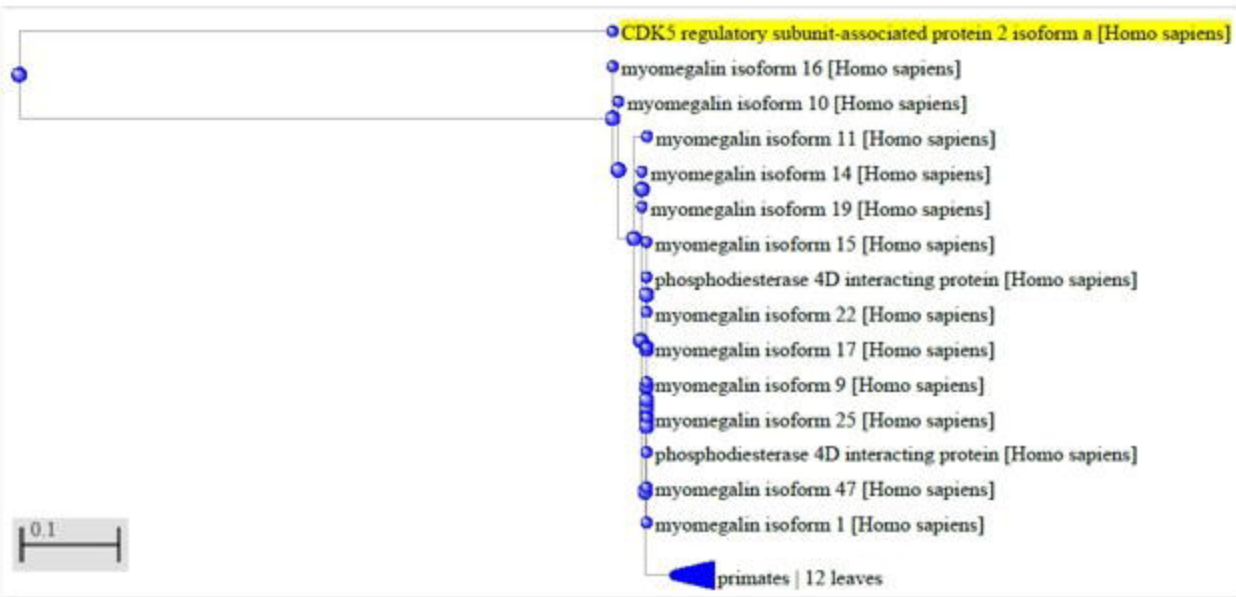
CDK5RAP2 and PDE4DIP

A

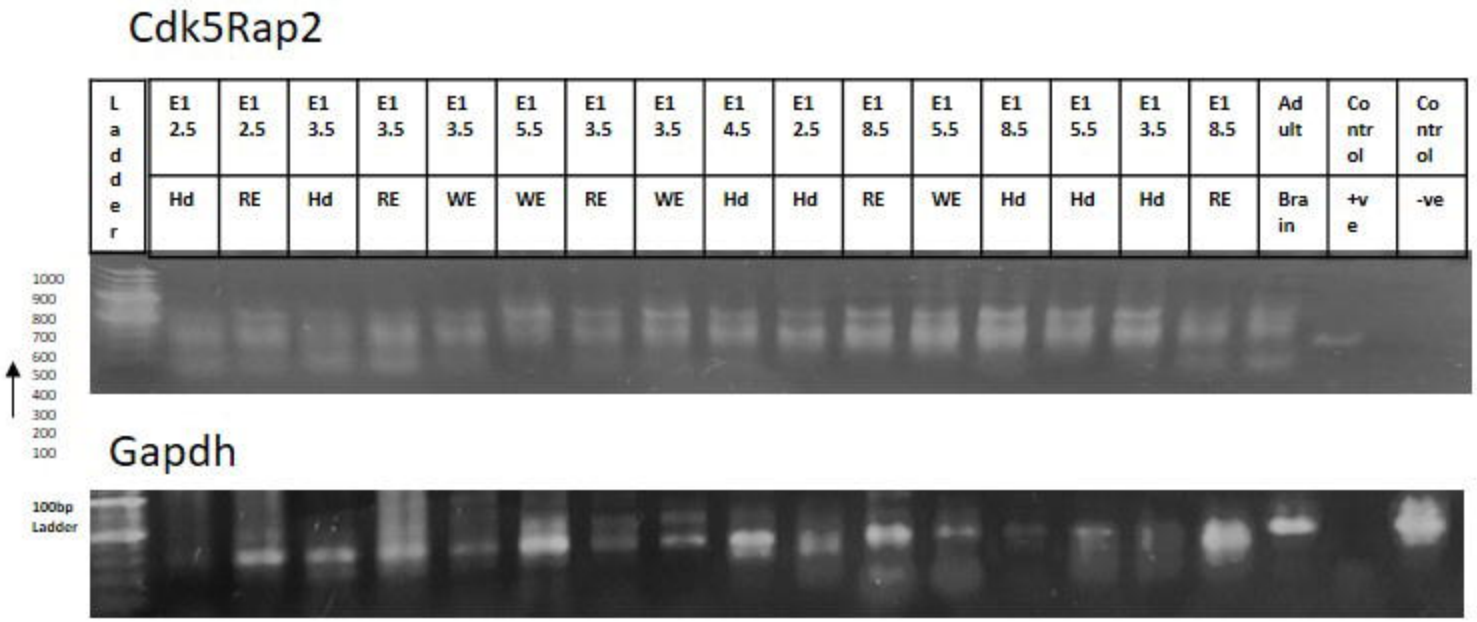


B

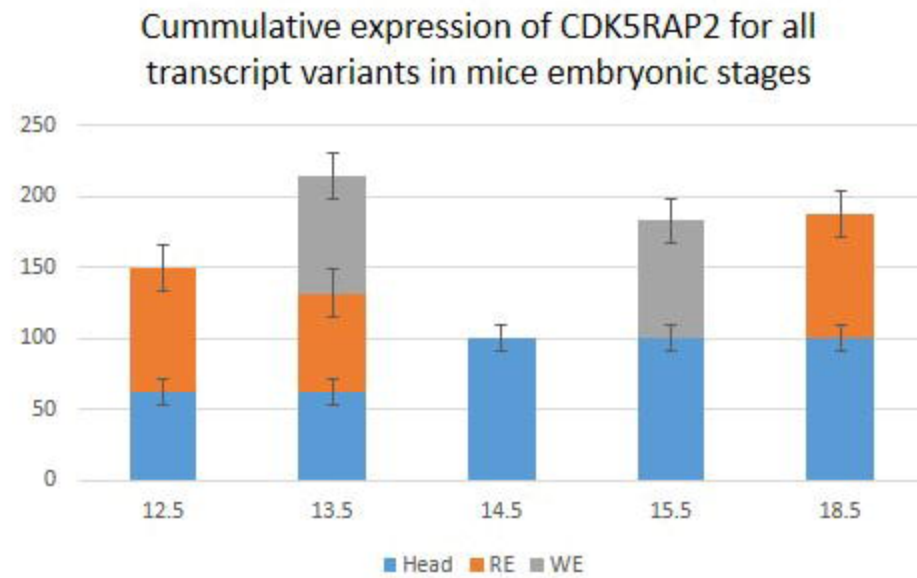
C



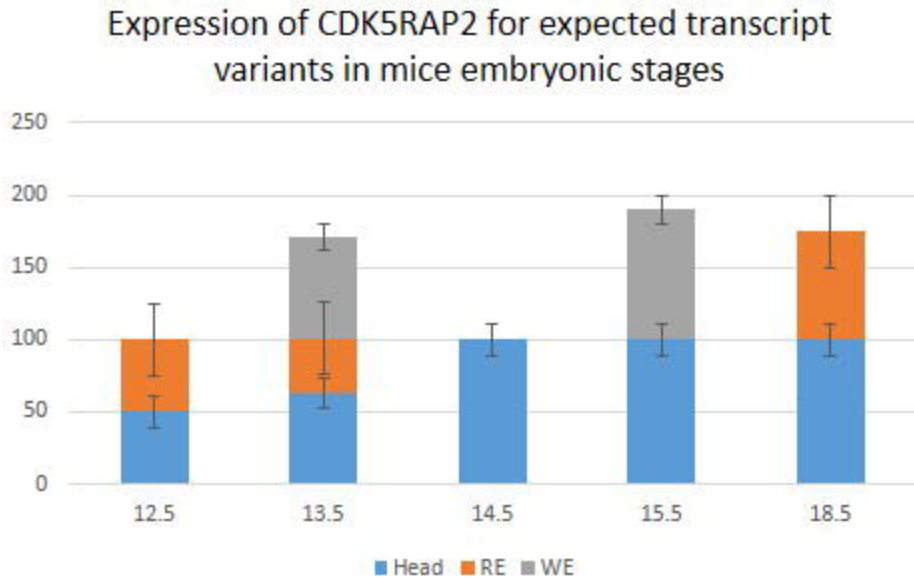
A


Cdk5Rap2
Ladder
E1 2.5 Hd
E1 2.5 RE
E1 3.5 Hd
E1 3.5 RE
E1 3.5 WE
E1 5.5 WE
E1 3.5 RE
E1 3.5 WE
E1 4.5 Hd
E1 2.5 Hd
E1 8.5 RE
E1 5.5 WE
E1 8.5 Hd
E1 5.5 Hd
E1 3.5 Hd
E1 8.5 RE
Adult Brain
Control +ve
Control -ve
1000
900
