Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Zhongqi Hou April 10, 2023

Investigation into the effect of RNASEH2B mutation on R-loop dysregulation in the pathogenesis of Aicardi-Goutières syndrome (AGS)

by

Zhongqi Hou

Dr. Bing Yao
Adviser

Biology

Dr. Bing Yao

Adviser

Dr. Peng Jin

Committee Member

Dr. Arri Eisen

Committee Member

2023

Investigation into the effect of RNASEH2B mutation on R-loop dysregulation in the pathogenesis of Aicardi-Goutières syndrome (AGS)

By

Zhongqi Hou

Dr. Bing Yao

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Biology

2023

Abstract

Investigation into the effect of RNASEH2B mutation on R-loop dysregulation in the pathogenesis of Aicardi-Goutières syndrome (AGS)
By Zhongqi Hou

R-loop is a three-stranded structure composed of a DNA:RNA hybrid and a displaced single-stranded DNA. Unscheduled R-loops level will lead to DNA damage and genome instability. To maintain R-loop homeostasis, the RNASEH2 gene can resolve excessive R-loop by cleaving RNA strands from R-loops. Clinically, more than 50% of diagnosed patients with Aicardi-Goutières syndrome (AGS), a severe autosomal recessive inflammatory disease that affects a patient's neurological development when born, have been found to have the mutation in RNASEH2 gene. However, the mechanism of mutations in RNASEH2 in the pathogenesis of AGS remains unclear. RNASEH2B is a subunit of RNASEH2, that we focused on in this study to examine its effect on R-loop regulation. Using the human iPSC-derived neuron method, we successfully characterized iPSC from an AGS patient's PBMC and differentiated it into neurons. An increased R-loop level was shown in the AGS patient. Using DRIP-seq, we profiled genome-wide R-loop and identified differential R-loop peaks in the AGS patient. We also found that upregulated R-loop peaks fall in the intergenic regions and downregulated R-loop peaks fall in the gene body regions. Altogether, our data suggest that mutation in RNASEH2B may cause R-loop dysregulation, ultimately leading to AGS.

Investigation into the effect of RNASEH2B mutation on R-loop dysregulation in the pathogenesis of Aicardi-Goutières syndrome (AGS)

By

Zhongqi Hou

Dr. Bing Yao

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Biology

2023

Acknowledgements

I want to express my deepest appreciation to my advisor, Dr. Bing Yao, for providing me with continuous support and the opportunity to learn in the lab. I would also like to thank my committee members, Dr. Arri Eisen and Dr. Peng Jin. Lastly, I would like to give special thanks to my mentor, Dr. Yingzi Hou, who has supported me throughout my thesis and my time in the lab.

Table of Contents

Introduction……..……1

Results……..……...3

Discussion……...……....12

Methods…….....……..…….....14

References……...…….....17

Introduction

R-loop, also indicated as RNA/DNA hybrid, is formed when RNA hybridizes to a complementary DNA strand which causes displacement of single-stranded DNA (Fig. 1). Studies have shown that R-loops are required for several cellular processes, such as regulating gene expression and attracting repair proteins to DNA double-strand breaks (Sollier & Cimprich, 2015). On the other hand, unscheduled R-loops can cause DNA damage and genome instability by inducing transcription-associated recombination and chromosome breaks/loss (Cristini, *et al*., 2022). R-loops must be kept in an optimal range; when R- loops are low, it causes gene silencing and telomere shortening (Wells, *et al*., 2019). DNA damage and genome instability will occur when the R-loop level is high (Lin, *et al*., 2022). Various genes are found to be associated with R-loop homeostasis, among which RNASEH1 and RNASEH2 can resolve R-loops by directly cleaving RNA strands from R-loops (Wahba, *et al*., 2011). Mutation in RNASEH2 has been reported to cause more than 50% of cases in Aicardi-Goutières syndrome (AGS) (Perrino, *et al*., 2008), which drove us to examine if R-loops might play a role in the pathogenesis of AGS.

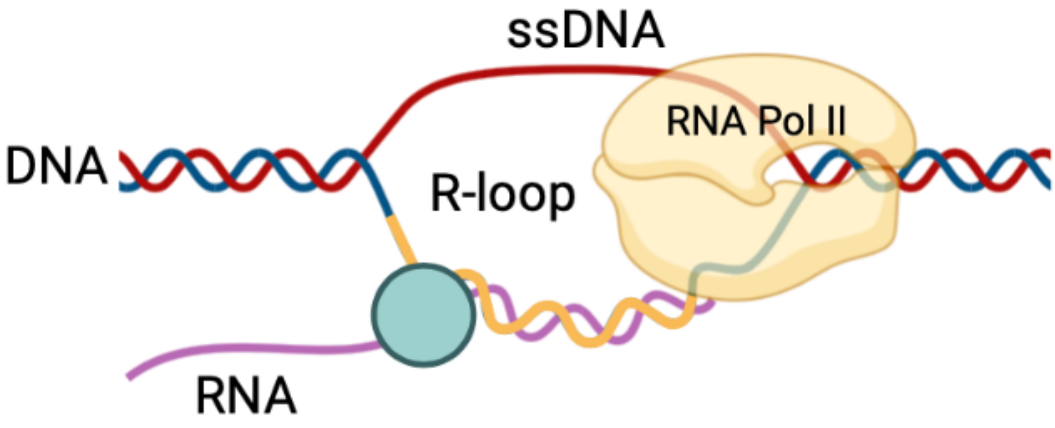


Figure 1. Schematic representation of an R-loop structure (Created by Biorender). The RNA polymerase II (yellow) progresses along the DNA double helix (red and blue). The transcribed RNA (purple) hybridized with the complementary DNA template strand (orange). The non-template strand is shown as Single-stranded DNA (ssDNA).

AGS is a rare and severe familial genetic disorder that affects an individual's brain, skin, and immune system (Crow. *et al*., 2006). The disease could be diagnosed clinically using magnetic resonance imaging (MRI) and cerebral spinal fluid (CSF) testing at an early age. The disease is an autosomal recessive inflammatory encephalopathy that could lead to various physical and mental handicaps (Cristini, *et al*., 2022). Early-onset and later-onset are two forms of AGS that begin in infancy. Patients typically experience intermittent fever, feeding difficulties, seizures, microcephaly, and skin rashes. Clinical studies showed that most AGS patients experience permanent neurological dysfunction with elevated levels of interferon-alpha (Rice, *et al*., 2009).

Mutations in several AGS genes have been identified as causes of AGS, including RNA adenosine deaminase (ADAR), 3' to 5' single-stranded DNA exonuclease (TREX1), RNA/DNA hybrid-specific ribonuclease H2 subunits (RNASEH2A, RNASEH2B, and RNASEH2C), and SAMHD1 (3' to 5' dNTP hydrolase) (Lim, *et al*., 2015). Mutations in any of these genes could trigger the accumulation of unprocessed nucleic acid and DNA:RNA hybrids, ultimately leading to an autoimmune response (Rabe, 2013). More importantly, the mutation in RNASEH2B is the major cause of AGS (Mackenzie, *et al*., 2016).

