Original Article

Whole-exome sequencing analysis to identify novel potential pathogenetic mutations in fetuses with abnormal brain structure

Lili Shi[1,2#], Ming Li[3#], Hong Qi[4], Jianjiang Zhu[4], Jing Yang[5], Jie Tang[6], Longxia Wang[6]

[1]Medical School of Chinese PLA, Beijing, China; [2]Department of Ultrasound Medicine, Beijing Chaoyang Hospital, Capital Medical University, Beijing, China; [3]Department of Laboratory, Tianyou Hospital Affiliated to Wuhan University of Science & Technology, Wuhan, China; [4]Department of Prenatal Diagnosis Center, Haidian Maternal and Child Health Hospital, Beijing, China; [5]Department of Obstetrics & Gynecology, Peking University Third Hospital, Beijing, China; [6]Department of Ultrasound, First Medical Center of Chinese PLA General Hospital, Medical School of Chinese PLA, Beijing, China

Contributions: (I) Conception and design: L Shi, M Li; (II) Administrative support: J Tang, L Wang; (III) Provision of study materials or patients: L Shi, J Zhu; (IV) Collection and assembly of data: L Shi, J Yang; (V) Data analysis and interpretation: L Shi, H Qi; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this study.

Correspondence to: Jie Tang. Department of Ultrasound, First Medical Center of Chinese PLA General Hospital, Medical School of Chinese PLA, Beijing 100853, China. Email: txiner@vip.sina.com; Longxia Wang. Department of Ultrasound, First Medical Center of Chinese PLA General Hospital, Medical School of Chinese PLA, Beijing 100853, China. Email: 13693685342@163.com.

Background: Genetic mutations in genes related to the production, migration, or differentiation of cortical neurons can result in malformations of cortical development (MCDs). However, a large number of MCD-related pathogenetic mutations remain unknown. This study aimed to investigate the genetic cause of MCDs and to identify the new MCD-associated mutations through whole - exome sequencing (WES) in fetuses with abnormal brain structure.

Methods: Cord venous blood samples were collected from 11 fetuses with MCDs. Whole-genome DNA was extracted from the blood, and WES was performed. Single nucleotide substitutions, insertions, and deletions were detected by bioinformatics analysis. Genetic mutations in genes associated with MCD were identified.

Results: A total of 1035 genes with high-impact genetic variants in at least 1 fetus were identified. The results of gene ontology enrichment analysis were consistent with those of previous studies and also indicated new potential MCD-related pathogenetic genetic mutations. Genes with high-impact mutations in multiple fetuses, such as *CTDSP2* and C-terminal binding protein 2 (*CTBP2*), were more likely to be the genes affecting normal brain development.

Conclusions: This study has characterized variations in fetuses with MCDs and identified potential genetic mutations causing MCDs. Our findings extend the mutation spectrum of MCDs and provide a promising source for the identification of MCD-related pathogenetic mutations.

Keywords: Malformations of cortical development (MCDs); whole-exome sequencing (WES); fetus, mutations; *CTDSP2*; C-te rminal binding protein 2 (*CTBP2*)

Submitted Mar 03, 2021. Accepted for publication May 08, 2021.
doi: 10.21037/atm-21-1477
View this article at: http://dx.doi.org/10.21037/atm-21-1477

Introduction

Malformations of cortical development (MCDs) encompass heterogeneous groups of structural brain anomalies which commonly cause neurodevelopmental delay and epilepsy. MCDs have been a topic of interest to clinicians and neuroscientists for many decades, as the study of these anomalies considerably advances the understanding of normal brain development and its perturbations. The

© Annals of Translational Medicine. All rights reserved.

majority of MCDs are believed to be attributable to underlying genetic mutations which disturb the proteins and associated signaling pathways involved in the development of the cerebral cortex (1).

Recent progress in the exploration of the genetic basis of brain malformation has been driven by advancements in high-throughput DNA sequencing technology. Mutations in different genes cause different MCDs including microcephaly (*WDR62* and *CDK5RAP2*) (2,3), megalencephaly (*AKT3* and *PIK3CA*) (4,5), and lissencephaly (*LIS1* and *TUBA1A*) (6,7). Moreover, mutations in genes involved in the RTK-PI3K-AKT-mTOR pathway are known to be associated with microlissencephaly, hemimegalencephaly, and focal cortical dysplasia. Although numerous genetic causes of MCDs have been identified, other genes potentially involved in cortical development and the pathogenesis of MCDs have yet to be identified.

In the present investigation, we attempted to identify the genetic mutations associated with MCDs by whole-exome sequencing (WES) in 11 fetuses with abnormal brain structure. Dozens of candidate genes were identified as being potentially involved in cortical development. Gene ontology (GO) enrichment analysis revealed that these genes were involved in the synapse, spindle pole, centrosome, and microtubule, which was consistent with the findings of previous studies. Further, we focused on genes with mutations within multiple fetuses and predicted the protein structure variations when these mutations occurred. Our findings provide new candidate genes that are involved in cortical development and are associated with neurodevelopmental disorders.

We present the following article in accordance with the STREGA reporting checklist (available at http://dx.doi.org/10.21037/atm-21-1477).

Methods

Sample collection

Villous tissue (10–15 mg) from the placenta of pregnant women with a gestational age of 11–13^{+6} weeks was obtained via abdominal puncture using an 18G puncture needle (PE18/15, Italian Gallini company) under ultrasound guidance. After a gestational age of >24 weeks had been reached, 2 mL of cord venous blood was obtained via percutaneous cord blood puncture through the abdomen of each woman using a 22G puncture needle (model: B22G, Japan Hakko company) under ultrasound guidance. After the extraction of whole-genome DNA from the cord venous blood, WES was performed. This study was approved by the ethical committee of Beijing Haidian maternal and Child Health Hospital and was performed in accordance with the Helsinki Declaration (as revised in 2013). Individual consent for this retrospective analysis was waived.

Next-generation sequencing and data analysis

Sequencing was performed on an Illumina Novaseq sequencer, with a library size of 200 bp. The adaptors and low-quality reads were removed with fastp (v0.12.3, http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/) under default parameters. The remaining sequencing reads were mapped to the human reference genome (hg19, http://hgdownload.cse.ucsc.edu/goldenPath/hg19/bigZips/hg19.fa.gz) with Burrows-Wheeler Aligner (BWA, v0.7.10, http://bio-bwa.sourceforge.net) using the MEM alignment mode under default parameters. Sambamba (v0.6.8, http://lomereiter.github.io/sambamba/) was used to remove the polymerase chain reaction (PCR) duplicates.

Copy number variation identification

Each chromosome was separated into 10-kb windows, and the sequencing coverage of each 10-kb region was estimated. For reasonable comparison of the sequencing depth among different chromosomes, the base coverage of each 10-kb window was normalized by dividing it by the median coverage for the corresponding chromosome. The copy number was taken as the normalized base coverage estimated for each 10-kb region. The normalized base coverage of most genomic regions was around 1, representing 2 copies, and those of the X and Y chromosomes in male fetuses were around 0.5, representing 1 copy.

Identification of single nucleotide substitutions, insertions, and deletions

The HaplotypeCaller module in GATK (v3.8, https://github.com/broadgsa/gatk/releases) was used to detect single nucleotide substitutions (SNSs), and insertions and deletions (inDels) in each fetus under default parameters. Next, both variant types were filtered against the information in 1KGP and dbSNP130. Variants located in intergenic or intronic regions were subsequently discarded. SnpEff (v4.3t, http://snpeff.sourceforge.net/SnpEff.html#intro) was used to predict the effects of the remaining variants

© Annals of Translational Medicine. All rights reserved.

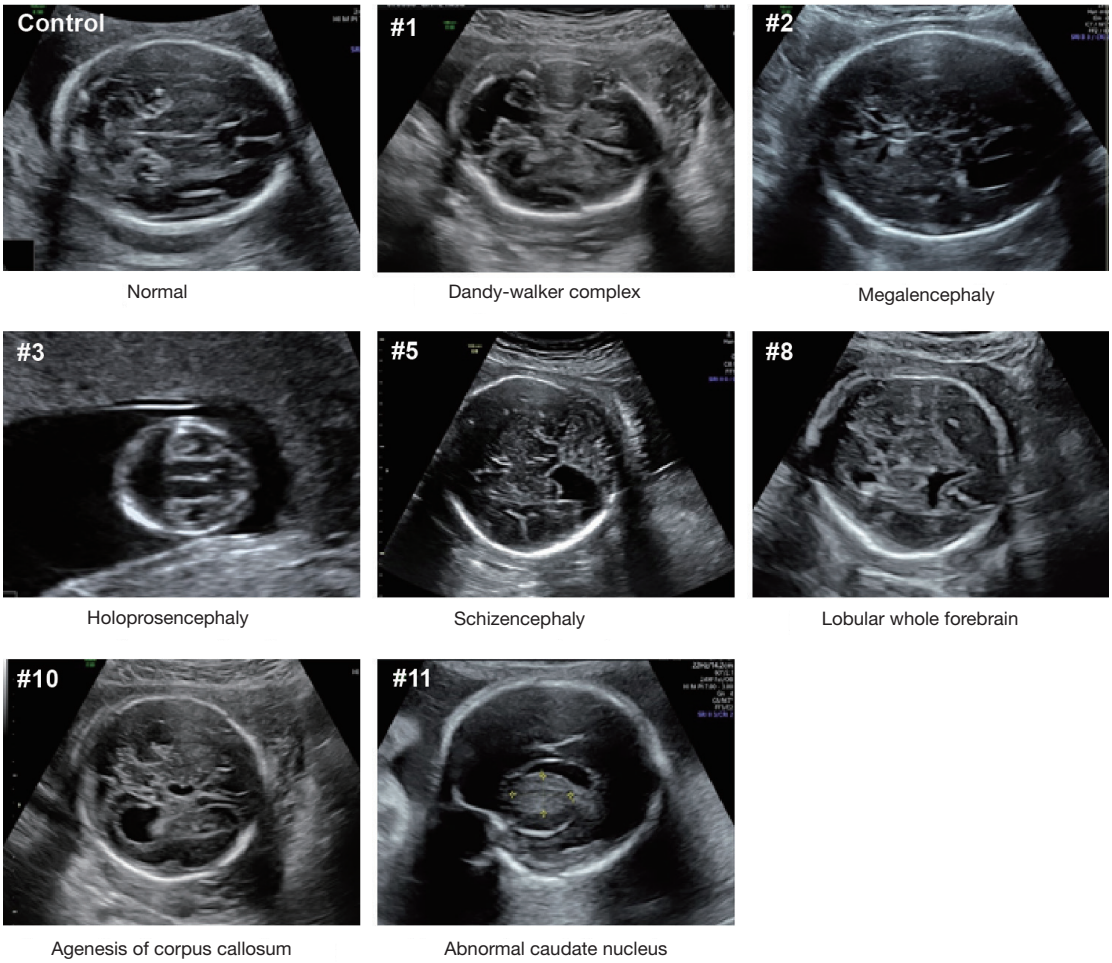


Figure 1 Phenotypes of fetuses with abnormal brain structure. One normal fetus and 7 fetuses with typical brain structural abnormalities are shown. The type of abnormality and the sample number are also shown.

on genes (such as amino acid changes). Only variants with a high impact were left. Considering the randomness of the mutations, identical SNSs or inDels were unlikely to occur in more than 1 fetus. The identical genetic mutations detected in more than 2 fetuses were also discarded.