800
700
600
500
400
300
200
100
Gapdh
100bp Ladder

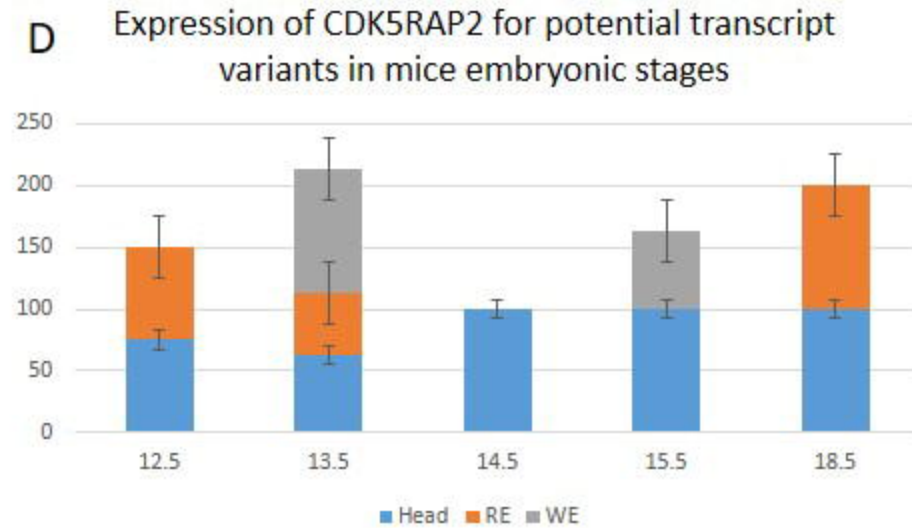
B


Cummulative expression of CDK5RAP2 for all transcript variants in mice embryonic stages
250
200
150
100
50
0
12.5
13.5
14.5
15.5
18.5
Head
RE
WE

C

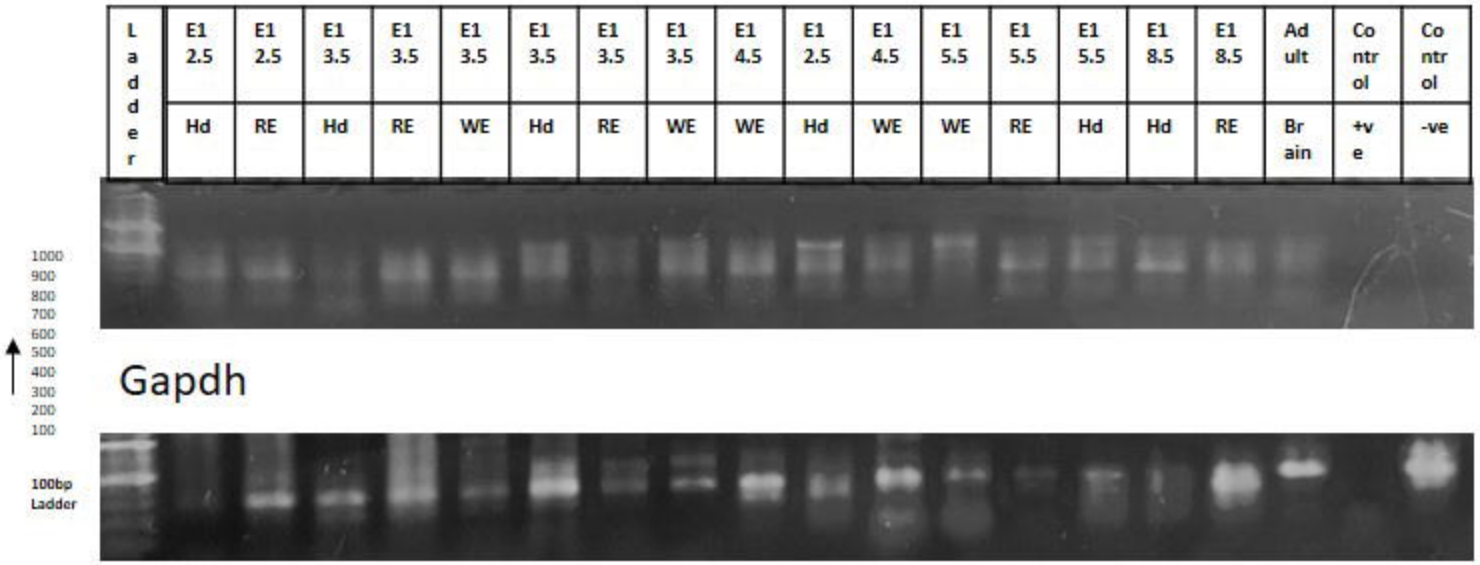

Expression of CDK5RAP2 for expected transcript variants in mice embryonic stages
250
200
150
100
50
0
12.5
13.5
14.5
15.5
18.5
Head
RE
WE