In the present study, our objective is to investigate how RNASEH2B acts on gene expression dysregulation through excessive R-loop loading, ultimately leading to neurological impairment in AGS patients. By characterizing human iPSC from a patient's PBMC and differentiating it to neuron cells, we can use various advanced methods to

identify at what specific sites those R-loops are formed in the whole genome. Overall, our findings are used to better understand the impact of R-loop accumulation on gene expression, how it leads to genomic instability, and ultimately provide insight into the pathogenesis of AGS.

Results

Characterization of human induced pluripotent stem cell (iPSC) line reprogrammed from human peripheral blood mononuclear cell (PBMC) of an AGS patient

PBMCs collected from a male patient with confirmed AGS disease carrying an RNASEH2B mutation (homo. c529G>A) from Children's Hospital of Philadelphia (AGS) and a healthy control individual (WT) were reprogrammed into iPSC lines by using Sendai virus vectors. The newly generated human iPSCs lines were cultured and then characterized. The cell line showed compact colonies with distinct borders and large nucleus presenting typical iPSC morphology (Fig. 2a). The expression of pluripotency markers is indicated by positive immunofluorescence staining for iPSC marker genes, including NANOG and OCT4, for both WT and the AGS patient (Fig. 2b). We performed PCR and sanger sequencing to validate the mutation of the AGS patient iPSC line. The PCR product was indicated in 563 base pair amplicons (Fig. 2d) and compared with WT iPSC line, the AGS iPSC line showed a homozygous c.529G>A mutation in exon 7 of RNASEH2B gene (Fig. 2c). Together, these results provide evidence of successfully characterized iPSC lines and confirmed mutation of the AGS patient.

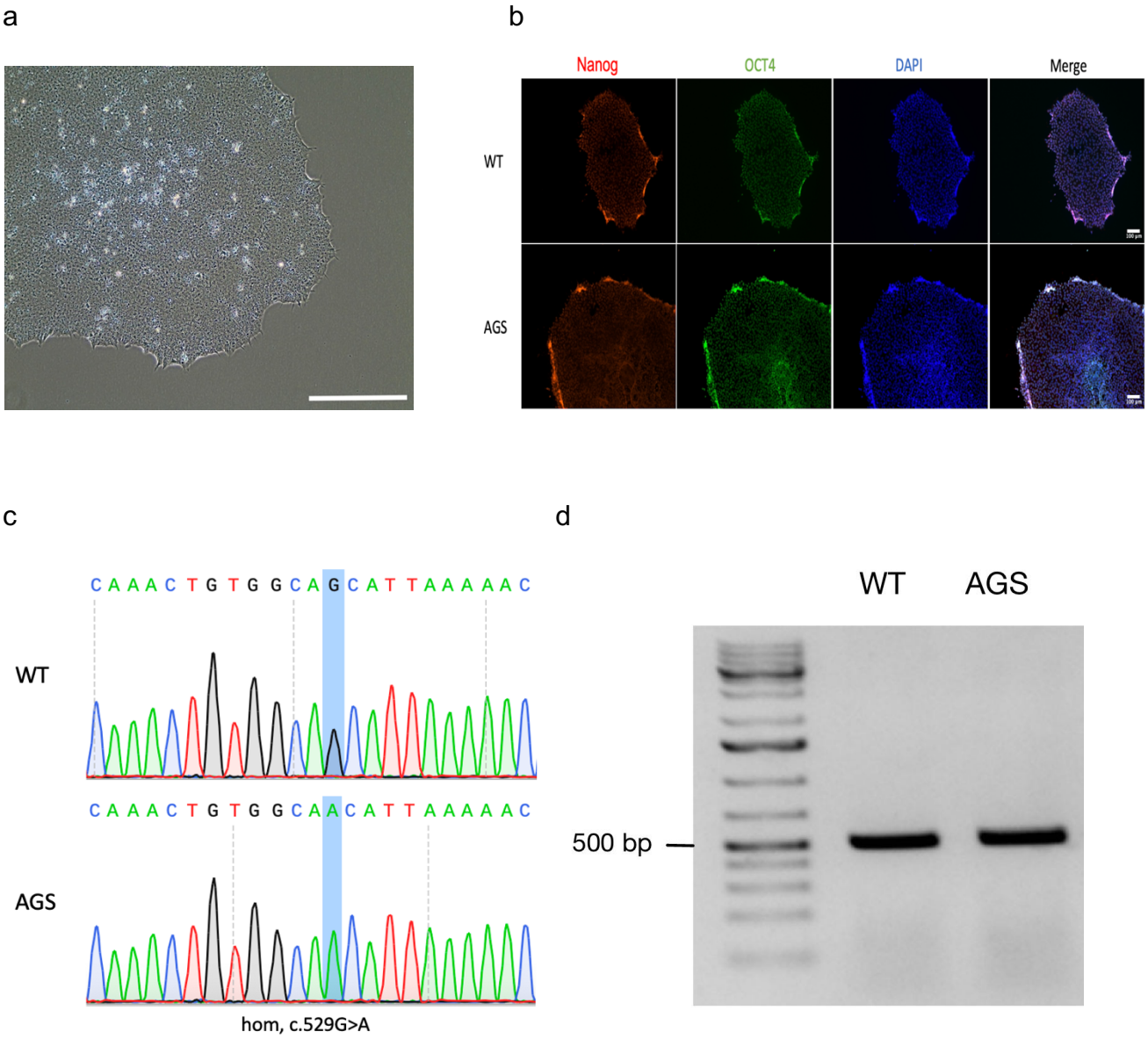
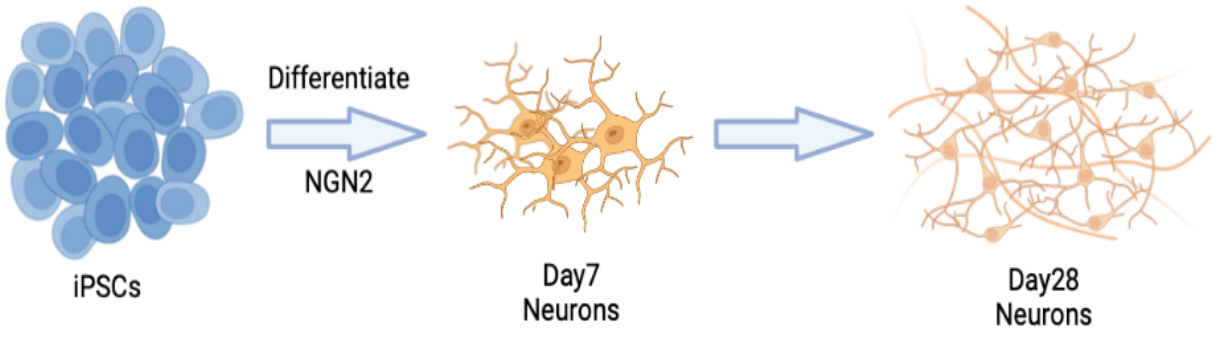


Fig. 2. Characterization of human iPSC line and disease validation. (a) Bright field image of cultured 28 days AGS patient iPSC line (Scale bar 100 μm). (b) Expression of iPSC markers Nanog (red), OCT4 (green) and DAPI (blue) in WT and the AGS patient 28 days iPSC by immunofluorescence staining (Scale bar 100 μm). (c) Electropherograms showing the compound homozygous c.529G>A mutation in exon 7 of RNASEH2B gene (highlighted blue box). (d) Verified mutation in 563 base pair amplicons by PCR.