Results

No large-scale CNV was detected in any of the 11 fetuses

This study included 11 fetuses with abnormal brain structure (*Figure 1* and Table S1). The brain abnormalities affecting the fetuses included dandy-walker complex, megalencephaly, holoprosencephaly, schizencephaly, lobular whole forebrain, agenesis of the corpus callosum, and abnormal caudate nucleus (*Figure 1* and Table S1). To explore the genetic mechanisms of these abnormalities at the molecular level, umbilical blood or villus tissue samples were collected and subjected to WES. The sequencing reads were mapped to the human reference genome (hg19). To explore if there were chromosomal or genomic large-scale CNVs, the sequencing depth for each chromosome was estimated. The sequencing depth was consistent among different regions of the genome, suggesting that there were no large-scale CNVs present in the included cases (*Figure 2* and Figure S1).

Genetic variants with high impact

To identify the genetic variants associated with the brain structural abnormalities of the fetuses, we performed SNS

© Annals of Translational Medicine. All rights reserved.

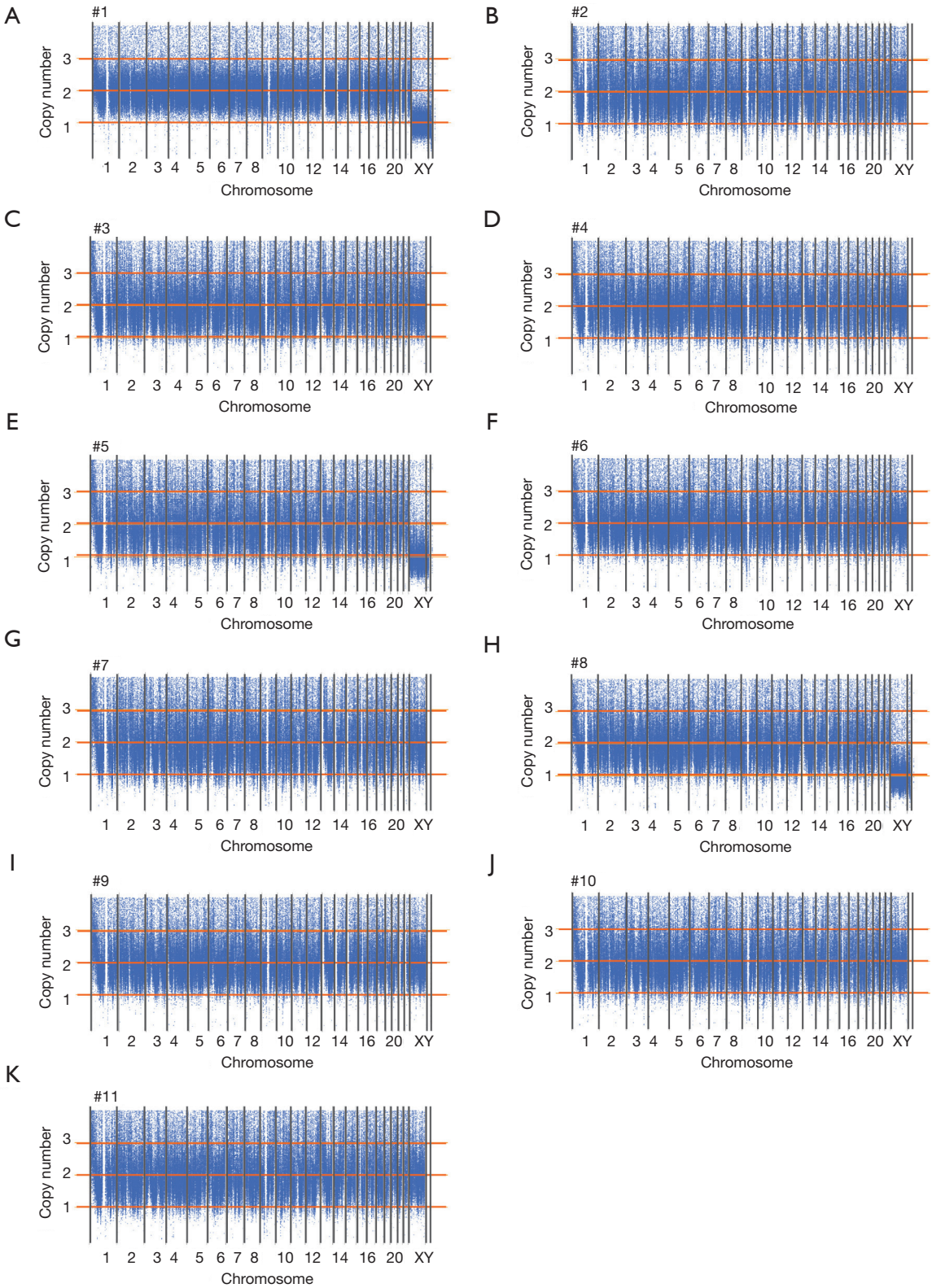


Figure 2 Karyotypes of fetuses with abnormal brain structure. Each chromosom is separated into 10-kb windows. Each point represents a 10-kb region. The x-axis and y-axis show the genomic position and the normalized sequencing depth of the corresponding 10-kb region.

© Annals of Translational Medicine. All rights reserved.

Table 1 Summary of genetic variants in 11 fetuses

Sample ID	Total No. of SNSs	After 1KG filtering, n (%)	After dbSNP filtering, n (%)	After removal of intronic and intergenic mutations, n (%)	Leading to nonsynonymous mutation, n (%)	With high impact*, n (%)
SNSs						
#1	1,278,399	279,872 (21.89)	170,343 (13.32)	5,183 (0.41)	429 (0.034)	43 (0.003)
#2	1,333,084	301,683 (22.63)	192,668 (14.45)	5,628 (0.42)	415 (0.031)	32 (0.002)
#3	1,326,336	310,812 (23.43)	208,080 (15.69)	7,366 (0.56)	973 (0.073)	96 (0.007)
#4	1,376,684	327,728 (23.81)	214,583 (15.59)	5,876 (0.43)	428 (0.031)	34 (0.002)
#5	1,372,275	440,550 (32.10)	337,287 (24.58)	8,692 (0.63)	504 (0.037)	39 (0.003)
#6	1,538,946	427,878 (27.80)	311,882 (20.27)	9,076 (0.59)	798 (0.052)	91 (0.006)
#7	1,446,788	396,661 (27.42)	285,334 (19.72)	7,331 (0.51)	447 (0.031)	35 (0.002)
#8	1,597,324	363,375 (22.75)	232,314 (14.54)	6,254 (0.39)	507 (0.032)	41 (0.003)
#9	1,629,349	539,857 (33.13)	426,183 (26.16)	11,326 (0.70)	895 (0.055)	80 (0.005)
#10	1,472,279	428,632 (29.11)	318,958 (21.66)	7,788 (0.53)	376 (0.026)	30 (0.002)
#11	1,480,871	492,572 (33.26)	387,997 (26.20)	10,901 (0.74)	1184 (0.080)	105 (0.007)
inDels						
#1	204,209	98,956 (48.46)	64,359 (31.52)	2,295 (1.12)	2 (0.001)	69 (0.034)
#2	210,641	101,285 (48.08)	63,986 (30.38)	2,481 (1.18)	6 (0.003)	94 (0.045)
#3	237,898	127,319 (53.52)	88,041 (37.01)	4,026 (1.69)	2 (0.001)	351 (0.148)
#4	188,582	86,911 (46.09)	55,221 (29.28)	2,236 (1.19)	0 (0.000)	79 (0.042)
#5	199,339	102,833 (51.59)	69,681 (34.96)	2,599 (1.30)	2 (0.001)	98 (0.049)
#6	216,128	106,906 (49.46)	73,435 (33.98)	3,245 (1.50)	4 (0.002)	236 (0.109)
#7	180,998	85,274 (47.11)	55,591 (30.71)	2,260 (1.25)	2 (0.001)	84 (0.046)
#8	211,380	93,368 (44.17)	58,222 (27.54)	2,289 (1.08)	3 (0.001)	83 (0.039)
#9	220,887	114751 (51.95)	81,586 (36.94)	3,350 (1.52)	2 (0.001)	296 (0.134)
#10	199,011	96,822 (48.65)	64,827 (32.57)	2,381 (1.20)	2 (0.001)	77 (0.039)
#11	196,554	104,308 (53.07)	75,116 (38.22)	3,581 (1.82)	3 (0.002)	379 (0.193)

*, SNSs with a high putative impact were annotated with SnpEff, including SNSs or inDels leading to exon loss, frame shift, rare amino acid, splice acceptor, splice donor, start codon lost, stop codon gained, stop codon lost, and transcript ablation variants. SNSs, single nucleotide substitutions; inDels, insertions and deletions

calling and inDels calling using the Genome Analysis Toolkit (GATK) (8). On average, 1,441,121 SNSs and 205,966 inDels were detected per sample (*Table 1*). The adjacent SNSs or inDels were mainly located within 500 bp (Figures S2 and S3). To identify genetic variants associated with brain structure abnormalities, SNSs annotated in 1,000 Genome (1KGP) and dbSNP were filtered, and those within intronic and intergenic regions were removed. To further remove the genetic variants making an insignificant or no contribution to severe brain phenotypic defects, we predicted the impacts by the SNSs or inDels on genes using SnpEff (9). Finally, on average, 56 SNSs with high impact were identified per sample (*Table 1*), including SNSs leading to exon loss, frame shift, rare amino acid, splice acceptor, splice donor, start codon lost, stop codon gained, stop codon lost, and transcript ablation variants. Further, on average, 168 inDels with high impact were identified per sample (*Table 1*).

© Annals of Translational Medicine. All rights reserved.

GO enrichment analysis of gene sets with high-impact mutations

To identify the features of the SNSs and inDels identified from the samples, we studied the corresponding genes of these variants. There were 1,035 genes with high-impact genetic variants in the 11 fetuses. To explore if these genes have a common cellular function, we performed GO enrichment analysis (Table S2). Enriched biological process terms included cell migration, cell differentiation, synapsis, and forebrain development (*Figure 3A*). Consistently, many genes linked to the regulation of cell migration and cell differentiation are known to also be associated with abnormal neuronal migration (10-12). Cellular component enrichment analysis showed that the genes were enriched in the synapse, spindle pole, and centrosome (*Figure 3B*). These results are in accordance with observations by previous studies that centrosome maturation (2) and interference with mitotic spindle formation (3,13) and the spindle pole (14) affect neurogenesis, especially the cell cycle phases of mitosis. Furthermore, microtubule binding was an enriched term in molecular function GO analysis. Mutations affecting microtubule proteins have consistently been reported as being associated with abnormal neuronal migration and postmigrational development (15,16). The consistency between our results and those of previous studies demonstrate that our WES data could be used to identify genes with potential involvement in the pathogenesis of MCD. In addition to the genes known to be associated with MCDs, several other genes were identified, including a gene related to ATP binding and the regulation of Pho protein signal transduction, which proved the new genes to potentially be involved in MCD pathogenesis.