D

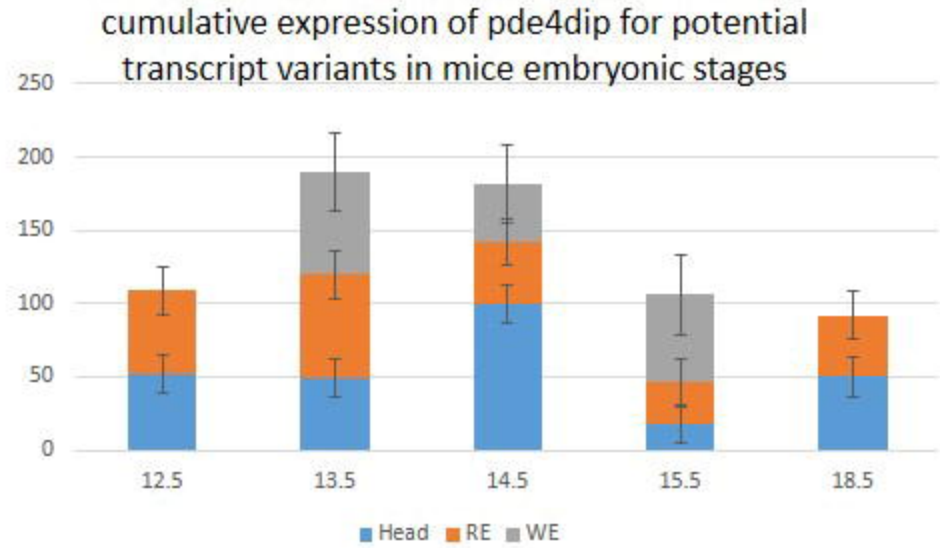

Expression of CDK5RAP2 for potential transcript variants in mice embryonic stages
250
200
150
100
50
0
12.5
13.5
14.5
15.5
18.5
Head
RE
WE