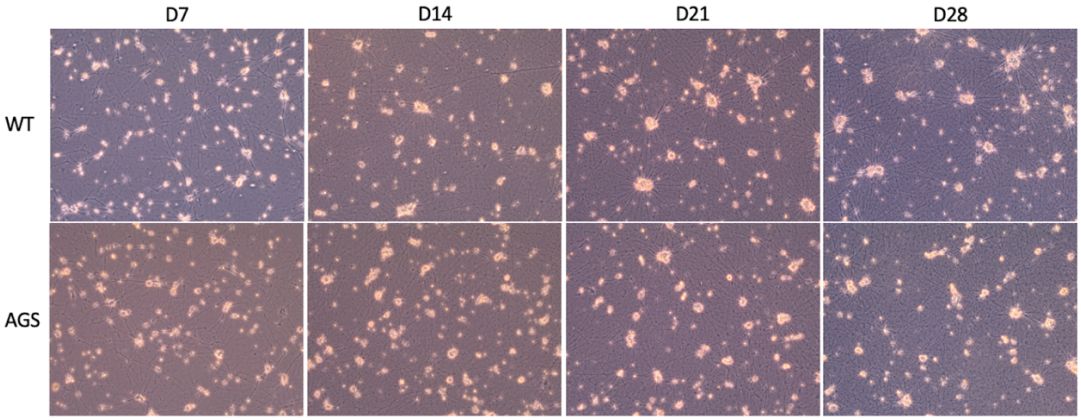
Accumulation of R-loop in differentiated iPSC-derived neuron of AGS patient

WT and the AGS patient iPSC lines were differentiated using a standardized 28-day direct induced neuron protocol that induced lentivirus expression of neurogenin 2 (NGN2) (Fig. 3a). Typical neural morphology was observed at day 7 and the neurons continued to grow to day 28 (Fig. 3b). Differentiating cells expressed typical nerve markers including MAP2 and TUJ1, indicating committed neuronal lineage (Fig. 3c & 3d). Interestingly, we used western blot to look at protein levels between WT and the AGS patient, and there were no differences presented (Fig. 4a). R-loops were quantified using dot-blots assay with S9.6 antibody, and the result showed increased R-loop level in AGS (Fig. 4b-4d). Hence, iPSC-derived neurons were successfully differentiated, and similar RNASEH2B expression and accumulation of R-loop was found in the AGS neurons.

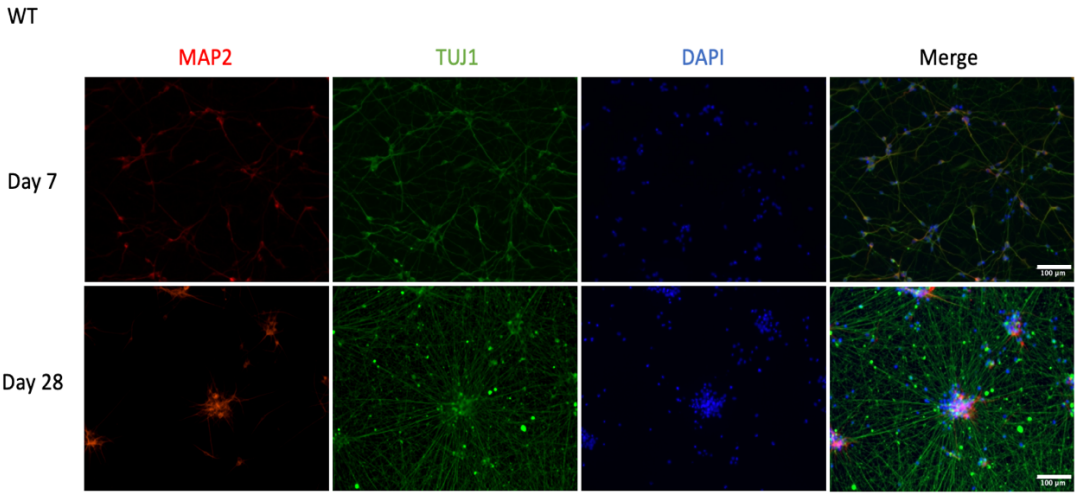
a



b


D7
D14
D21
D28
WT
AGS

c


WT
MAP2
TUJ1
DAPI
Merge
Day 7
Day 28

d

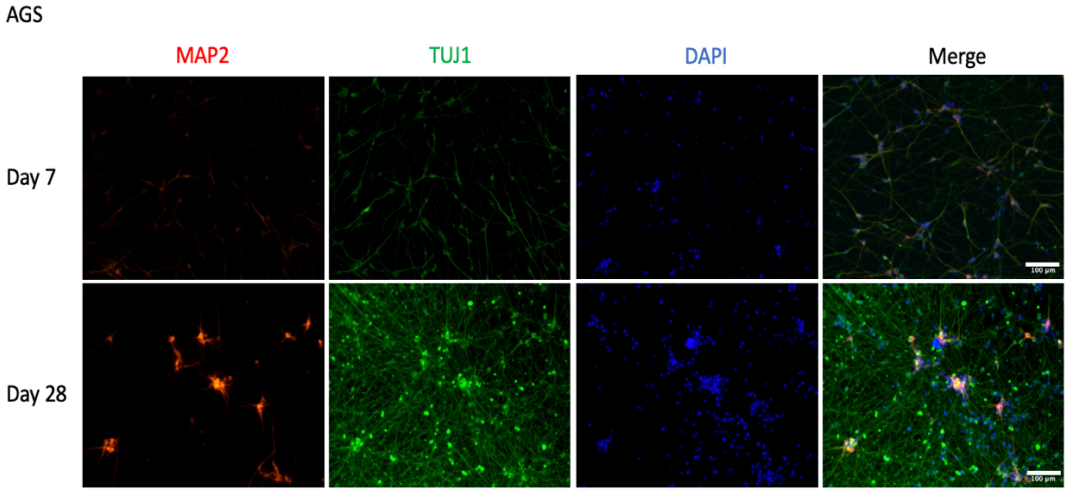

AGS
MAP2
TUJ1
DAPI
Merge
Day 7
Day 28

Figure 3. 28 days iPSC-derived neuron differentiation. (a) Schematic representation of 28 days neuron differentiation from iPSC line using induced Ngn2. (b) Bright field images of cultured neurons from day 7 to day 28 of WT and the AGS patient. (c-d) Immunofluorescence staining of WT and AGS patient neurons (day 7 and day 28) expressed neuronal marker MAP2 (red), TUJ1 (green), and DAPI (blue) (Scale bar 100 µm).