Mutations in CTDSP2

Genes with high-impact mutations detected in multiple fetuses may be more likely to be associated with brain development. We identified 7 genes with high-impact mutations in at least 3 fetuses (Table S3). These genes included *CTDSP2*, which is known to catalyze the dephosphorylation of 'Ser-5' within the tandem 7 residue repeats in the C-terminal domain (CTD) of the largest RNA polymerase II subunit POLR2A. *CTDSP2* negatively regulates RNA polymerase II transcription, possibly through controlling the transition from initiation/capping to processive transcription elongation (17). CTDSP2 is also recruited by RE1-silencing transcription factor (REST) to neuronal genes containing RE-1 elements, resulting in neuronal gene silencing in non-neuronal cells (18). We retrieved the expression data of *CTDSP2* from the Genotype-Tissue Expression (GTEx) project, and observed variation in the expression levels of *CTDSP2* among different human brain tissues (Figure S4). Among 13 types of brain tissue, cerebellar hemisphere tissue exhibited the highest expression of *CTDSP2* (*Figure 4A*), suggesting that mutations in *CTDSP2* may lead to severe brain phenotypic defects, mainly through inaction in the cerebellar hemisphere.

We checked 4 mutations (T111, I102, V105, and D260) in *CTDSP2* in 5 fetuses carefully. All of the mutations were associated with amino acid changes in the protein sequence (*Figure 4B*). To understand the impact of these mutations, we predicted the protein structure. Three of these 4 mutations (I102, V105, and D260) were in the beta-pleated sheets but extremely close to the alpha-helix, which suggested that mutations in these sites may alter the protein structurally, further destroying its function (*Figure 4C*).

Mutations in C-terminal binding protein 2 (CTBP2)

Another gene with a high-impact mutation in 5 fetuses was the *CTBP2* gene. The mammalian *CTBP2* gene produces alternative transcripts encoding 2 distinct proteins, 1 of which is a transcriptional repressor (19), with the other isoform being a major component of specialized synapses known as synaptic ribbons. Both proteins contain a nicotinamide adenine dinucleotide (NAD+) binding domain similar to NAD+-dependent 2-hydroxyacid dehydrogenases. The *CTBP2* expression data from GTEx showed that the expression levels of *CTBP2* varied among different human brain tissues (Figure S5). Among 13 types of brain tissue, *CTBP2* was also highly expressed in the cerebellar hemisphere and cerebellum (*Figure 5A*), indicating that mutations in *CTDSP2* may cause severe brain phenotypic defects, mainly through inaction in the cerebellum.

Identification of the mutation sites of the CTBP2 protein revealed that a frame shift had occurred in 2 of the 5 fetuses, while 2 nonsynonymous mutations (D112A and P34L, *Figure 5B,C*) had occurred in the other 3 fetuses. The predicted homodimer protein structure of CTBP2 showed that these 2 amino acid changes were outside the interaction domain, which suggested that these mutations may destroy the function of the protein by changing its structure rather than by destroying its interaction ability.

© Annals of Translational Medicine. All rights reserved.

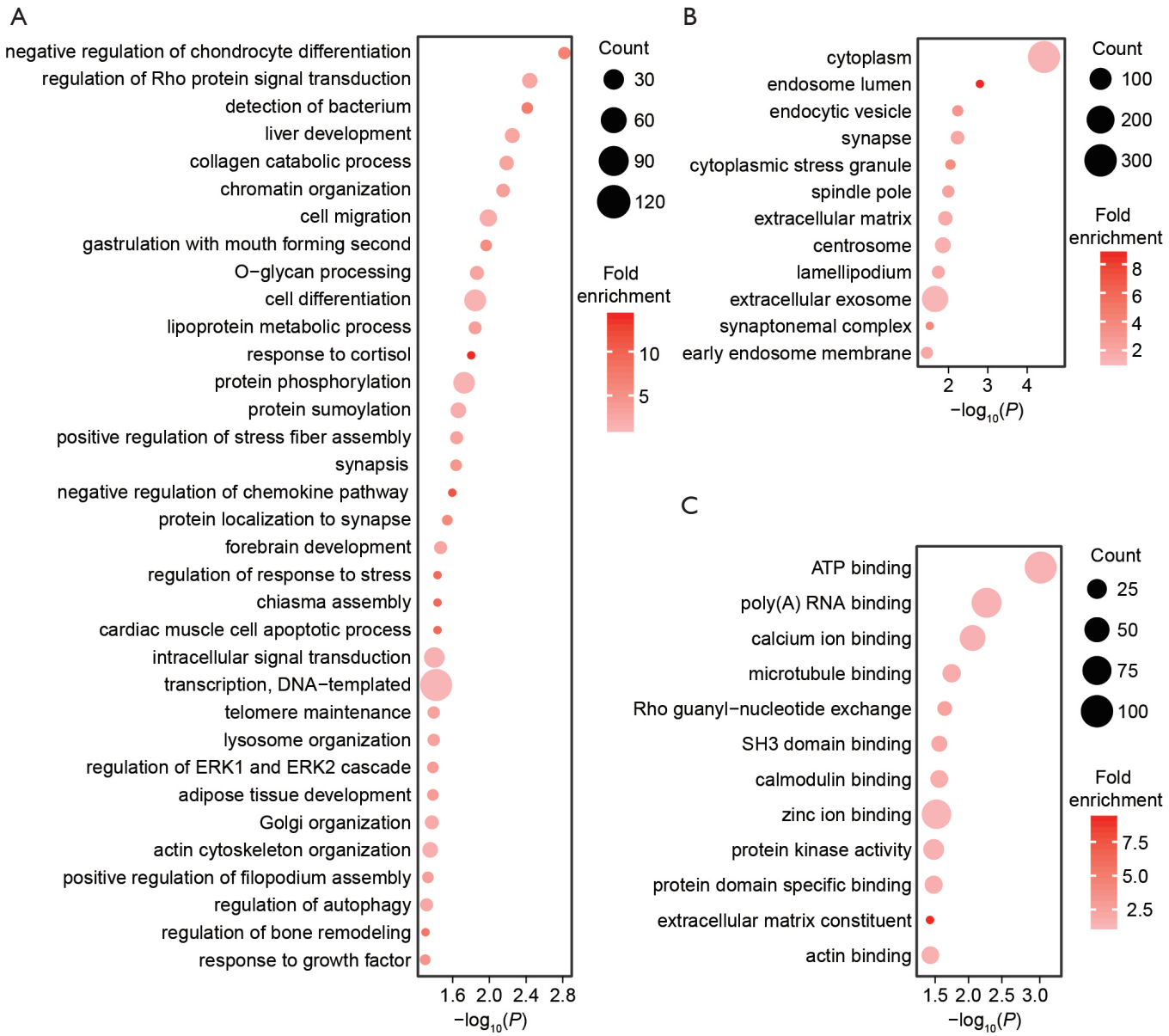


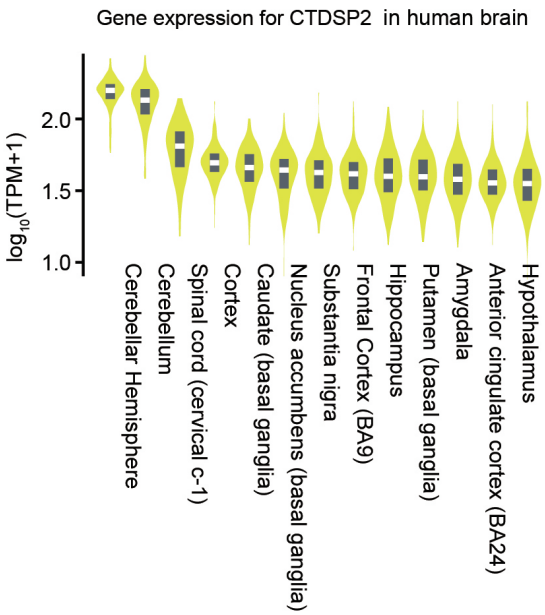
Figure 3 GO enrichment analysis of 1035 genes occurring with nonsynonymous mutations in 11 fetuses. Enriched biological process (A), cellular component (B), and molecular function (C) terms for 1,035 genes. Terms with P<0.05 are shown. All protein coding genes are included as the background. The size of each point represents the number of genes in the corresponding term.

Discussion

In this study, we identified novel genetic mutations potentially associated with MCDs, and our results expand the mutation spectrum of the disease. The consistency between our observations and those of previous studies indicate the high efficiency and sensitivity of WES, as an extremely helpful approach to identifying the genetic causes of MCDs. Considering that the time and economic cost of WES have been remarkably decreased, and the interpretation of sequencing data has been largely improved, WES provides a promising application for the diagnosis of disease. Although we have provided data suggesting the potential roles of genes in the pathogenesis of MCDs, namely *CTBP2*, *CTDSP2*, *HLA-DRB5*, *LZTFL1*, *MUC19*, *MUC4*, and *MUC6* (Table S3), it is necessary

© Annals of Translational Medicine. All rights reserved.

A



B

Patient	Genome variant	cDNA variant	Protein variant
#1	chr12:g.58217422T>C	c.779A>G	p.D260G
#4	chr12:g.58220819A>G	c.314T>C	p.V105A
#8	chr12:g.58220801G>T	c.332C>A	p.T111N
#9	chr12:g.58220827G>C	c.306C>G	p.I102M
#10	chr12:g.58220801G>T	c.332C>A	p.T111N

C

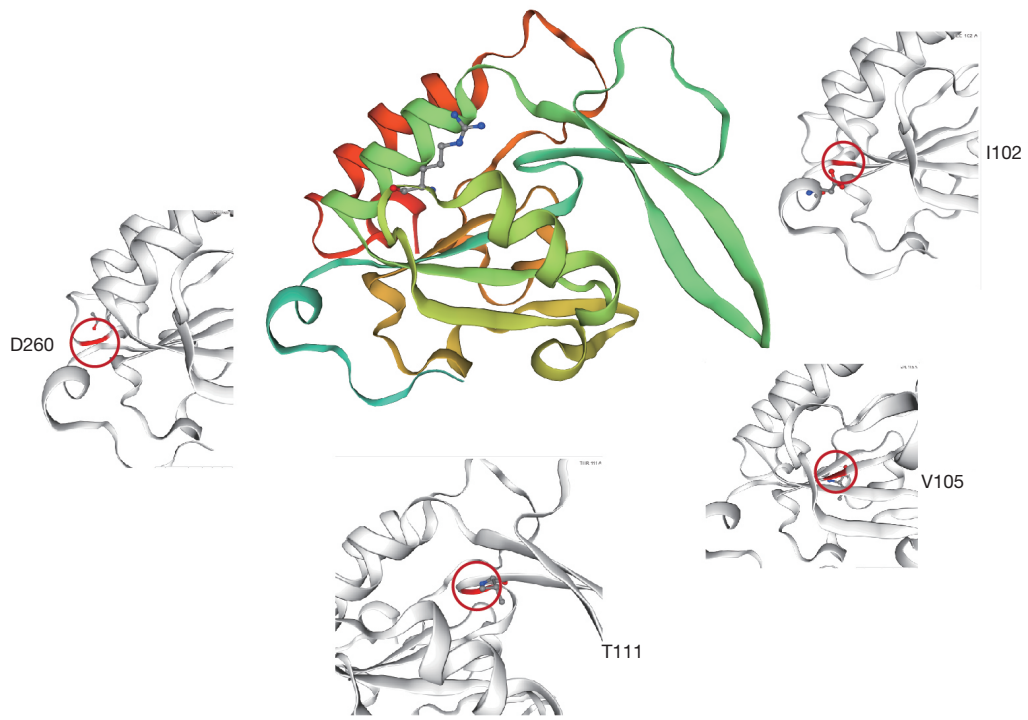
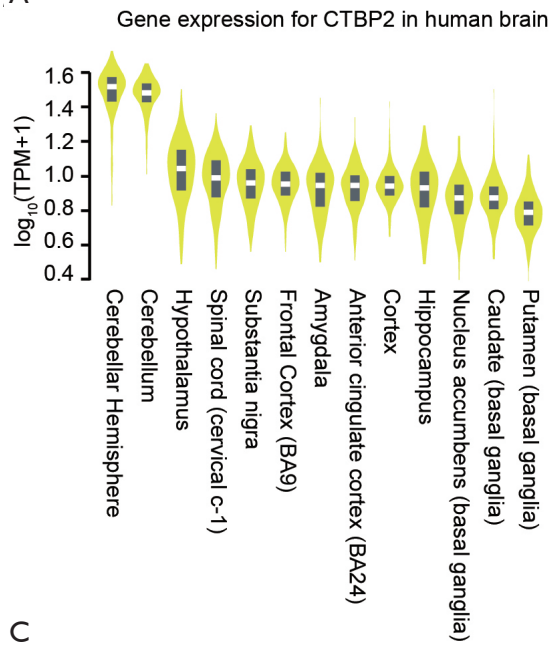