A

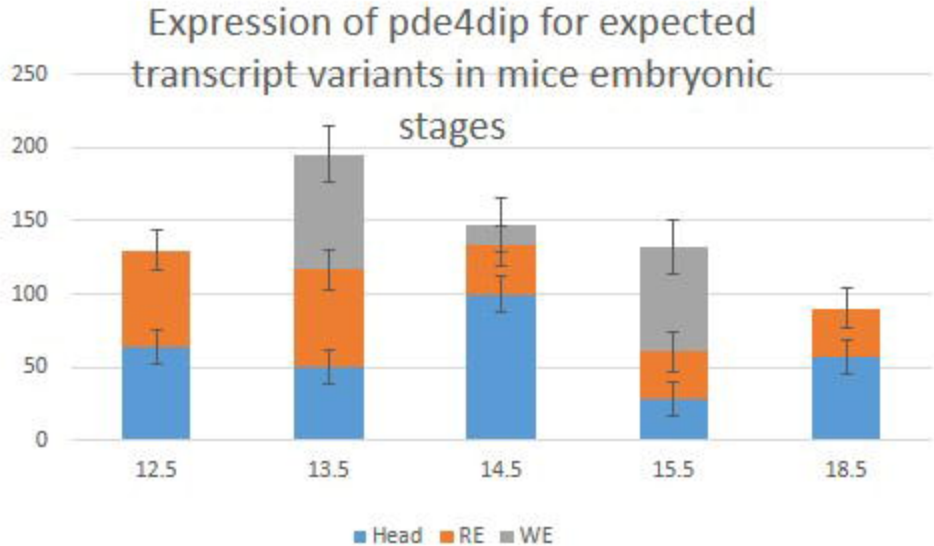
Pde4Dip



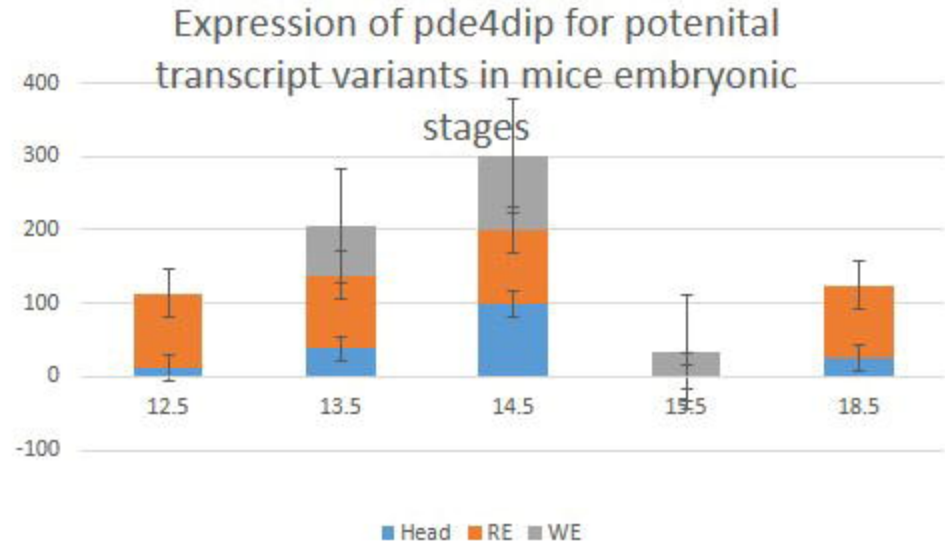
B



C

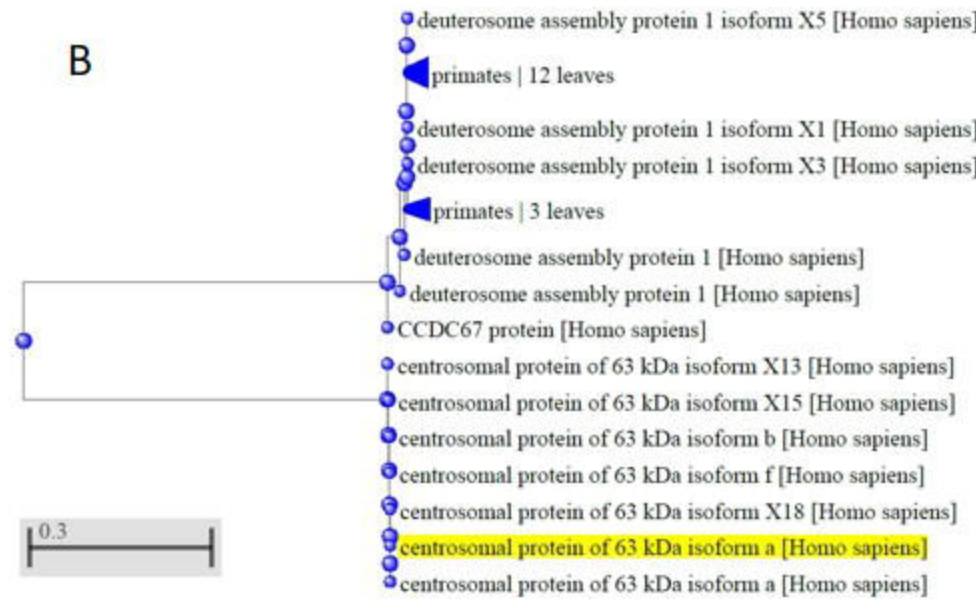


D



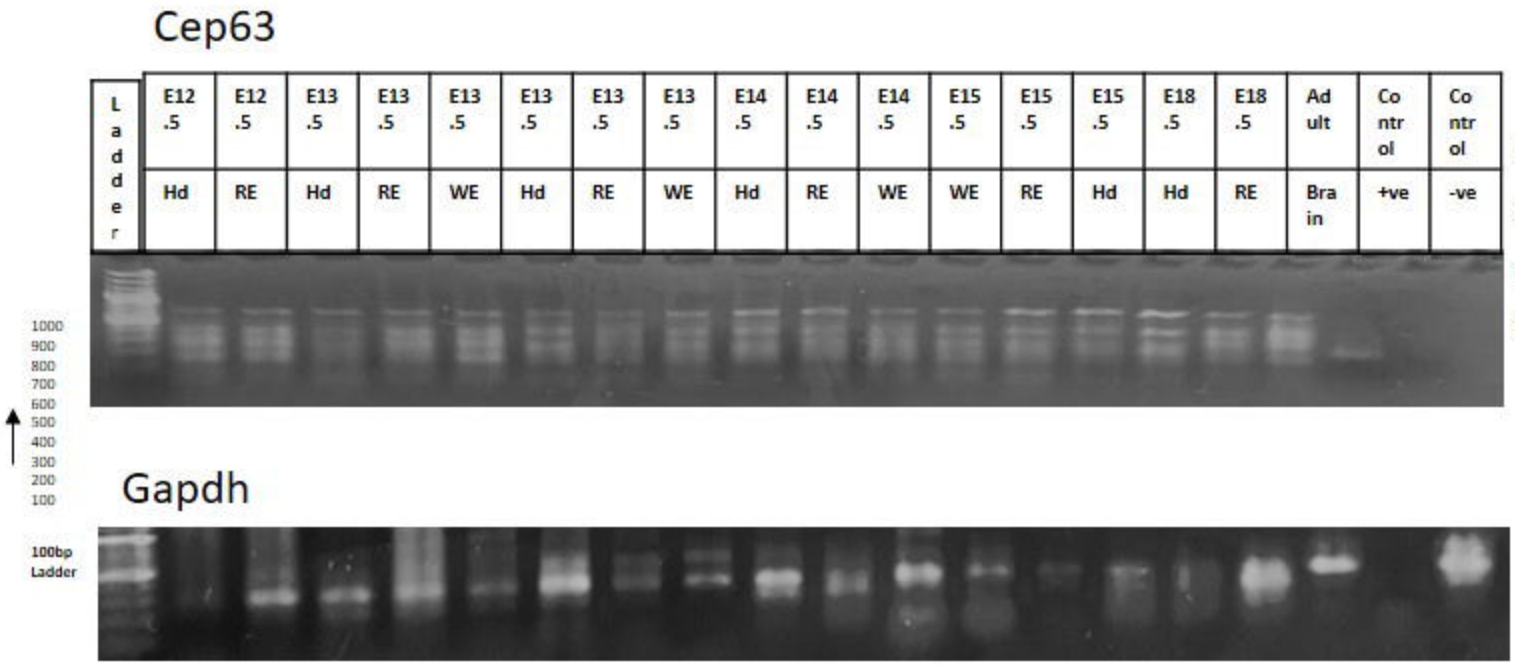
A

B

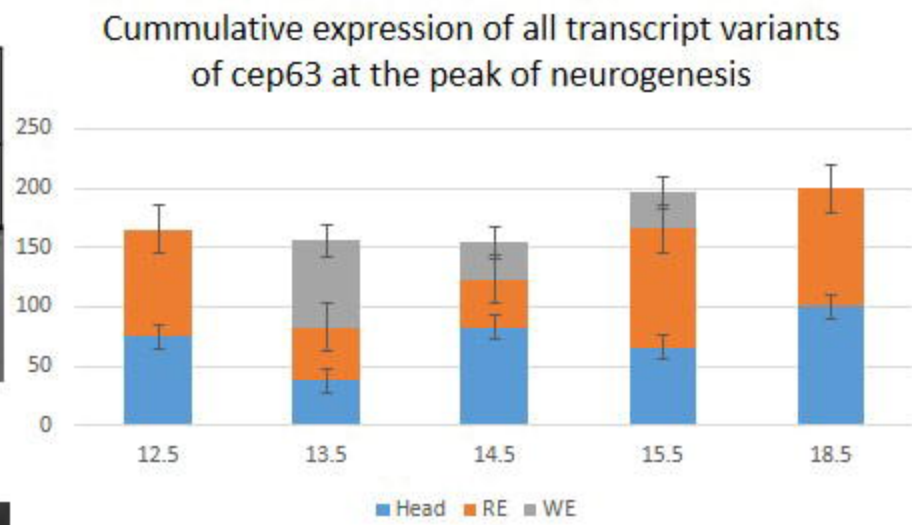


CEP63 and DEUP1

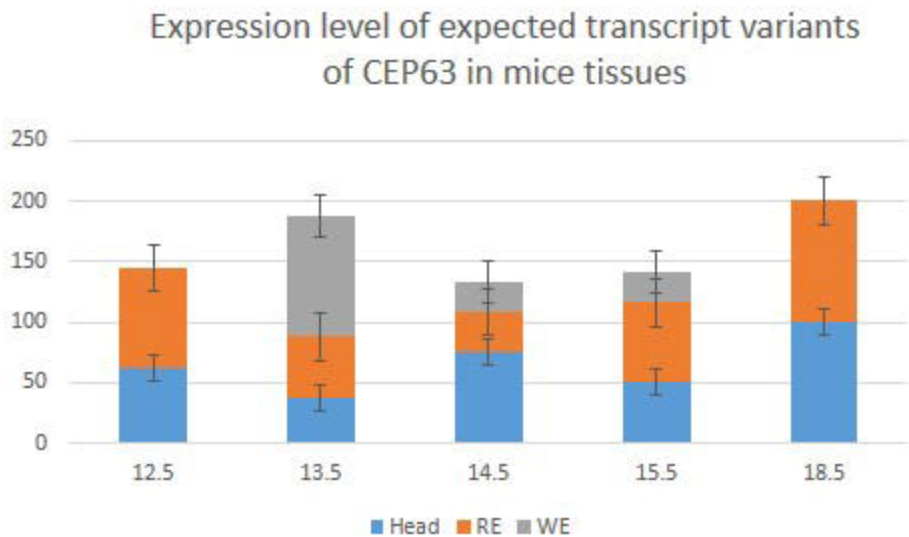
A


Cep63
Ladder
E12.5 Hd
E12.5 RE
E13.5 Hd
E13.5 RE
E13.5 WE
E13.5 Hd
E13.5 RE
E13.5 WE
E14.5 Hd
E14.5 RE
E14.5 WE
E15.5 WE
E15.5 RE
E15.5 Hd
E18.5 Hd
E18.5 RE
Adult Brain
Control +ve
Control -ve
1000
900
800
700
600
500
400
300
200
100
Gapdh
100bp Ladder

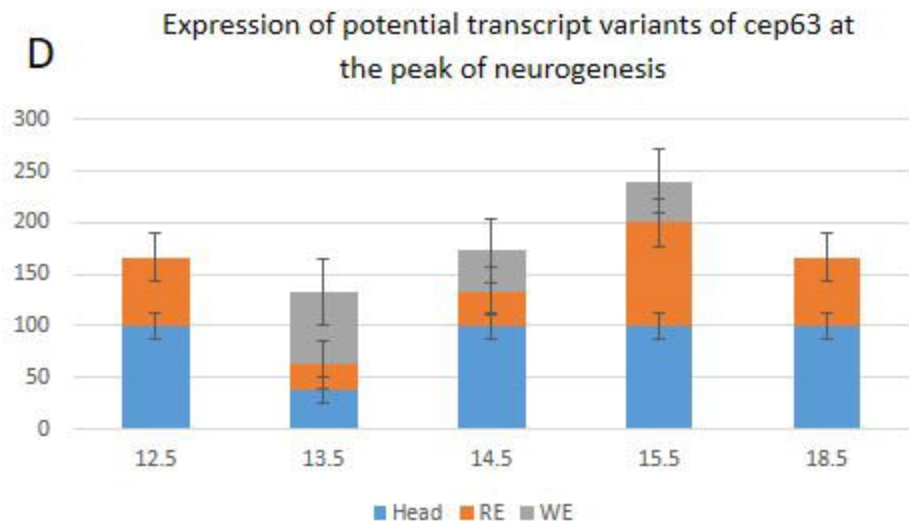
B


Cummulative expression of all transcript variants of cep63 at the peak of neurogenesis
250
200
150
100
50
0
12.5
13.5
14.5
15.5
18.5
Head
RE
WE