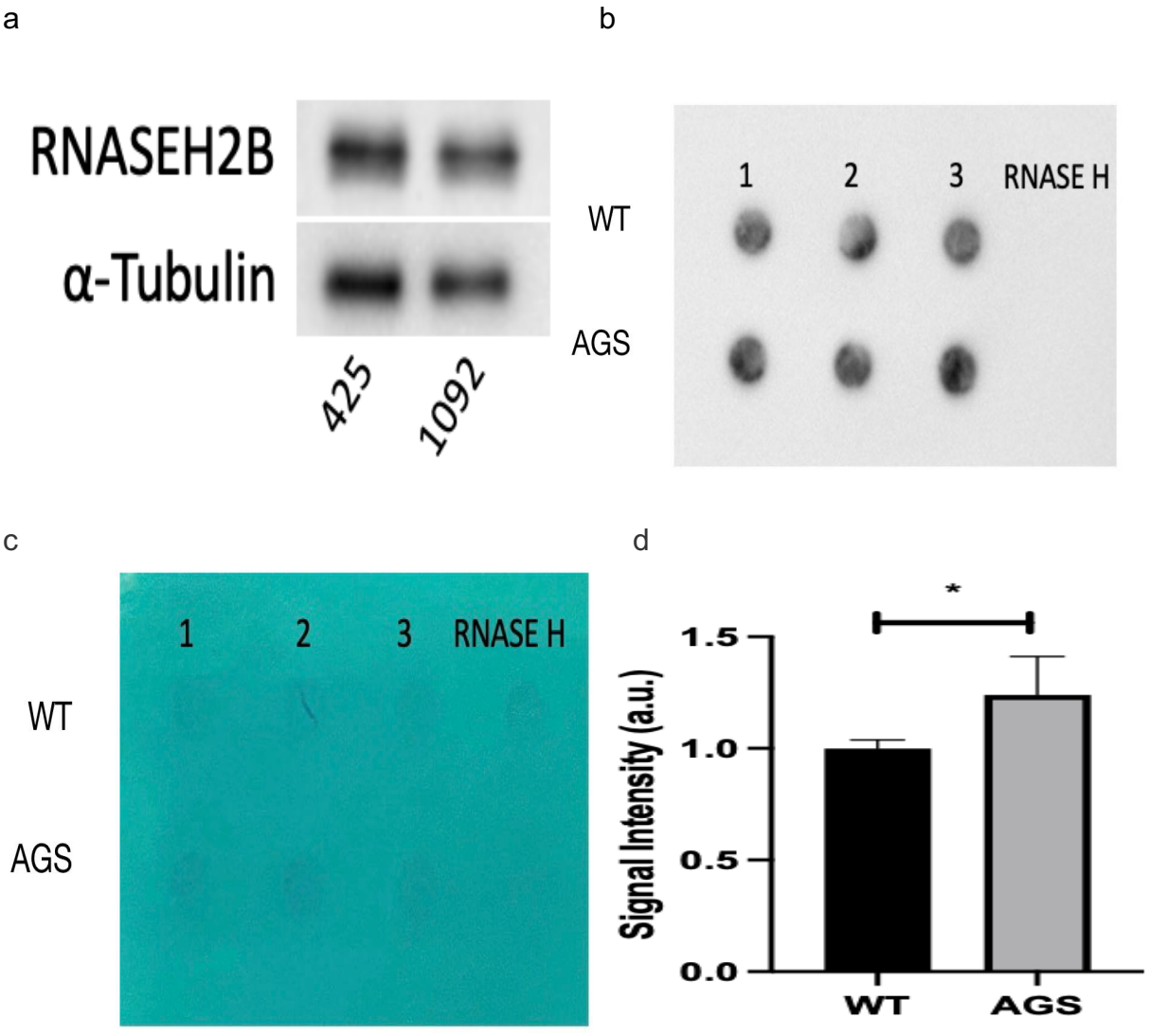
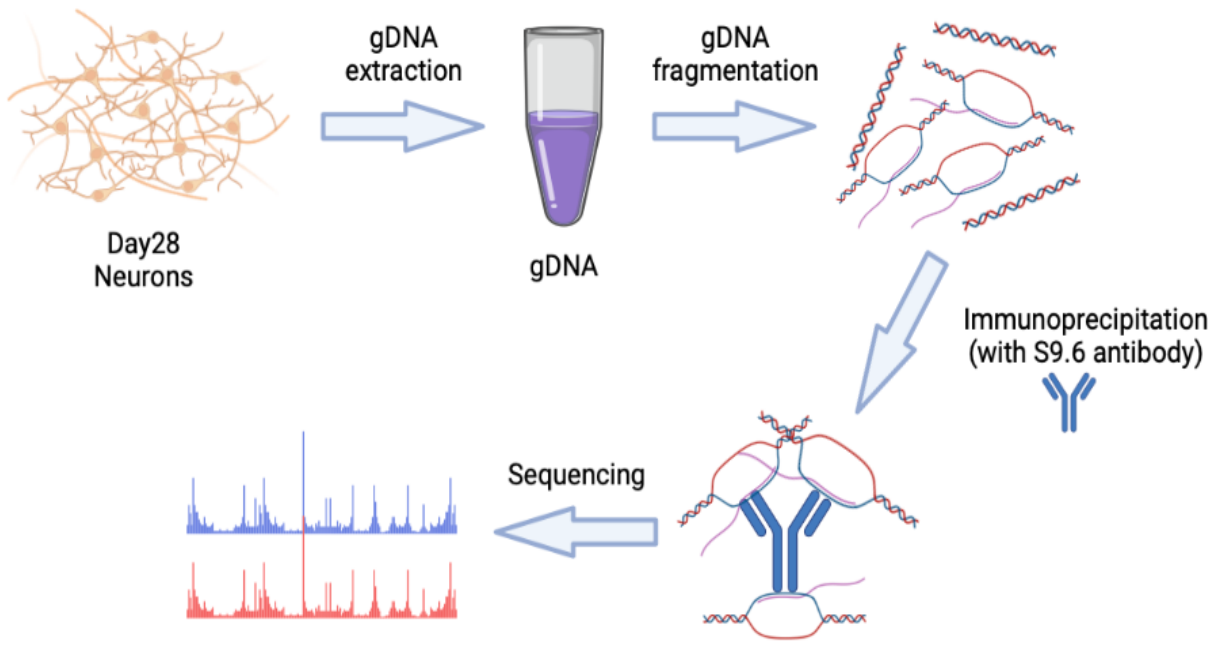


Figure 4. Compared protein level and R-loop level in WT and the AGS patient. (a) Western blot indicating the same protein level between WT and the AGS patient. α-Tubulin was used as loading control. (b-d) Dot blot assay showing higher R-loop concentration in the AGS patient than WT with equal loading sample in each trial.