Figure 4 Mutations in *CTDSP2*. (A) Gene expression levels for *CTDSP2* in 13 different types of brain cells. The plot was generated by GTEx (https://www.gtexportal.org/home/); (B) summary of *CTDSP2* mutations in 5 fetuses; (C) the protein structure of *CTDSP2*. The protein structure homology modelling was generated using SWISS-MODEL (https://swissmodel.expasy.org/). The monomer structure of CTDSP2 is shown (GMQE =0.64, QMEAN =−1.05). Nonsynonymous mutation sites in 5 fetuses are shown individually.

© Annals of Translational Medicine. All rights reserved.

A



B

Patient	Genome variant	cDNA variant	Protein variant
#1	chr10:g.126691552T>G	c.335A>C	p.D112A
#4	chr10:g.126727724C>16inserts		5aa inserts
#5	chr10:g.126692019G>A	c.101C>T	p.P34L
#9	chr10:g126714705C>56inserts		Frameshift
#10	chr10:g.126691552T>G	c.335A>C	p.D112A

C

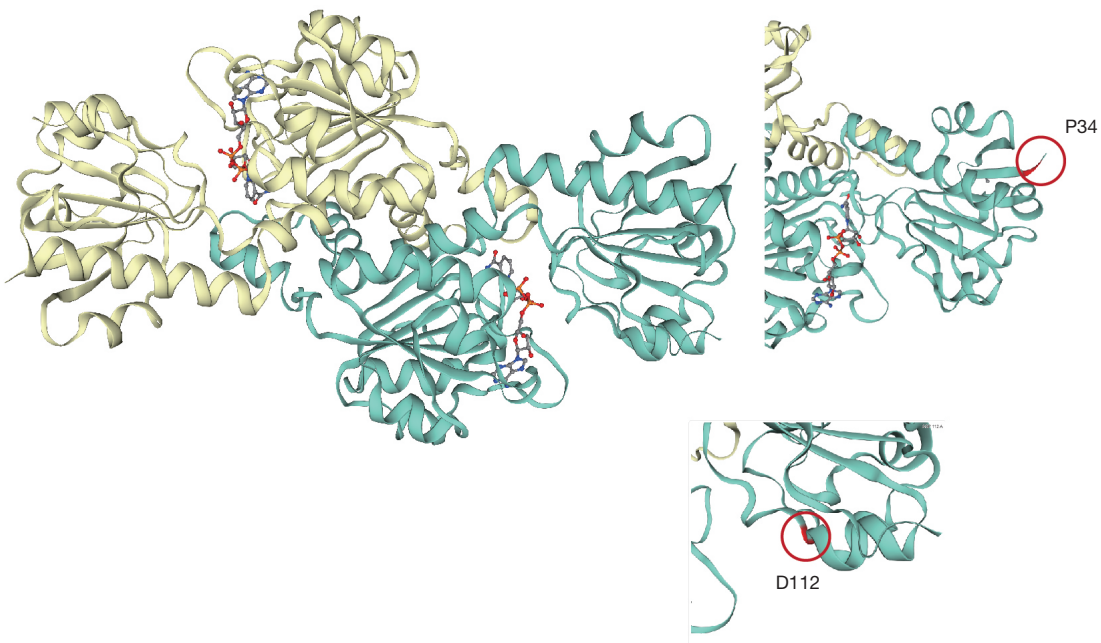


Figure 5 Mutations in *CTBP2*. (A) Gene expression levels for *CTBP2* in 13 different types of brain cells. The plot was generated by GTEx (https://www.gtexportal.org/home/); (B) summary of *CTBP2* mutations in 5 fetuses; (C) the homodimer protein structure of *CTBP2*. The protein structure homology modelling was generated using SWISS-MODEL (https://swissmodel.expasy.org/). The homodimer structure of *CTDSP2* is shown (GMQE =0.76, QMEAN =0.17). Two nonsynonymous mutation sites in cases #1, #5, and #10 are shown individually.

© Annals of Translational Medicine. All rights reserved.

to discuss the limitations of our work. Firstly, to further confirm the genetic causes of MCDs the mutations in corresponding samples should be detected by Sanger sequencing. However, due to the difficulty in obtaining villous and cord venous blood from the fetuses and in order to obtain high-quality sequencing data, WES was used for all samples collected. Secondly, whether the developmental phenotype of brain tissue is abnormal should be studied in related gene knockout animal models. We intend to carry out such experiments in our future study.

The expression levels and protein structures of candidate genes should be considered in the exploration of the underlying genetic mechanisms of MCDs. In this study, we retrieved the expression data from the GTEx project and compared the expression levels of genes among different tissues. The tissues exhibiting higher expression levels were more likely to be those for which the normality of the gene should be kept. Furthermore, mutation sites and protein structural domains can together provide clues for improving the understanding of the molecular consequences of genetic mutations.

Acknowledgments

Funding: This work was supported by a grant from the Ministry of Science and Technology of the People's Republic of China (2016YFC1000400).

Footnote

Reporting Checklist: The authors have completed the STREGA reporting checklist. Available at http://dx.doi.org/10.21037/atm-21-1477

Data Sharing Statement: Available at http://dx.doi.org/10.21037/atm-21-1477

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/atm-21-1477). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study was approved by the ethical committee of Beijing Haidian maternal and Child Health Hospital and was performed in accordance with the Helsinki Declaration (as revised in 2013). Individual consent for this retrospective analysis was waived.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Barkovich AJ, Guerrini R, Kuzniecky RI, et al. A developmental and genetic classification for malformations of cortical development: update 2012. Brain 2012;135:1348-69.
2. Thornton GK, Woods CG. Primary microcephaly: do all roads lead to Rome? Trends Genet 2009;25:501-10.
3. Bilgüvar K, Ozturk AK, Louvi A, et al. Whole-exome sequencing identifies recessive WDR62 mutations in severe brain malformations. Nature 2010;467:207-10.
4. Poduri A, Evrony GD, Cai X, et al. Somatic activation of AKT3 causes hemispheric developmental brain malformations. Neuron 2012;74:41-8.
5. Rivière JB, Mirzaa GM, O'Roak BJ, et al. De novo germline and postzygotic mutations in AKT3, PIK3R2 and PIK3CA cause a spectrum of related megalencephaly syndromes. Nat Genet 2012;44:934-40.
6. Reiner O, Carrozzo R, Shen Y, et al. Isolation of a Miller-Dieker lissencephaly gene containing G protein beta-subunit-like repeats. Nature 1993;364:717-21.
7. Fallet-Bianco C, Loeuillet L, Poirier K, et al. Neuropathological phenotype of a distinct form of lissencephaly associated with mutations in TUBA1A. Brain 2008;131:2304-20.
8. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 2011;43:491-8.
9. 9. Cingolani P, Platts A, Wang le L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 2012;6:80-92.
10. Wynshaw-Boris A. Lissencephaly and LIS1: insights into

© Annals of Translational Medicine. All rights reserved.

the molecular mechanisms of neuronal migration and development. Clin Genet 2007;72:296-304.
11. Ferland RJ, Batiz LF, Neal J, et al. Disruption of neural progenitors along the ventricular and subventricular zones in periventricular heterotopia. Hum Mol Genet 2009;18:497-516.
12. Pramparo T, Youn YH, Yingling J, et al. Novel embryonic neuronal migration and proliferation defects in Dcx mutant mice are exacerbated by Lis1 reduction. J Neurosci 2010;30:3002-12.
13. Yu TW, Mochida GH, Tischfield DJ, et al. Mutations in WDR62, encoding a centrosome-associated protein, cause microcephaly with simplified gyri and abnormal cortical architecture. Nat Genet 2010;42:1015-20.
14. Nicholas AK, Khurshid M, Desir J, et al. WDR62 is associated with the spindle pole and is mutated in human microcephaly. Nat Genet 2010;42:1010-4.
15. Poirier K, Keays DA, Francis F, et al. Large spectrum of lissencephaly and pachygyria phenotypes resulting from de novo missense mutations in tubulin alpha 1A (TUBA1A). Hum Mutat 2007;28:1055-64.
16. Abdollahi MR, Morrison E, Sirey T, et al. Mutation of the variant alpha-tubulin TUBA8 results in polymicrogyria with optic nerve hypoplasia. Am J Hum Genet 2009;85:737-44.
17. Yeo M, Lin PS, Dahmus ME, et al. A novel RNA polymerase II C-terminal domain phosphatase that preferentially dephosphorylates serine 5. J Biol Chem 2003;278:26078-85.
18. Yeo M, Lee SK, Lee B, et al. Small CTD phosphatases function in silencing neuronal gene expression. Science 2005;307:596-600.
19. Turner J, Crossley M. The CtBP family: enigmatic and enzymatic transcriptional co-repressors. Bioessays 2001;23:683-90.

(English Language Editor: J. Reynolds)

Cite this article as: Shi L, Li M, Qi H, Zhu J, Yang J, Tang J, Wang L. Whole-exome sequencing analysis to identify novel potential pathogenetic mutations in fetuses with abnormal brain structure. Ann Transl Med 2021;9(9):807. doi: 10.21037/atm-21-1477

© Annals of Translational Medicine. All rights reserved.