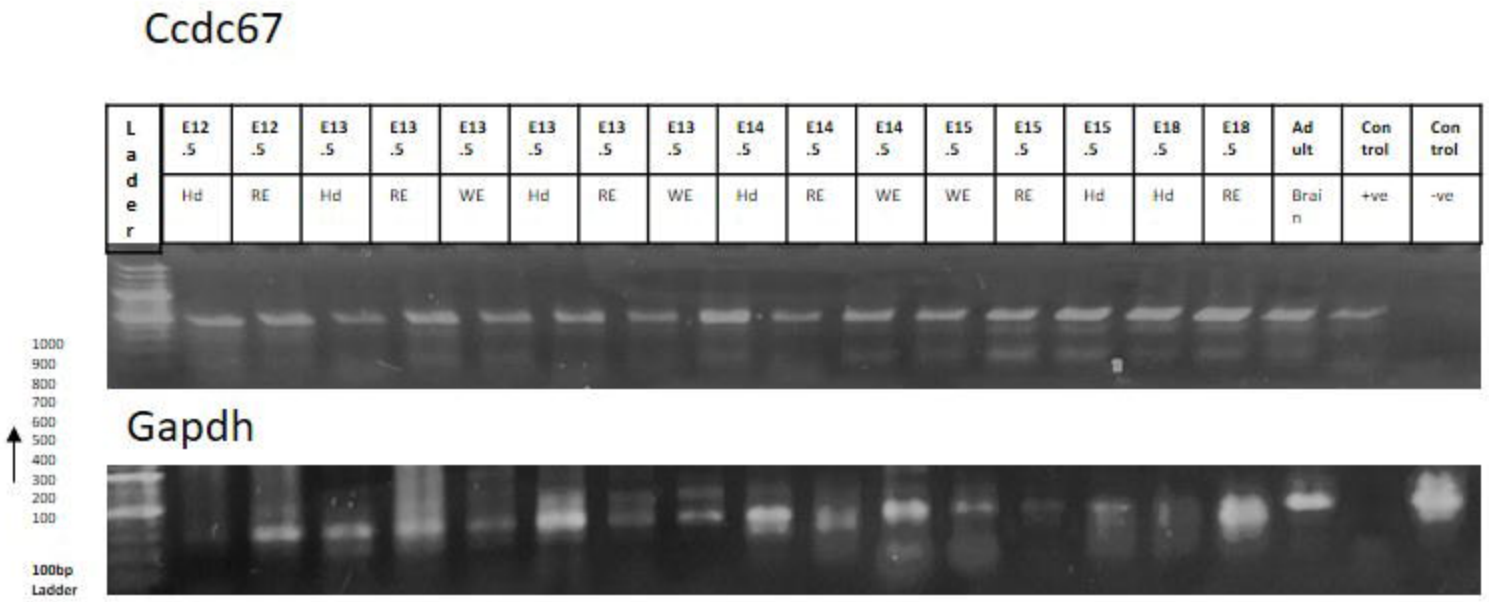
C


Expression level of expected transcript variants of CEP63 in mice tissues
250
200
150
100
50
0
12.5
13.5
14.5
15.5
18.5
Head
RE
WE

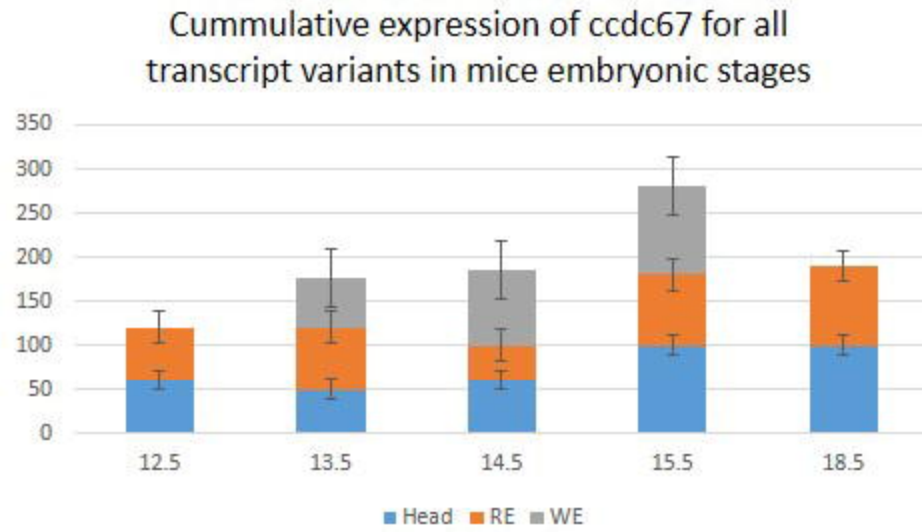
D


Expression of potential transcript variants of cep63 at the peak of neurogenesis
300
250
200
150
100
50
0
12.5
13.5
14.5
15.5
18.5
Head
RE
WE

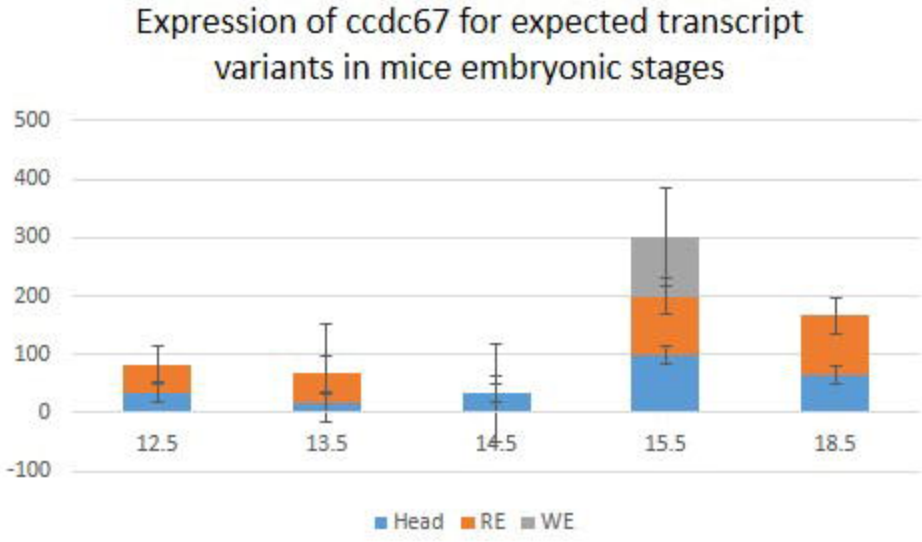
A


Ccdc67
Ladder
E12.5 Hd
E12.5 RE
E13.5 Hd
E13.5 RE
E13.5 WE
E13.5 Hd
E13.5 RE
E13.5 WE
E14.5 Hd
E14.5 RE
E14.5 WE
E15.5 WE
E15.5 RE
E15.5 Hd
E18.5 Hd
E18.5 RE
Adult Brain
Control +ve
Control -ve
1000
900
800
700
600
500
400
300
200
100
100bp Ladder
Gapdh