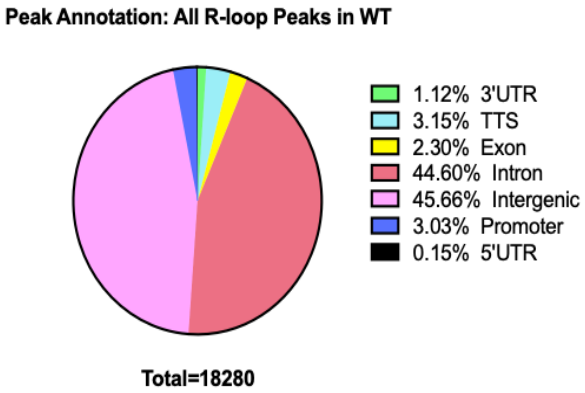
DNA:RNA immunoprecipitation sequencing (DRIP-seq) illustrates upregulated and downregulated R-loop peak between WT and the AGS patient

DRIP-seq, a technique initially developed to profile genome-wide R-loops, was performed to identify and compare R-loop peaks between WT and the AGS patient (Fig. 4a). We identified 18,280 R-loop peaks in the WT patient and 12,345 R-loop peaks in the AGS patient. The genomic distribution of all R-loop peaks in WT and AGS were revealed by HOMER peak annotation. A majority of R-loop peaks fall in intron regions (44.3%) and intergenic regions (46.43%) (Fig. 4b and 4c). Compared to WT, the R-loop peaks of the AGS patient increased in the intergenic region (45.66% to 47.20%) and decreased in both the promoter region (3.03% to 2.75%) and exon region (2.30% to 2.06%) (Fig. 4b and 4c).

a



b



c

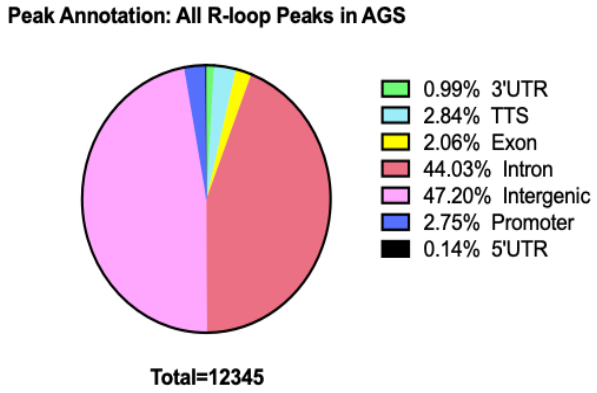
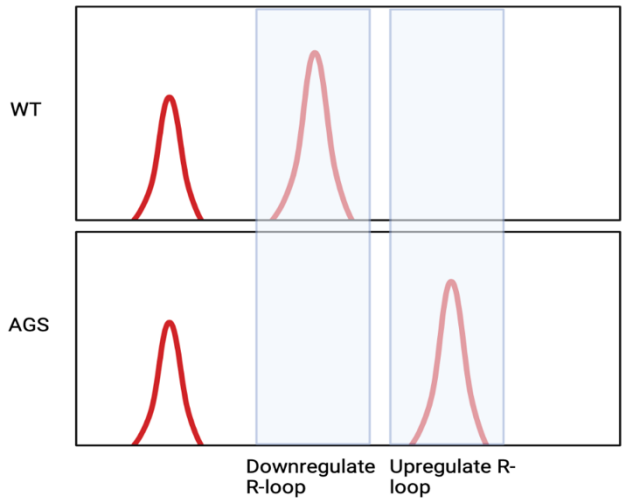


Figure 5. Drip-seq and peak annotation in WT and the AGS patient. (a) Schematic representation of Drip-seq protocol. (b) HOMER peak annotation pie chart shows more than half of peaks of all R-loop peaks in WT falling into the intergenic region (pink, 45.66%) and intronic region (red, 44.6%). Only 3.03% of peaks fall in the promoter region (dark blue). (c) HOMER peak annotation pie chart shows more than half of peaks of all R-loop peaks in the AGS patient falling into the intergenic region (pink, 47.2%) and intronic region (red, 44.03%). Only 2.75% of peaks fall in the promoter region (dark blue).

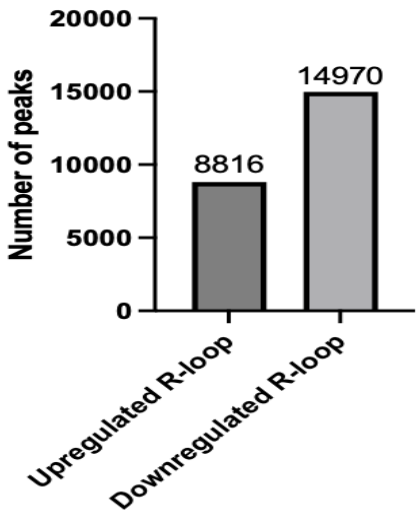
To determine the R-loops dysregulation in the AGS patient, we called differential R-loop peaks between WT and AGS. If the peak was present in WT but not in the AGS patient, it was identified as downregulated R-loop peak; conversely, upregulated R-loop peak was shown in the AGS patient but absent in WT (Fig. 6a). Across all detected R-loop peaks between WT and AGS, a total of 8816 peaks were upregulated, and a total of 14970 peaks were downregulated (Fig. 6b). Interestingly, there were more upregulated R-loop peaks in the intergenic region and more downregulated peaks fall in the exon and promoter region (Fig. 6c & 6d). We then used genome-wide R-loop signal profiling for WT and the AGS patient to confirm upregulated and downregulated R-loop peaks. We counted R-loop reads over the upregulated and downregulated R-loop peak regions using

ngsplot. For upregulated R-loop peaks, the results revealed higher reads in the AGS patient compared to WT (Fig. 6e and 6f). Conversely, higher reads were shown in WT for downregulated R-loop peaks than in the AGS patient (Fig. 6g and 6h). Gene ontology analysis for all downregulated R-loop peaks shows the most significantly enriched terms related to neuron differentiation and DNA repair (Fig. 7).