Supplementary

Table S1 Phenotypes of 11 fetuses

Case no.	Ultrasound diagnostic phenotype
#1	The shape of the fetal head is "strawberry"; the posterior horn of both lateral ventricles is about 1.0 cm wide; and the cisterna magna is about 1.2 cm wide
#2	The fetal head circumference is enlarged; there is forehead protuberance; the intracranial frontal lobe is full and disordered; and the midline, thalamus, and cerebellum are visible
#3	Fetus shows a single ventricle of approximately 0.5 cm wide
#4	Fetal head circumference ultrasound value is lower than M-2SD line
#5	The intracranial left cerebellar hemisphere of the fetus has a visible range of about 3.3 cm×1.4 cm without echo. The lower margin is immediately above the cerebellar curtain; the medial margin is near the midline of the brain; and the upper margin is near the thalamus. The shape is irregular, and there is no obvious communication with the left ventricle
#6	The fetal head circumference ultrasound value is lower than the M-2SD line
#7	The fetal head circumference ultrasound value is lower than the M-2SD line
#8	The midline of the fetus' brain is not continuous, and the anterior horns of the lateral ventricles communicate with each other
#9	The fetal biparietal diameter and head circumference ultrasound measurements are lower than the M-2SD line
#10	The posterior horns of the bilateral ventricles of the fetus are about 1.2 cm wide, and the anterior horns are abducted. The bilateral ventricles showed water drop-like appearance.. The transparent compartment is not shown. The third ventricle is raised and about 0.4 cm wide. The cisterna magna is about 1.2 cm wide. The cerebellar vermis is not shown. The cisterna magna is connected to the 4^{th} ventricle, which measures about 0.6 cm wide
#11	The left lateral ventricle of the fetus is about 1.15 cm wide, and the right lateral ventricle is about 1.09 cm wide

M-2SD, Mean-2 Standard Deviation; cm, Centimeter

© Annals of Translational Medicine. All rights reserved.
http://dx.doi.org/10.21037/atm-21-1477

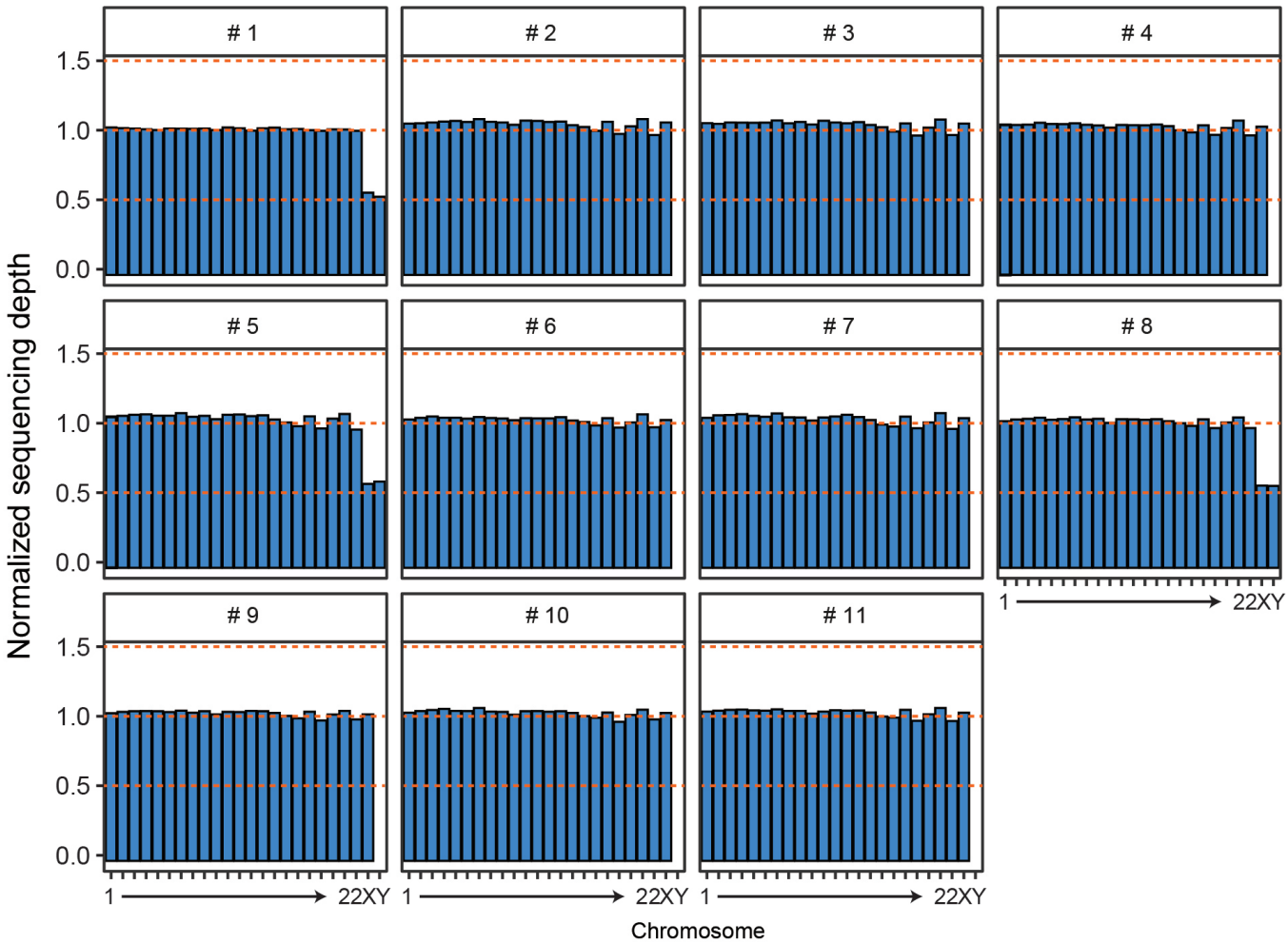


Figure S1 The average normalized sequencing depth of each chromosome for 11 fetuses. The sequencing coverage of each 10-kb genomic region is normalized by the median coverage of all 10-kb regions.

© Annals of Translational Medicine. All rights reserved.
http://dx.doi.org/10.21037/atm-21-1477

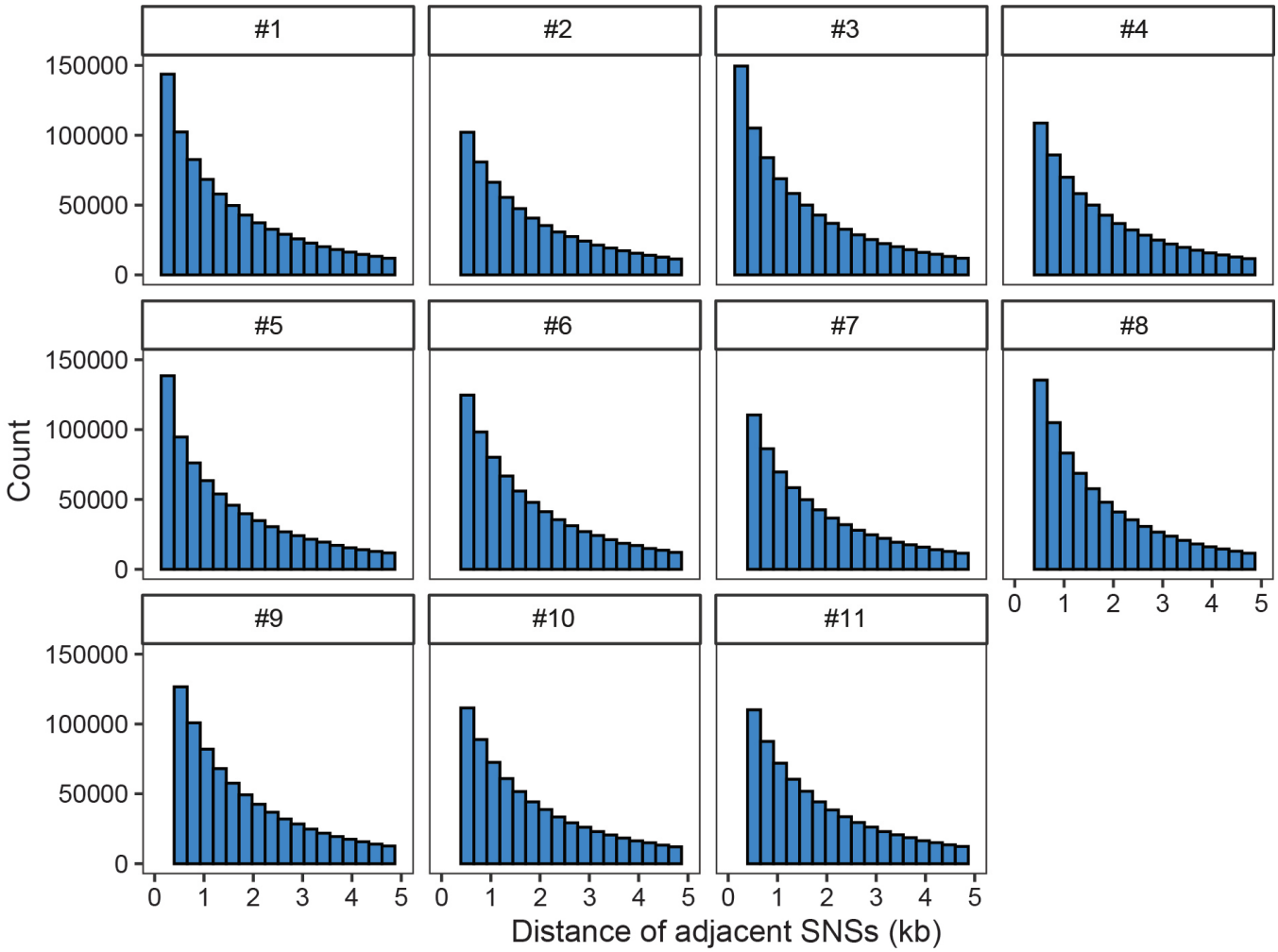


Figure S2 The distribution of the distance between adjacent single nucleotide substitutions (SNSs).

© Annals of Translational Medicine. All rights reserved.
http://dx.doi.org/10.21037/atm-21-1477