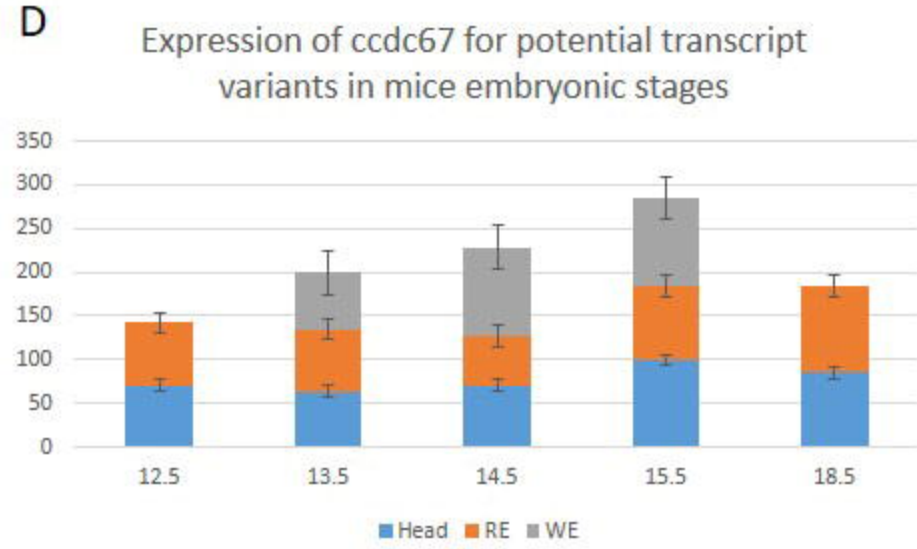
B


Cummulative expression of ccdc67 for all transcript variants in mice embryonic stages
350
300
250
200
150
100
50
0
12.5
13.5
14.5
15.5
18.5
Head
RE
WE

C


Expression of ccdc67 for expected transcript variants in mice embryonic stages
500
400
300
200
100
0
-100
12.5
13.5
14.5
15.5
18.5
Head
RE
WE

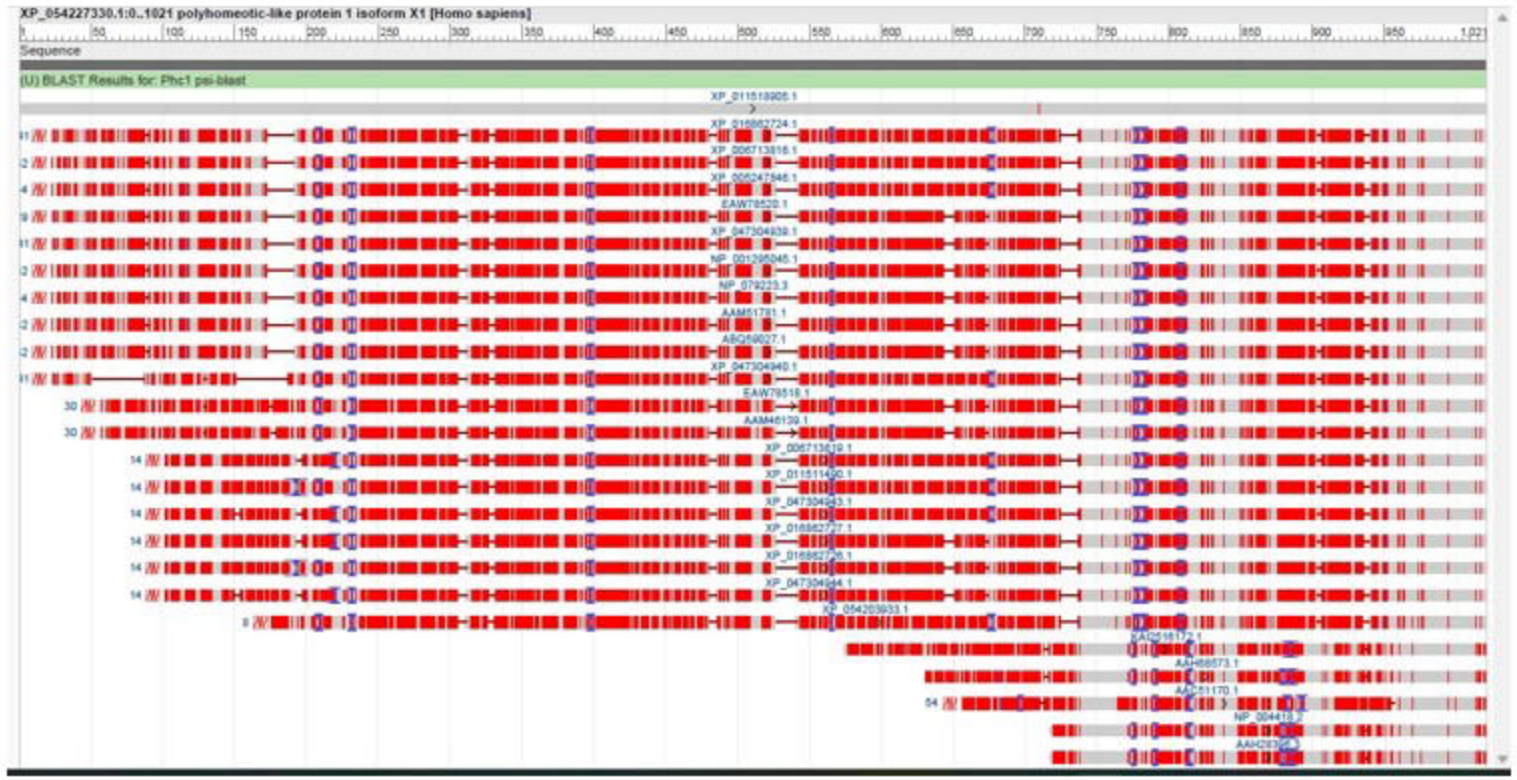
D


Expression of ccdc67 for potential transcript variants in mice embryonic stages
350
300
250
200
150
100
50
0
12.5
13.5
14.5
15.5
18.5
Head
RE
WE