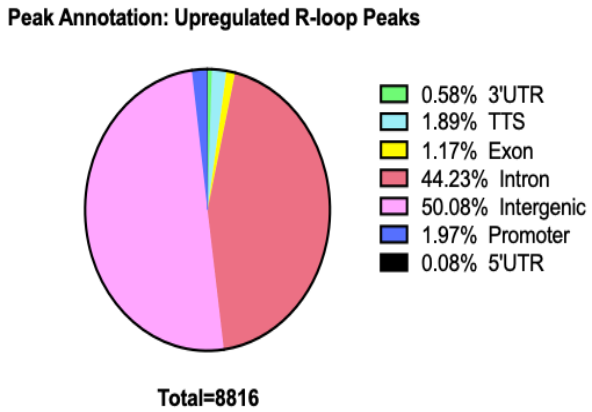
a



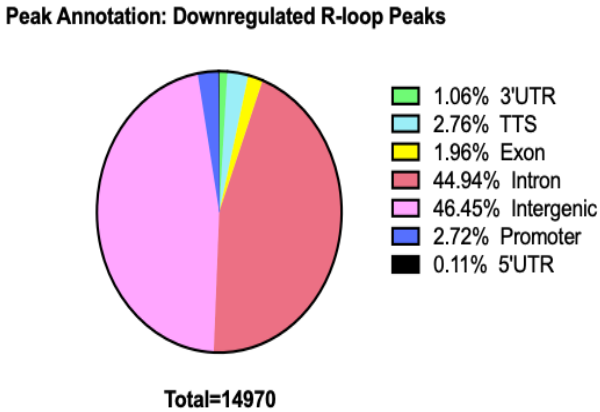
b



c



d



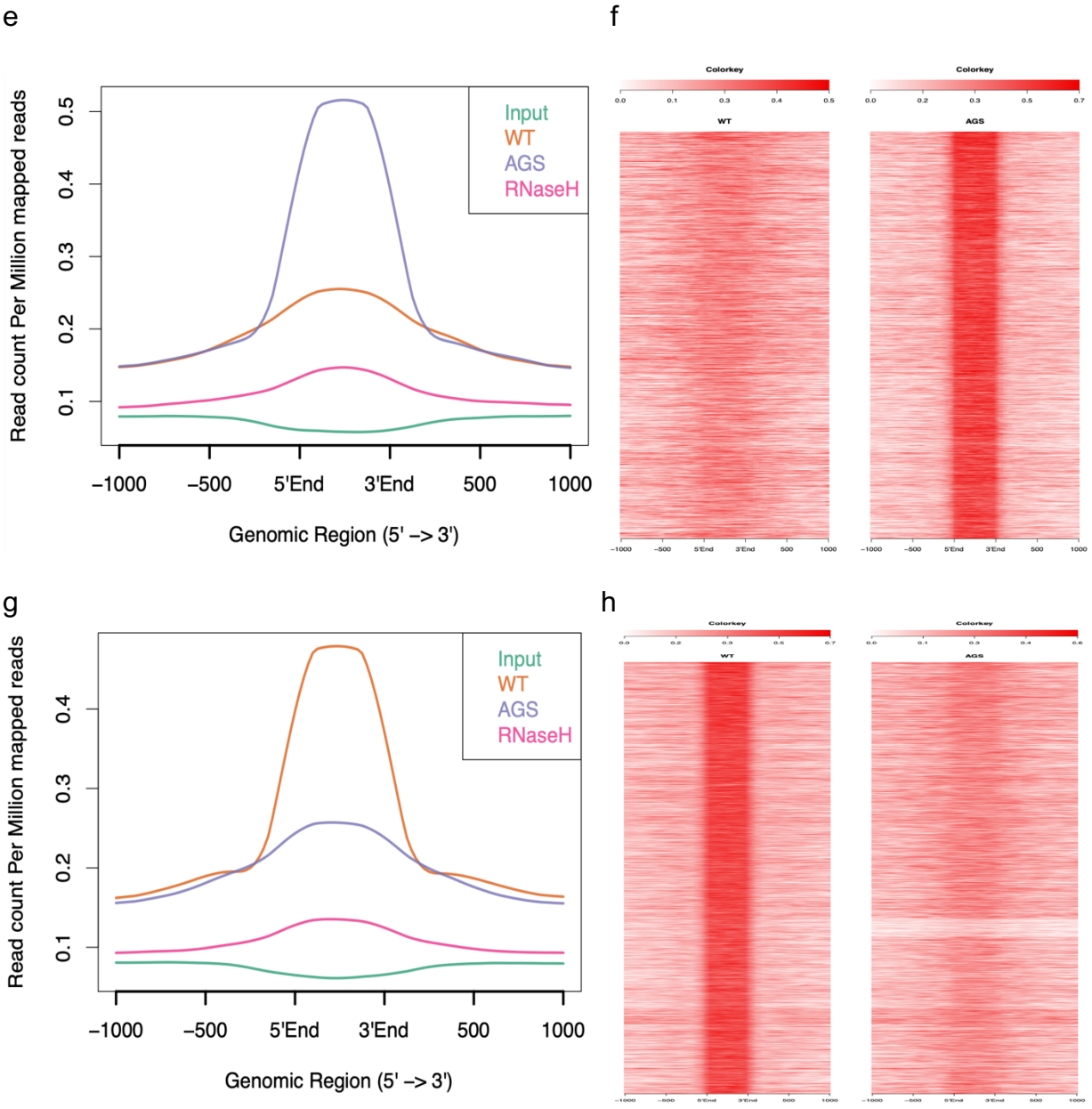


Figure 6. Upregulated and downregulated R-loop peaks. (a) Schematic diagram representing how upregulated and downregulated R-loops peaks were identified. (b) Bar graph shows upregulated R-loop peaks (8816) and downregulated R-loop peaks (14970) in WT and the AGS patient. (c-d) Ngs.plot represents higher reads in the AGS patient (purple) than WT (orange) for upregulated R-loop peaks which correspond with the result in the heat map. (e-f) Ngs.plot represents higher reads in WT than the AGS patient for downregulated R-loop peaks which correspond with the result in the heat map.

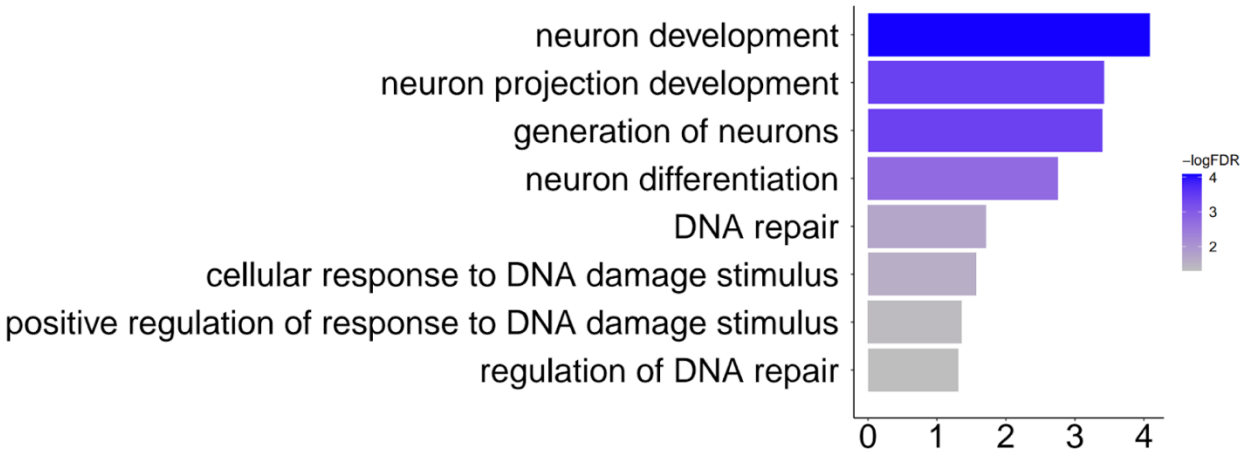


Figure 7. Gene ontology for all downregulated R-loop peaks.