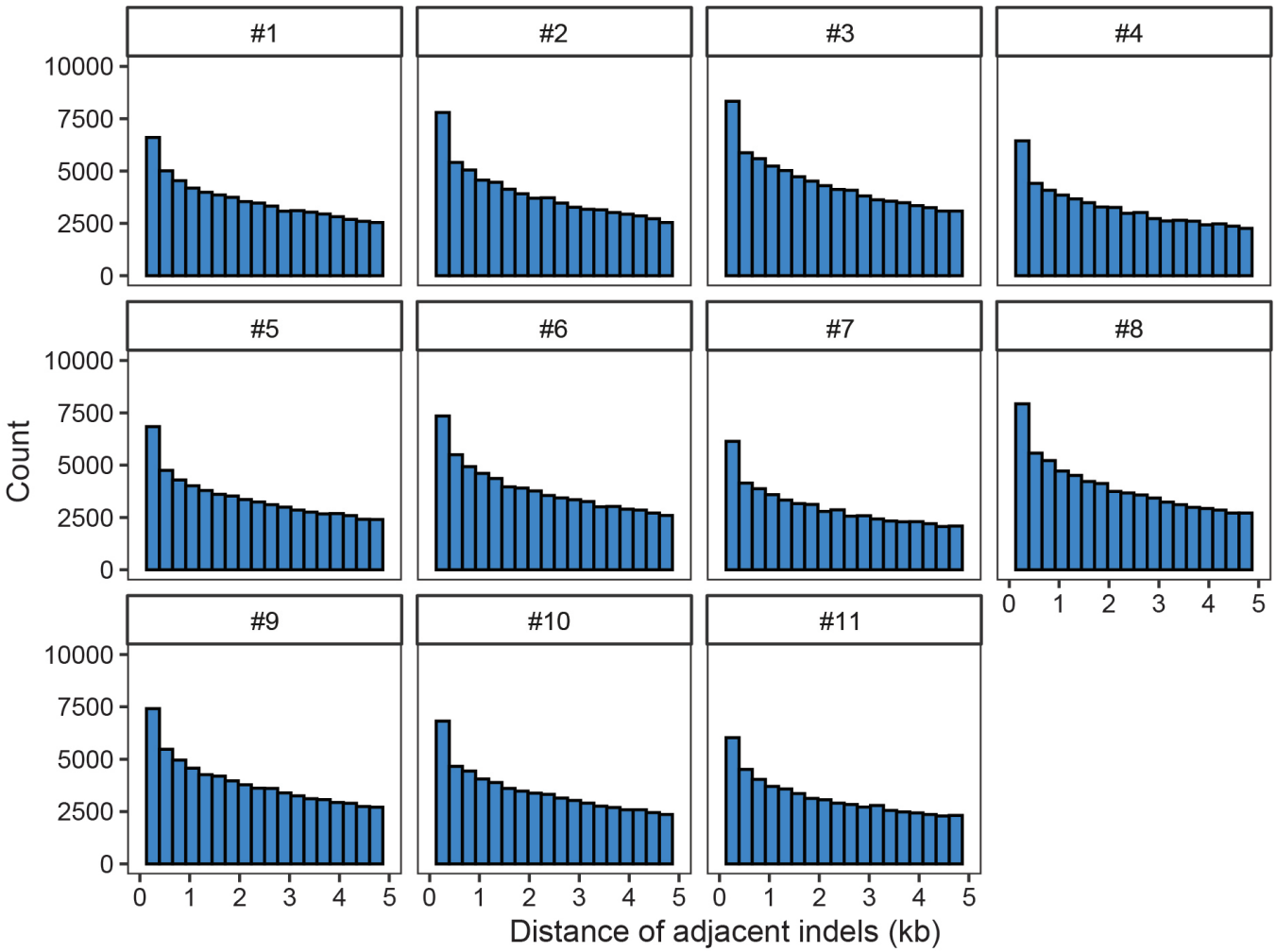


Figure S3 The distribution of the distance between adjacent insertions and deletions (inDels).

© Annals of Translational Medicine. All rights reserved.
http://dx.doi.org/10.21037/atm-21-1477