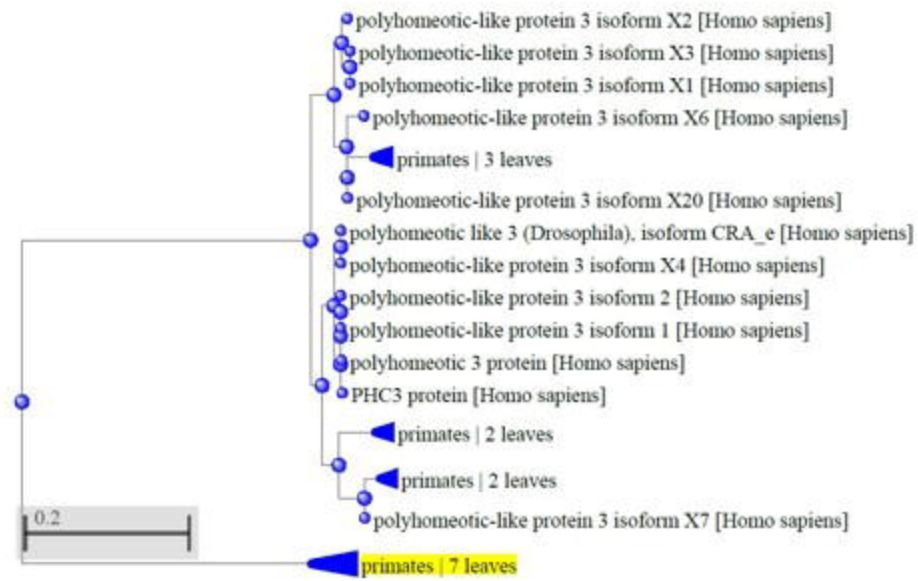
PHC1 and PHC3

A

B



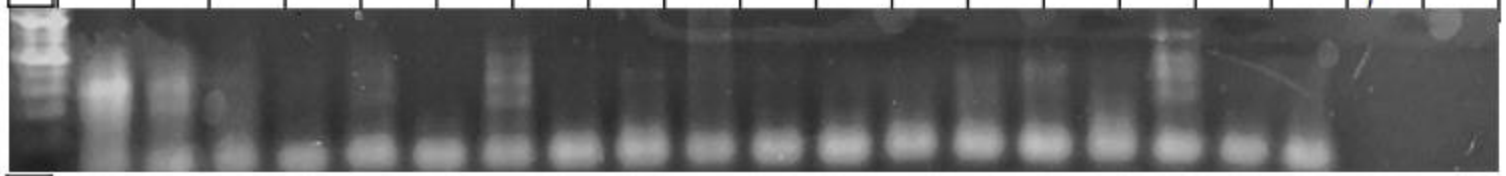
C



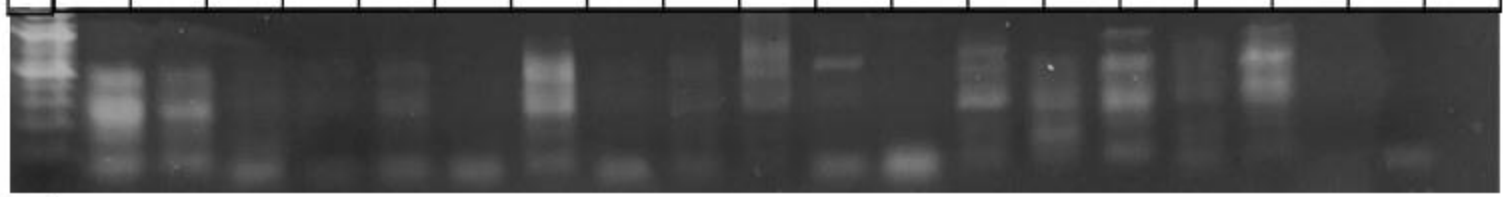
A

Phc1

Ladder	E12.5	E13.5	E13.5	Adult	E18.5	Adult	E18.5	E14.5	E14.5	E14.5	E15.5	E15.5	E13.5	E12.5	E12.5	E13.5	Adult	control	control
	Hd	RE	Hd	Heart	RE	Brain	Hd	RE	WE	Hd	WE	RE	WE	Hd	RE	Hd	Brain	+ve (Gapdh)	-ve



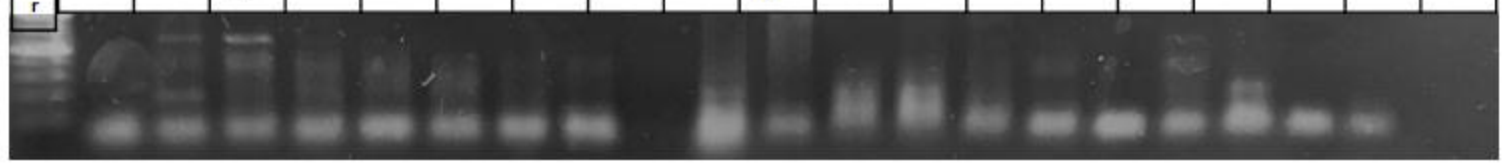
Ladder	E12.5	E13.5	E13.5	Adult	E18.5	Adult	E18.5	E14.5	E14.5	E14.5	E15.5	E15.5	E13.5	E12.5	E12.5	E13.5	Adult	Control	Control
	Hd	RE	Hd	Heart	RE	Brain	Hd	RE	WE	Hd	WE	RE	WE	RE	Hd	Hd	Brain	+ve	-ve



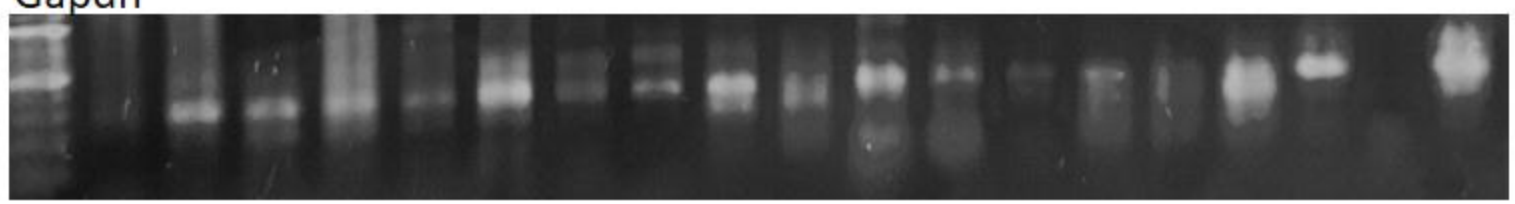
1000
900
800
700
600
500
400
300
200
100

100bp Ladder

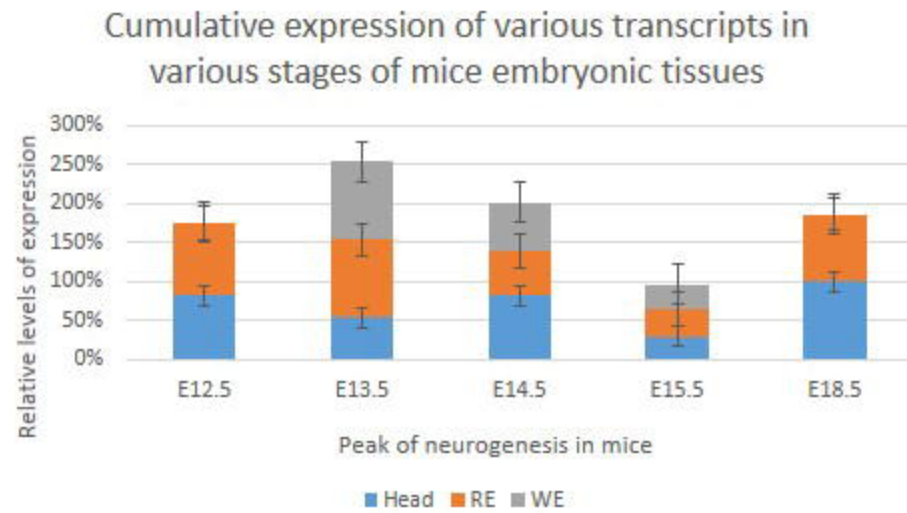
Ladder	E12.5	E12.5	Adult	E18.5	E13.5	E13.5	E13.5	E13.5	E13.5	Adult	E14.5	E15.5	E15.5	E14.5	E14.5	E15.5	E14.5	E18.5	Control
	Hd	RE	Heart	RE	Hd	WE	WE	RE	Hd	Brain	RE	RE	Hd	Hd	WE	WE	RE	Hd	-ve