Discussion

In this study, we aimed to use the human iPSC-derived neuron model to understand how the mutation in RNASEH2B could affect R-loop dysregulation in AGS patients. In the past, it has been challenging to study neurological disorders in the laboratory due to the need for predictive models to study the pathogenesis of disease. Previous studies have used genetically engineered animal models; however, it is difficult to replicate when looking at various mutations that are causing the disease. The mutations identified for neurologic diseases, like AGS, also have contributions from other genetic, epigenetic, or environmental factors (Vassos, *et al*., 2010). So, using human iPSCs allows researchers to study the effect of various human genetic backgrounds on disease penetrance and severity (Dolmetsch & Geschwind, 2011). Furthermore, we believe AGS disease affects a patient's neuron development, so using human iPSC-derived neurons allows us to target specific human brain development. Some clinical studies realized the newly discovered treatment that works in animal models was ineffective in human patients due to different electrophysiological phenotype between humans and other animals

(Steffenhagen, *et al*., 2011). Thus, our result used successful differentiated iPSC-derived neuron reproducibility which allows us to look at the effect of R-loop level in the AGS patient (Fig 3).

Interestingly, we did not see an expression change in the RNASEH2B gene between our healthy control and the AGS patient as we first looked at protein levels (Fig. 4a). However, increased R-loop level observed in the AGS patient (Fig. 4b) indicates a failure to remove accumulated R-loop during transcription (Cristini, *et al*., 2022). Therefore, we focused on where the R-loop peaks fall in the whole genome. The significant observation from DRIP-seq profiling reveals a major accumulation of upregulated R-loop peaks falls in intergenic regions and away from gene body regions (exon and promoters); downregulated R-loop peaks, in contrast, fall in gene body regions (Fig 6c & 6d). Enriching upregulated R-loop peaks in the intergenic region could cause DNA damage (Richard & Manley, 2017). The accumulation of downregulated R-loop peaks in exon and promoter regions leads to binding alteration of transcription factors (Allison & Wang, 2019), potentially lowering gene expression, failing to repair DNA damage, and decreasing neuron differentiation corresponding to our gene ontology result (Fig, 7).

As mentioned before, RNASEH2B is a subunit of RNASEH2 (Lim, *et al*., 2015). RNASEH2 is a complex that contains RNASEH2A, RNASEH2B, and RNASEH2C. This RNASEH2 complex structure was found in previous study to define critical interaction interfaces relevant to various human diseases and the function of enzyme (Reijns, *et al*., 2011). We predicted two possible mechanisms of how mutations in RNASEH2B could

lead to R-loop accumulation. The first one is when a mutation occurs in RNASEH2B; it causes complex formation defects in which RNASEH2B cannot bind with RNASEH2A and RNASEH2C, ultimately failing to remove R-loops. Another mechanism we predict is when mutations occur in RNASEH2B, it will not affect the RNASEH2 complex. However, it changes the genomic binding regions, which fail to recognize R-loops.

In conclusion, our study illustrated how human iPSC-derived neuron model was a powerful tool for investigating neurological disease pathogenesis. We identified the accumulation of R-loop in the AGS patient, which led to DNA damage and neurological impairment. We also predicted two mechanisms for the causation of R-loop accumulation for future study.

Methods

iPSC characterization and neuron differentiation

The ethics committee approved this study, and informed consent was obtained from the patient. Our lab successfully reprogrammed PBMC samples that were collected from a male patient from the Children's Hospital of Philadelphia to iPSC cell lines using Sendai virus vectors. All cell culture procedures were performed in the safety cabinet using an aseptic technique to ensure sterility. Established iPSC lines were cultured in a Matrigel-coated plate for 28 days. The cells were cultured at 37°C, and the culture medium (mTeSR1) was replaced every day. Successfully characterized iPSCs were differentiated into neurons based on a standardized 28-day direct Ngn2-induced neuron protocol (Zhang, *et al*., 2013). Before differentiation, iPSC cells were infected with NGN2-expressing lentivirus. Starting from day 0, cells were split and plated to around 200,000

cells/cm^2 with mTeSR and ROCKi (Y-27632). KSR media and Doxycycline were added to induce Ngn2 expression. Dead cells that were not expressing NGN2/puro were washed off. On day 4, Neurons were replated with 3 mls of NBM, B27, and ROCKi. Media were changed every 4 days until day 28.

Immunofluorescence staining

iPSC and neurons were plated onto glass chamber slides coated with Matrigel to ensure strong adhesive. Cells were then fixed with 1 ml of 4% paraformaldehyde in PBS, permeabilized using 0.5% Triton X-100 and blocked with 5% goat serum at room temperature. Primary antibodies (anti-OCT4 & anti-Nanog for iPSC, anti-MAP2 & anti TUJ1 for neuron) were added to the cells and incubated overnight at 4°C. The cells were then incubated with secondary antibodies (anti-mouse IgG and anti-rabbit IgG) after washing and co-stained with DAPI. ImageJ was used to analyze morphology.

Western blot

Cultured 28 days neurons were collected, washed in PBS, and lysed in RIPA buffer. Protein concentration was measured using Nanodrop. The sample with loading dye were loaded in 8% polyacrylamide gel and run for an hour. The gel was then transferred to the PVDF membrane using the Trans-Blot Turbo Transfer System (Biorad). Primary antibodies (anti-RNASEH2B and anti-α-tubulin) were added after blocking the membrane with 5% skim milk in TBST for one hour and incubated on a shaker overnight (4°C). A secondary antibody was added in 5% skim milk in TBST after 3 times washing with TBST

and shaking for 1 hour. After washing three times with TBST, signals were detected in the dark chamber UVITEC Cambridge.

Dot blot assay

Neuron cells were washed with PBS twice, removed from the cell culture dish to 1.5 mL tube, and lysed with a cell lysis buffer and proteinase K overnight at 37℃. The gDNA was isolated using phenol:chloroform:isoamyl ethanol and precipitated in 100% ethanol. gDNA was fragmented using a cocktail of restriction enzymes (BamHI, EcoRI, HindIII, SspI, XbaI) and purified using phenol:chloroform:isoamyl ethanol. Next, the purified DNA was added with 2 U RNaseH per microgram of gDNA for 4 hours as negative control. gDNA and RNase H-treated gDNA were loaded on a nylon membrane and crosslinked using a UV linker. 5% skim milk was used to block the nylon membrane, and the S9.6 antibody (1:1000) was then added and incubated overnight. Secondary antibodies were loaded onto the membrane and incubated for one hour. After detection of the S9.6 signal, methylene blue was added to stain and determine loaded total gDNA. ImageJ was used to analyze and quantify the level of the R-loop.