Table S2 Genes listed in each enriched GO term

Term	Genes
Negative regulation of chondrocyte differentiation	*ADAMTS7, TGFBR1, GDF5, ADAMTS12, GLI2, SNAI2*
Regulation of Rho protein signal transduction	*DLC1, MURC, OBSCN, ARHGEF33, PLEKHG6, ARHGEF28, ARHGEF5, TRIO, ARHGEF9, SPATA13, ARHGEF10L, ARHGEF10*
Detection of bacterium	*NLRC4, NOD1, PGLYRP4, HLA-DRB1, HLA-DRB5*
Liver development	*ACADM, XBP1, GATA6, IGF2R, RB1CC1, NKX2-8, PKD1, PCSK9, CAD, AACS, LSR*
Collagen catabolic process	*COL4A4, COL4A1, COL19A1, COL6A6, COL7A1, COL13A1, MMP27, CTSS, COL5A2, MMP12*
Chromatin organization	*SMARCAD1, HMGB2, CDAN1, SOX2, PADI2, ZNF462, PADI4, CHD4*
Cell migration	*FGFR4, TMEM18, PLXNB1, CD248, ELMO3, GAS6, EPHA3, FAM83D, NCK2, PLCG1, PSTPIP1, DGKZ, JAK2, IL12B, ADAMTS12, USP24, SPATA13, ADAM9*
Gastrulation with mouth forming second	*RNF2, UGDH, AMOT, CHRD, LRP5*
O-glycan processing	*GALNT10, B3GNT5, MUC5AC, MUC19, MUC6, MUC17, MUC12, MUC4, MUC16*
Cell differentiation	*ETV7, TMEM189-UBE2V1, FOXA3, CREM, ARHGEF28, NAA15, ANPEP, MYBL2, CFAP54, MUSK, FLI1, OSR2, SMOC1, ZC3H12A, ROS1, C9ORF24, BCL7B, KIF2A, INA, COL13A1, CENPF, MORC1, ABCB5, CEL, MURC, SPAG8, COL19A1, TEX15, HUWE1, TXNRD3, MYCBPAP, JAK2, ADAM18, TNK1, APAF1, NHLH1, DUSP6*
Lipoprotein metabolic process	*APOB, APOA1, LRP1, HSPG2, NPC1L1, PCSK9, ABCA1*
Response to cortisol	*CAD, SLIT2, SLIT3*
Protein phosphorylation	*CDK19, FASTKD1, FASTKD2, PRPF4B, NEK2, BRSK1, CAMKK2, CSNK2A1, PRKRA, ROS1, AATK, PRKCA, OBSCN, BRAF, TGFBR1, NEK10, TRIO, BIRC6, NEK11, GAS6, CDC25B, FNIP2, DAPK1, MAPK1, HUNK, CDC42BPG, RPS6KA3, TEX14, CAMK4, ULK1, RSRC1, NEK8, BUB1B, JAK2, CPNE3, TNK1*
Protein sumoylation	*TOP1, NUP98, HNRNPK, SP100, BLM, INCENP, RNF2, TDG, PIAS2, UBE2I, HERC2, NUPL2, RNF212*
Positive regulation of stress fiber assembly	*APOA1, BRAF, TGFBR1, ARHGEF5, AMOT, ARHGEF10L, ARHGEF10*
Synapsis	*TEX15, FANCD2, MEIOB, SYCP1, RNF212*
Negative regulation of chemokine-Mediated signaling pathway	*PADI2, SLIT2, SLIT3*
Protein localization to synapse	*DLG4, NLGN2, SHANK1, PCLO*
Forebrain development	*NOTCH3, DLC1, SOX2, VAX2, APAF1, SETD2, CHRD*
Cardiac muscle cell apoptotic process	*HSP90AA1, GNB1, APAF1*
Chiasma assembly	*MSH5, SYCP1, RNF212*
Regulation of DNA-templated Transcription in response to stress	*RPS6KA3, BCLAF1, RGS14*
Intracellular signal transduction	*ADCY4, ARHGEF28, BRSK1, STAC3, NOD1, DGKB, PIKFYVE, ANP32A, PRKCA, MAGI3, PLXNB1, ZP4, TGFBR1, ARHGEF5, SPHK1, DGKK, RGS14, NEK11, DAPK1, MAST4, HUNK, PDZD8, RPS6KA3, CDC42BPG, PLCG1, DUSP1, CAMK4, RASSF1, DGKZ, JAK2, UNC13C*
Transcription, DNA-templated	*HKR1, THRA, TMEM18, NAA15, ASCC1, ZNF250, CNOT1, FOXO6, MAMSTR, CSNK2A1, SMARCD3, ZNF302, ZNF578, ZNF391, ZNF595, RBL2, NKX2-8, VAX2, TOX2, TRIM37, MAPK1, KDM2A, BAZ1B, RFC1, MCIDAS, ZNF239, HOXD4, MNX1, ZNF384, PIAS2, ASCL5, HMGB2, ZNF517, HOXA13, SETD1A, CXXC1, MED12L, TAL1, TCF20, ZNF429, LYL1, ZNF697, MN1, ASXL2, KCTD1, SKI, TEAD2, ZSCAN4, MYPOP, DPPA4, ZFHX4, ZNF311, RNF2, TDG, DNMT1, RFX2, ZFHX2, PHF6, HDAC6, TAF1B, TCEANC, POU6F1, E2F4, ZNF81, FOXA3, ZKSCAN1, ZKSCAN4, ZNF726, ANP32A, HSF5, LRRFIP1, INO80D, SP100, CTBP2, ZNF544, ZNF142, NCOA7, IRF2BP2, PADI4, ZNF335, MURC, NAB2, ZNF717, PERM1, ZNF710, ZFPM2, CUX2, COMMD6, BCLAF1, ABLIM3, CREM, WTIP, CIC, MYCBP2, ZNF512B, NPAS3, TCEA3, RB1CC1, PRDM10, ZC3H12A, GATAD2B, PER3, ZSCAN23, KDM3B, BCAS3, CHD6, FOXD2, CHD4, FOXD4, TXNIP, TBX6, TRNP1, ZNF454, POLR3A, SNAI2, SNAI3, ATXN3, ZNF860, SFPQ, ZBTB5, ZNF462, ZNF573*
Telomere maintenance	*RAD51D, SP100, ACD, POLD2, SMG1, CTC1*
Lysosome organization	*GNPTAB, TPP1, LYST, ABCA1, VPS11, ZKSCAN4*
Adipose tissue development	*LEP, AMER1, XBP1, AACS, LRP5*
Regulation of ERK1 and ERK2 cascade	*FAM83D, ARHGEF5, NEK10, RRAS, ROS1*
Golgi organization	*ARHGAP21, SYNE1, CUL7, SPTBN5, GORASP1, HTT, ATP8B2, BCAS3, KIFC3*
Actin cytoskeleton organization	*DLC1, WASF3, SPTBN5, FOXJ1, WASF1, ABLIM3, ARHGEF5, ATXN3, NISCH, XIRP2, AMOT, PDE4DIP, IQSEC1*
Positive regulation of filopodium assembly	*DNM3, CCR7, TGFBR1, DPYSL3, BCAS3*
Regulation of autophagy	*XBP1, SOGA3, ULK1, USP13, DAPK1, PIK3R2, HDAC6*
Regulation of bone remodeling	*LEP, LRP5, SUCO*
Response to growth factor	*GATA6, SOX2, DUSP6, HDAC6*
Negative regulation of phosphatase activity	*SH2D4A, PPP1R37, CEP192, MYO1D, SFI1, WDR81, RBM26*
Cytoplasm	*RAD51D, DLC1, ADCY4, NAF1, HRNR, TMEM18, FAM110C, NAA15, KPRP, NAA16, CDC16, CD2AP, DUPD1, APOB, CUL7, AFDN, DHX35, LRRC4, SH3GL3, KRR1, CEP295, DBNL, PTPRM, BRAF, UBR3, HCFC1R1, MYH8, VASH1, PCF11, MAPK1, SPAG8, HUWE1, RFC1, CAMK4, ZWINT, BHMT, NEK8, PDE4DIP, PIDD1, UNC13C, AHCY, NEK2, DNAH17, PABPC4, ARF5, KRT27, ACD, UBASH3B, RCCD1, MYO15B, PABPC3, PABPC1, ARHGEF10L, IQSEC1, SGIP1, PTGR1, TRPC6, LGALS3, PCNT, SPHK1, EPM2A, SKI, CDC27, LARP4B, FNIP1, GAS6, GART, FNIP2, NOTCH3, CAPN15, CEL, PLA2G4A, SYNE1, CDC42BPG, SYNE2, KIAA0922, TNK1, MUC5AC, UTP20, BACH1, CRNKL1, CPEB2, WASF3, HIP1R, TCOF1, TTC28, ZNF346, CAMKK2, CKB, MSRA, SMAP1, NOD1, ANP32A, TSEN2, CATIP, ASTL, ACO1, CLMN, INPPL1, ARHGEF5, MICAL3, RPH3AL, HERC3, CCT6A, ARHGEF9, HERC2, DECR1, ELMO3, SLIT2, NAPRT, ANKRD27, CARD11, TAF10, MAST4, NCK2, EFHC1, DNPH1, FANCD2, PLA2G6, IL12B, THOC5, USP24, SSU72, PPFIA2, RAB3B, CREM, USP4, FKBP1A, CEP128, FEM1A, STAU1, NECAB1, CFAP54, CYP27B1, DGKB, ATXN2L, C19ORF68, XBP1, PLEKHG6, RB1CC1, FASN, PKD1, PCSK9, PER3, BCAS3, DTNA, ATRN, CAPN9, RTN4R, DGKK, FZD2, ITPR3, SNAI2, RAPH1, SAMD1, CABYR, RPS6KA3, TEX14, TEX15, LRP1, DUSP1, RPL21, RASSF1, SFPQ, AOX1, BLVRB, MEX3C, MYCBPAP, USP47, AMOT, DGKZ, SLFN11, DUSP6, ASPN, KAT14, TMEM189-UBE2V1, TBK1, TBP, CNOT1, ALDH1L2, UBQLN1, FOXO6, CTNNBL1, CHAD, CRYGD, KRT80, SMARCD3, DNAJC9, ZNF578, DND1, ACIN1, IFNL4, ASPM, CASKIN1, CFAP126, RBL2, DARS, NUDT3, SLC34A1, LRTOMT, VAX2, SOCS5, NUPL2, TRIM37, RALGAPA2, SGSM1, LYST, MNX1, SPATA13, DST, FILIP1L, CRACR2B, FBXO10, FGFR4, HMGB2, PODN, BLM, SOX2, GIPC2, MROH2B, CXXC1, WAPL, HNRNPK, OSBPL1A, GMPPA, PRKRA, THAP12, SLC30A3, ERCC6L, ZMYM1, SPTBN5, MYO1C, HTT, SMG5, SPTBN4, FANK1, SMG1, ABHD14B, SHANK1, TRIM62, GCN1, ARID3C, MPRIP, TRIM56, CASS4, PLCG1, NEDD4, ARMC7, ULK1, BEX5, CPNE3, THNSL2, RFX2, TMPO, PLEKHA1, HDAC6, IER2, OSBP, TLN2, ESD, GABBR2, SH2D4A, FAM83D, ALS2CR12, DLG4, SNTB2, LRRFIP1, ZCCHC7, CDK13, PRKCA, SP100, HSP90AA1, G3BP1, MEIOB, NEK10, PADI2, UBE2I, IRF2BP2, PADI4, EML5, DAPK1, HAGH, ACADVL, UBE2O, HUNK, MURC, UMPS, EVPL, PRDX6, IGF2R, ADGRV1, PERM1, BTG3, BUB1B, ZFPM2, COMMD6, DNAH9, BCLAF1, CD248, ABLIM3, EGLN1, BRSK1, DNAH5, DNAH6, NPAS3, CCDC125, THYN1, ZC3H12A, KLHL42, CHD4, RTCB, TXNIP, GEMIN2, CEP192, MAP1A, DTX3, CENPF, RGS18, DPYSL3, RGS14, CDC25B, SMC4, LEP, CYP4A11, ATXN3, ITPA, CCDC116, CDAN1, RSRC1, KCNN3, JAK2, SH3BP1*
Endosome lumen	*APOB, PRLR, JAK2, AP4M1, AP4S1*
Endocytic vesicle	*APOA1, RAB5B, RAB5C, IGF2R, AMOT, ABCA1, VPS11, CD2AP, DPP4*
Synapse	*PPFIA2, ACHE, CPEB2, GABRB3, TLN2, WASF1, SDK1, NLGN2, PCLO, CEL, NCK2, DMD, ADGRV1, SYN2, DOK7, DLG4, SNTB2, SLC5A7, DTNA*
Cytoplasmic stress granule	*ATXN2L, PABPC4, G3BP1, PUM1, PABPC1, LARP4B, STAU1*
Spindle pole	*NEK2, FAM110C, BIRC6, CENPF, TTC28, CEP128, RGS14, FBF1, FRY, CDC25B, RASSF1, PLEKHG6, KIF2A*
Proteinaceous extracellular matrix	*COL4A4, ASPN, LTBP1, ADAMTSL1, PODN, ZP4, ADAMTS16, CD248, VIT, COL5A2, MMP12, SLIT2, MUC4, SLIT3, CHAD, ADAMTS7, ADAMTS6, WISP2, ADAMTS8, COL19A1, COL6A6, SMOC1, FBN3, ADAMTS12*
Centrosome	*RAD51D, NEK2, BRSK1, CDC16, TTC28, WDR81, FBF1, KIFC3, CUL7, CEP250, TSEN2, CHD4, KIF2A, ERCC6L, CEP295, CEP192, PCNT, CENPF, NEK10, PDE4D, SKI, SFI1, CEP152, CENPJ, CDC27, ARHGEF10, RGS14, CDC25B, EFHC1, CCDC116, HSPB11, PDE4DIP, SPATC1L, WDR43*
Lamellipodium	*DBNL, PTPRM, HAX1, WASF3, INPPL1, WASF1, ABLIM3, KCNA2, DPYSL3, RAPH1, PLCG1, PSTPIP1, DGKZ, AMOT, SPATA13, DPP4*
Extracellular exosome	*HRNR, TMEM189-UBE2V1, SLC22A13, MASP2, KPRP, KIAA1324, NAA16, PDLIM2, TTN, ALDH1L2, UXS1, CD2AP, KIFC3, APOB, APOA1, VPS13D, AKR7A2, RRAS, PLA1A, SLC4A4, KRT84, IDUA, ADAM9, SH3GL3, DBNL, DARS, RBL2, PTPRG, NUDT3, PTPRS, HLA-C, COLEC12, AP4M1, PPP1CB, PCLO, SIRPA, MAPK1, CAMK4, RFC1, HUWE1, BHMT, MGAM, DST, TPST2, AHCY, HLA-DRB1, GIPC2, ARF5, AGMAT, NAGK, KRT27, HNRNPK, MTCH2, FOLR1, OSBPL1A, GMPPA, HLA-DRB5, PABPC3, PABPC1, ANGPTL2, DNAJA2, PTGR1, B4GAT1, LGALS3, SPTBN5, MYO1C, TBC1D10A, MYO1D, SPTBN4, AK2, ABHD14B, RBMX, GAS6, MUC4, GART, LAMA2, CEL, LAMA3, SYNE2, HSPB11, FAM151A, NEDD4, BEX5, CPNE3, MUC5AC, PLEKHA1, TAF1B, STEAP4, WASF3, RAB5B, RAB5C, SLC15A2, DUOX2, ESD, ANPEP, CAD, CDSN, LSR, CKB, WISP2, MSRA, CEP250, TPP1, DNAJC14, TKFC, CHST14, KLK11, SNTB2, SEMA3B, DPP4, SCN10A, PRKCA, HSP90AA1, ACADM, SLC25A5, ACO1, SDK1, PADI2, DECR1, CCT6A, EML5, SLIT2, NAPRT, HAGH, C8B, CARD11, EVPL, DNPH1, GNB1, PRDX6, IGF2R, ADGRV1, CUX2, WIZ, HIST1H2AB, PPFIA2, MEGF8, RAB3B, CD248, UGDH, BTN2A2, FKBP1A, STAU1, CBR1, PRSS3, FASN, PKD1, ZSCAN23, APMAP, QSOX1, THBS4, DNM3, SI, TMC4, ATRN, HSPG2, RTN4R, FRMPD1, A2ML1, CYP4A11, SDHB, TEX14, NEB, BLVRB, AOX1, SLC5A8, TCEB2, APAF1, MUC19, MUC16*
Synaptonemal complex	*WAPL, MSH5, UBE2I, SYCP1, RNF212*
Early endosome membrane	*SH3GL3, RAB5B, RCC2, EPHA8, RAB5C, ZFYVE9, PIKFYVE, INPP5F, HLA-C, CLIP3, SNX13, HLA-G*
ATP binding	*RAD51D, ADCY4, PRPF4B, TBK1, TTN, MTHFD1L, KIFC3, NLRC4, CSNK2A1, DHX35, ROS1, NMNAT3, MAGI3, DARS, BRAF, MYH8, CARNS1, MAPK1, BAZ1B, RFC1, CAMK4, NEK8, EP400, FGFR4, BLM, NEK2, DNAH17, ABCA1, NAGK, MUSK, MYO15B, ENTPD3, DDX43, DNAJA2, ERCC6L, CSF1R, OBSCN, NADK2, SHPRH, MYO1C, TGFBR1, MYO1D, MET, SPHK1, TRIO, SMG1, AK2, ATR, ABCB5, EPHA3, GART, CDC42BPG, EPHA8, ULK1, TNK1, CLCN5, CDK19, ATP10B, CAD, CAMKK2, CKB, FANCM, NOD1, TKFC, ACOT12, ATP8B2, CDK13, KIF2A, PRKCA, NLRP7, HSP90AA1, G3BP1, NEK10, UBE2I, CCT6A, NEK11, DAPK1, UBE2O, HUNK, MAST4, BUB1B, SLC27A5, SMARCAD1, DNAH9, BRSK1, BMS1, DNAH5, DNAH6, DGKB, MCMDC2, PIKFYVE, SYN2, CHD6, RTCA, CHD4, RTCB, AATK, ABCA10, MSH5, DGKK, SMC4, TEX14, RPS6KA3, ABCC3, SLFN12, DGKZ, JAK2, SLFN11, APAF1, ZRANB3, AACS*
Poly(A) RNA binding	*NAF1, FASTKD1, FASTKD2, PRPF4B, CPEB2, CRNKL1, TCOF1, KIAA1324, NAA15, CNOT1, ZNF346, TOP1, TRMT1L, PCBP4, URB1, ANP32A, PUM1, SNTB2, PUM3, RBM47, ACIN1, DHX35, RBM10, ZCCHC7, CDK13, MTERF1, KRR1, HSP90AA1, DARS, SLC25A5, G3BP1, UBE2I, CCT6A, NUPL2, UBE2O, NOP2, RCC2, HUWE1, FAM120C, CPSF4, CCDC9, WDR43, BCLAF1, HMGB2, PABPC4, FCF1, BMS1, STAU1, DROSHA, PPAN, TCF20, ATXN2L, HNRNPK, PRKRA, FASN, PCSK9, DMGDH, PABPC1, RBM28, RTCA, MEPCE, RBM26, RTCB, ZC3H13, LGALS3, SMG1, LRRC47, LARP4B, GCN1, RBMX, TRIM56, SON, SYNE1, MRPL21, LRP1, PTCD3, RPL21, SFPQ, MEX3C, CPNE3, UTP20, PHF6*
Calcium ion binding	*ASPN, EFCC1, HRNR, LTBP1, CLSTN2, MASP2, DUOX2, MMP27, FSTL4, TTN, CAMKK2, PCDH1, CD93, VWDE, PCDHGA10, BRAF, LRP1B, RPH3AL, STIM2, PADI2, CDHR4, PADI4, PCLO, SLIT2, MMP12, SLIT3, EFHC1, GRM7, ADGRV1, C2CD4A, DST, CRACR2B, DCHS2, MEGF8, CD248, NECAB1, DGKB, GNPTAB, FAT3, SMOC1, PRSS3, FBN3, PCDHA13, THBS4, CAPN9, HSPG2, ITPR3, GAS6, CABYR, NOTCH3, PLA2G4A, LRP1, PLCG1, EFHB, FLG*
Microtubule binding	*CAPN6, DNM3, MAP1A, WDR81, BCL2L11, RGS14, KIFC3, FAM83D, CFAP44, SPAG8, RCC2, CCSER2, CLIP3, MAP6, WDR43, JAKMIP3, CHD4, HDAC6, KIF2A, SGIP1*
Rho guanyl-nucleotide exchange factor activity	*OBSCN, ARHGEF33, PLEKHG6, ARHGEF28, ARHGEF5, TRIO, ARHGEF9, SPATA13, ARHGEF10L, ARHGEF10*
SH3 domain binding	*CTTNBP2, CASP9, INPPL1, DPYSL3, ADAM19, SH3BP1, SHANK1, SIRPA, ELMO3, CD2AP, CABYR, ADAM9, SGIP1*
Calmodulin binding	*OBSCN, MYO1C, PCNT, SPHK1, MYO1D, RRAD, TTN, MYH8, CAMKK2, DAPK1, IQCF2, KCNQ3, CAMK4, KCNN3, SNTB2, PLA2G6, MAP6, ASPM*
Zinc ion binding	*PNMA3, THRA, MMP27, PDLIM2, CAD, ANPEP, ZNF346, GLI2, RNF212, PGLYRP4, GATA6, RBCK1, RBM10, VPS11, ZCCHC7, USP13, PRKCA, PHRF1, SP100, ASTL, CIZ1, MTA2, MICAL3, UBR3, HERC2, MORC1, MMP12, ZSWIM5, TRIM37, ADAMTS7, ADAMTS6, ZDHHC16, ADAMTS8, KDM2A, BAZ1B, BHMT, CPXM1, RNF25, ZFPM2, PIAS2, CPSF4, TRIM64C, ADAMTSL1, ADAMTS16, ABLIM3, USP5, WTIP, CXXC1, TRIM73, MYCBP2, TCF20, TCEA3, C19ORF68, DMD, PIKFYVE, GATAD2B, ADAMTS12, CHD4, DTNA, CPA5, PTGR1, SHPRH, ZMYM1, DTX3, SKI, TRIM62, TRIM60, CAPN15, TRIM56, ZFHX4, RASSF1, RNF2, MEX3C, DNMT1, ZNF385B, ZRANB3, ZFHX2, PHF6, HDAC6*
Protein kinase activity	*CDK19, FASTKD1, FASTKD2, PRPF4B, TBK1, NEK2, CAD, BRSK1, TTN, CSNK2A1, CDK13, AATK, PRKCA, OBSCN, BRAF, TGFBR1, MET, NEK10, SMG1, ATR, NEK11, DAPK1, HUNK, RPS6KA3, TEX14, CAMK4, NEK8, BUB1B, JAK2*
Protein domain specific binding	*DBNL, PHRF1, HMGB2, OSBP, THRA, SP100, E2F4, FOXA3, MYO1D, CNOT1, SKI, CENPJ, CEP250, RFC1, RCC2, NEDD4, TDG, PKD1, VPS11*
Extracellular matrix constituent, lubricant activity	*MUC5AC, MUC17, MUC4*
Actin binding	*DBNL, SPTBN5, TRPC6, MYO1C, WASF3, TLN2, CLMN, INPPL1, WASF1, ABLIM3, SPTBN4, MICAL3, MPRIP, SYNE1, SYNE2, NEB, XIRP2, DMD, FHOD3, SNTB2, DST, BCL7B, HDAC6*