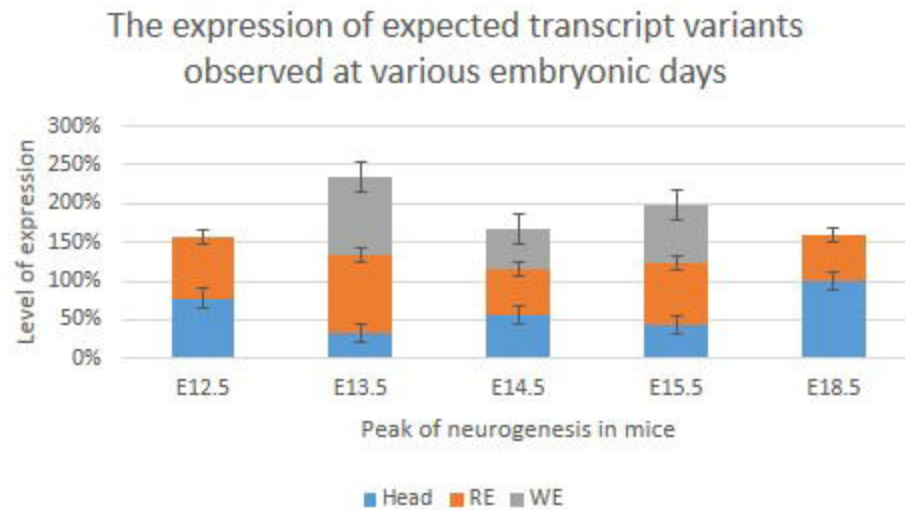
Gapdh



B

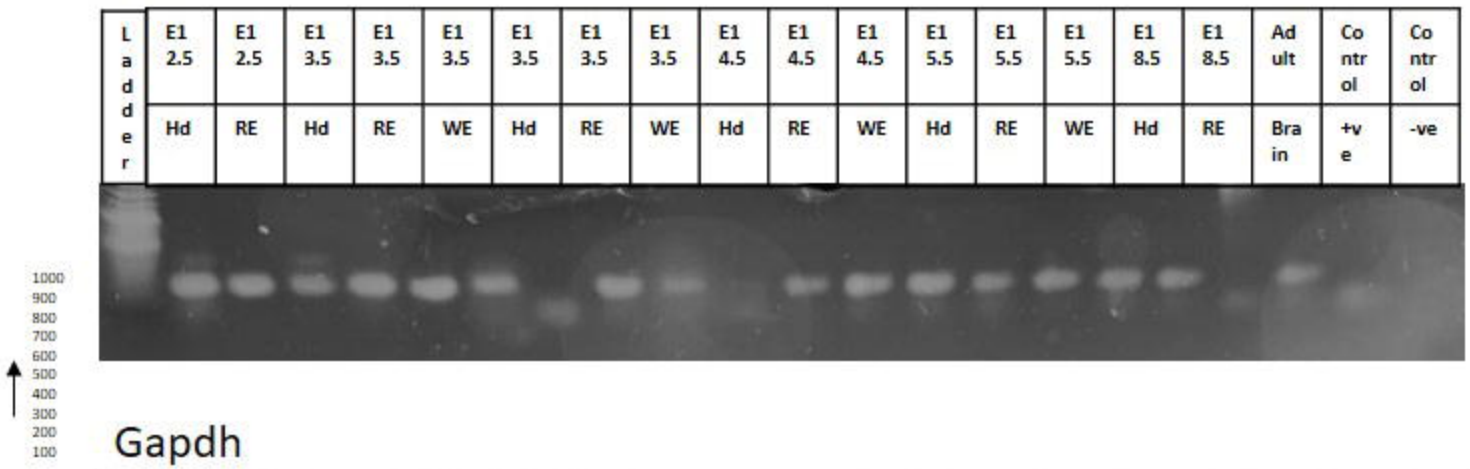


C

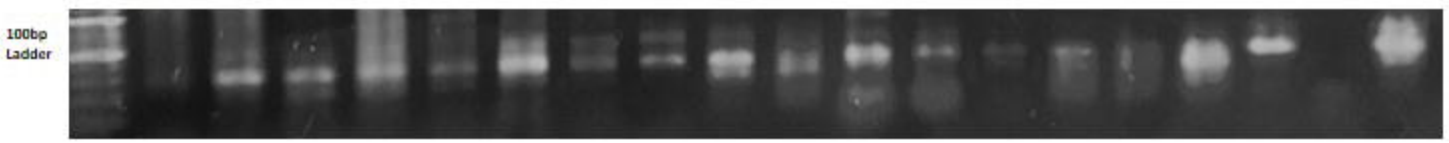


A

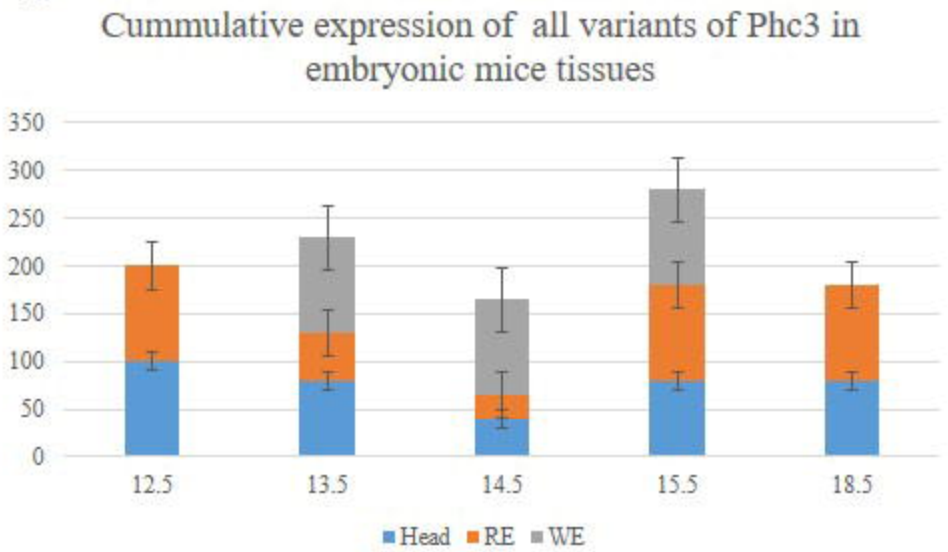
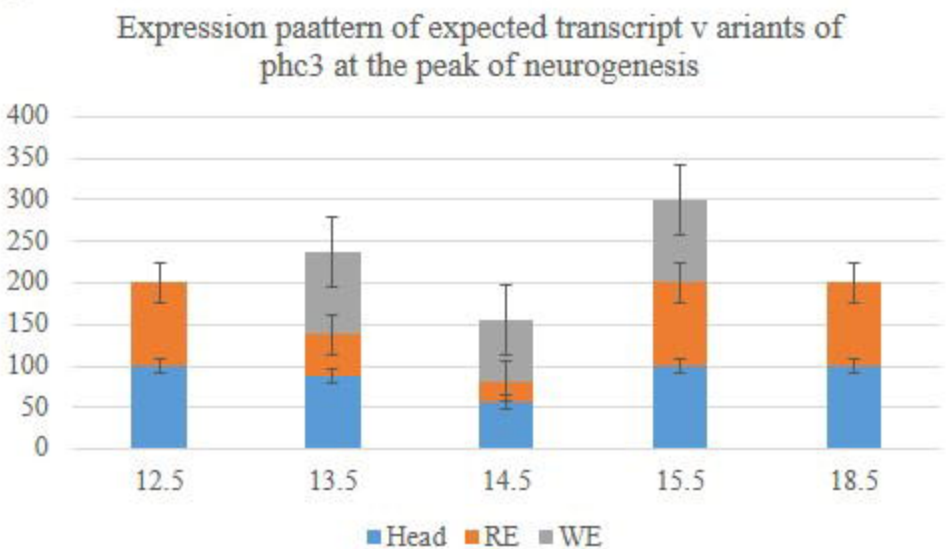
Phc3



Gapdh



B



C