DRIP-seq

After genomic extraction from the 28-day neuron, the gDNA was incubated with S9.6 antibody on a rotator at 4℃ overnight. In this process, S9.6 antibody binds to the specific R-loop region (DNA-RNA hybrids) across the genome. The antibody-DNA complexes were bound to magnetic beads and washed three times to remove any left gDNA that was not attached to the beads. R-loops were then eluted using SDS and purified using

phenol:chloroform:isoamyl ethanol. Immunoprecipitated R-loops were treated with RNase H and then sonicated to ~300 bp. Sonicated DNA was end-repaired adaptor-added. PCR was performed to add indexes to the library of the samples. The library was analyzed by bioanalyzer and subjected to sequencing.

Bioinformatics analysis

DRIP-seq reads were aligned to human genome sequence (hg38) by Bowtie2 version 2.4.4 with default parameter. Aligned reads in the bam file was sorted by genomic coordinates by samtools. Peaks for each library were identified by MACS2 version 2.2.6 with input bam file. Differential peaks were identified using bedtools (version 2.30.0) intersect. Peaks were annotated to genes by annotatePeaks.pl from HOMER version 4.11.

References

Allison, D. F., & Wang, G. G. (2019). R-loops: Formation, function, and relevance to cell stress. *Cell Stress*, *3*(2), 38–46. doi: 10.15698/cst2019.02.175.

Cristini, A., Tellier, M., Constantinescu, F., Accalai, C., Albulescu, L. O., Heiringhoff, R., Bery, N., Sordet, O., Murphy, S., & Gromak, N. (2022). RNase H2, mutated in aicardi-goutières syndrome, resolves co-transcriptional R-loops to prevent DNA breaks and inflammation. *Nature Communications*, *13*(1). doi: 10.1038/s41467-022-30604-0.

Crow, Y. J., Leitch, A., Hayward, B. E., Garner, A., Parmar, R., Griffith, E., Ali, M., Semple, C., Aicardi, J., Babul-Hirji, R., Baumann, C., Baxter, P., Bertini, E., Chandler, K. E., Chitayat, D., Cau, D., Déry, C., Fazzi, E., Goizet, C., … Jackson, A. P. (2006). Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutières syndrome and

mimic congenital viral brain infection. *Nature Genetics*, *38*(8), 910–916. doi: 10.1038/ng1842.

Dolmetsch, R., & Geschwind, D. H. (2011). The human brain in a dish: The promise of ipsc-derived neurons. *Cell*, *145*(6), 831–834. doi: 10.1016/j.cell.2011.05.034.

Lim, Y. W., Sanz, L. A., Xu, X., Hartono, S. R., & Chédin, F. (2015). Genome-wide DNA hypomethylation and RNA:DNA hybrid accumulation in aicardi–goutières syndrome. *ELife*, *4*. doi: 10.7554/eLife.08007.

Lin, W.-L., Chen, J.-K., Wen, X., He, W., Zarceno, G. A., Chen, Y., Chen, S., Paull, T. T., & Liu, H.-wen. (2022). DDX18 prevents R-loop-induced DNA damage and genome instability via PARP-1. *Cell Reports*, *40*(3), 111089. doi: 10.1016/j.celrep.2022.111089.

Mackenzie, K. J., Carroll, P., Lettice, L., Tarnauskaitė, Ž., Reddy, K., Dix, F., Revuelta, A., Abbondati, E., Rigby, R. E., Rabe, B., Kilanowski, F., Grimes, G., Fluteau, A., Devenney, P. S., Hill, R. E., Reijns, M. A. M., & Jackson, A. P. (2016). Ribonuclease H2 mutations induce a cGAS/STING-dependent innate immune response. *The EMBO Journal*, *35*(8), 831–844. doi: 10.15252/embj.201593339.

Perrino, F. W., Harvey, S., Shaban, N. M., & Hollis, T. (2008). RNASEH2 mutants that cause Aicardi–GOUTIERES syndrome are active nucleases. *Journal of Molecular Medicine*, *87*(1), 25–30. doi: 10.1007/s00109-008-0422-3.

Rabe, B. (2013). Aicardi–Goutières Syndrome: Clues from the RNase H2 knock-out mouse. *Journal of Molecular Medicine*, *91*(11), 1235–1240. doi: 10.1007/s00109-013-1061-x.

Reijns, M. A. M., Bubeck, D., Gibson, L. C. D., Graham, S. C., Baillie, G. S., Jones, E. Y., & Jackson, A. P. (2011). The structure of the human RNase H2 complex defines key interaction interfaces relevant to enzyme function and human disease. *Journal of Biological Chemistry*, *286*(12), 10530–10539. doi: 10.1074/jbc.m110.177394.

Richard, P., & Manley, J. L. (2017). R loops and links to human disease. *Journal of Molecular Biology*, *429*(21), 3168–3180. doi: 10.1016/j.jmb.2016.08.031.

Rice, G. I., Bond, J., Asipu, A., Brunette, R. L., Manfield, I. W., Carr, I. M., Fuller, J. C., Jackson, R. M., Lamb, T., Briggs, T. A., Ali, M., Gornall, H., Couthard, L. R., Aeby, A., Attard-Montalto, S. P., Bertini, E., Bodemer, C., Brockmann, K., Brueton, L. A., … Crow, Y. J. (2009). Mutations involved in Aicardi-Goutières syndrome implicate SAMHD1 as regulator of the innate immune response. *Nature Genetics*, *41*(7), 829–832. doi: 10.1038/ng.373.

Sollier, J., & Cimprich, K. A. (2015). Breaking bad: R-loops and genome integrity. *Trends in Cell Biology*, *25*(9), 514–522. doi: 10.1016/j.tcb.2015.05.003.

Steffenhagen, C., Kraus, S., Dechant, F.-X., Kandasamy, M., Lehner, B., Poehler, A.-M., Furtner, T., Siebzehnrubl, F. A., Couillard-Despres, S., Strauss, O., Aigner, L., & Rivera, F. J. (2011). Identity, fate and potential of cells grown as neurospheres: Species matters. *Stem Cell Reviews and Reports*, *7*(4), 815–835. doi: 10.1007/s12015-011-9251-9.

Wahba, L., Amon, J. D., Koshland, D., & Vuica-Ross, M. (2011). RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. *Molecular Cell*, *44*(6), 978–988. doi: 10.1016/j.molcel.2011.10.017.

Wells, J. P., White, J., & Stirling, P. C. (2019). *R loops and their composite cancer connections*. Trends in Cancer. Retrieved April 8, 2023. doi: 10.1016/ j.trecan.2019.08.006.