© Annals of Translational Medicine. All rights reserved.
http://dx.doi.org/10.21037/atm-21-1477

Table S3 Genomic variants in genes with high-impact mutations in at least 3 fetuses

Case no.	Name	Genomic variant
#1	*CTBP2*	chr10:g.126691552T>G
#4	*CTBP2*	chr10:g.126727724C>16inserts
#5	*CTBP2*	chr10:g.126692019G>A
#9	*CTBP2*	chr10:g.126714705C>56inserts
#10	*CTBP2*	chr10: g.126691552T>G
#1	*CTDSP2*	chr12:g.58217422T>C
#4	*CTDSP2*	chr12:g.58220819A>G
#8	*CTDSP2*	chr12:g.58220801G>T
#9	*CTDSP2*	chr12:g.58220827G>C
#10	*CTDSP2*	chr12:g.58220801G>T
#2	*HLA-DRB5*	chr6:g.32489755GT>G, chr6:g.32497957C>3inserts
#5	*HLA-DRB5*	chr6:g.32489937G>5inserts
#6	*HLA-DRB5*	chr6:g.32497905G>A
#10	*HLA-DRB5*	chr6:g.32489755G>3inserts
#6	*LZTFL1*	chr3:g.45877274A>15inserts
#9	*LZTFL1*	chr3:g.45877262G>6inserts, chr3:g.45877263T>24inserts
#11	*LZTFL1*	chr3:g.45877273C>9inserts
#1	*MUC19*	chr12:g.40880335A>G
#4	*MUC19*	chr12:g.40879082A>T, chr12:g.40881226T>29inserts
#7	*MUC19*	chr12:g.40882872A>G, chr12:g.40882796G>3inserts, chr12:g.40884241A>122inserts
#8	*MUC19*	chr12:g.40881235A>G
#2	*MUC4*	chr3:g.195507082G>3inserts
#3	*MUC4*	chr3:g.195518092G>6inserts
#9	*MUC4*	chr3:g.195510225A>3inserts, chr3:195510228C>47inserts, chr3:195518368C>23inserts
#1	*MUC6*	chr11:g.1018239G>5inserts, chr11:g.1018241T>14inserts
#5	*MUC6*	chr11:g.1017466T>A
#6	*MUC6*	chr11:g.1017466T>A

© Annals of Translational Medicine. All rights reserved.
http://dx.doi.org/10.21037/atm-21-1477

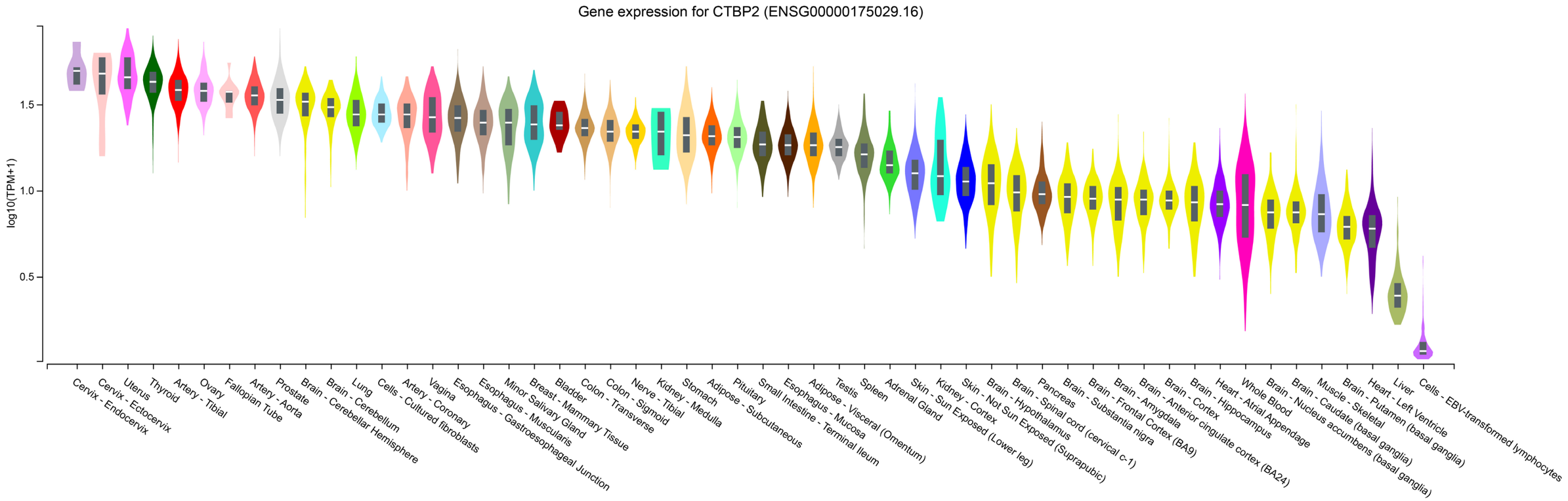


Figure S4 Gene expression levels of *CTDSP2* in different human tissues. The plot was generated by GTEx (https://www.gtexportal.org/home/).

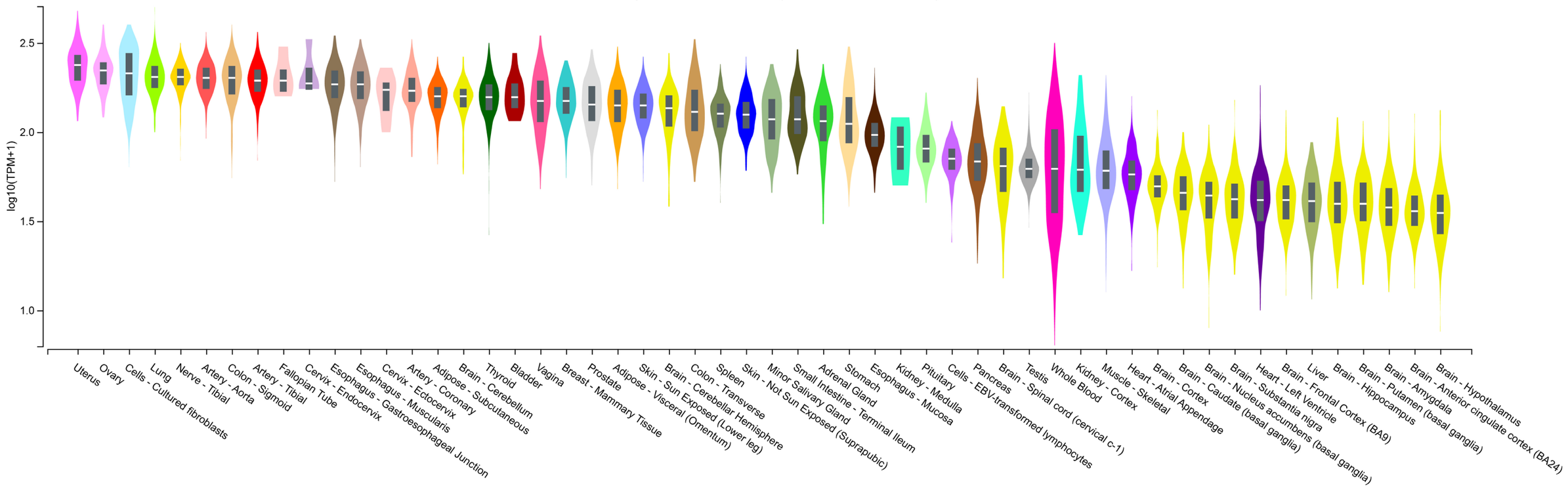


Figure S5 Gene expression levels of *CTBP2* in different human tissues. The plot was generated by GTEx (https://www.gtexportal.org/home/).

© Annals of Translational Medicine. All rights reserved.
http://dx.doi.org/10.21037/atm-21-1477