JOURNAL OF
EMERGING INVESTIGATORS

Article

Impact of TCERG1 SNP on gene expression and protein interactome in Huntington's disease

Sangmin Song[1, 2, 3], Yiyoung Kwon[2, 3], Kiuk Lee[2, 3], Ihn Sik Seong[2, 3]

[1] Phillips Exeter Academy, Exeter, New Hampshire

[2] Molecular Neurogenetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, Boston, Massachusetts

[3] Department of Neurology, Harvard Medical School, Boston, Massachusetts

SUMMARY

Huntington's disease (HD) is a genetic neurodegenerative disorder that usually presents symptoms in adulthood. An expansion of the CAG repeat in the *huntingtin* gene (*HTT*) creates a mutated huntingtin protein, subsequently causing disease phenotypes. However, the disease age of onset is significantly influenced by single nucleotide polymorphisms (SNPs) across a patient's genome. In this study, we investigated the role of the SNP rs79727797 located within *TCERG1* (Transcription Elongation Regulator 1), a well-known huntingtin interactor, at the chromosome 5 locus 5BM1. We hypothesized that this SNP would influence gene expression, alter protein interactions, and affect cell growth. Using lymphoblastoid cell lines (LCLs) from HD patients with and without SNP rs79727797, we first assessed the effect of rs79727797 on TCERG1 and huntingtin protein levels and found no significant changes. Next, we optimized an immunoprecipitation protocol to isolate TCERG1 and any interacting proteins, allowing for a detailed analysis of potential SNP-driven changes in TCERG1 interactions. This study lays the foundation for future research studying how SNP rs79727797 modifies HD onset, including potential alterations to the TCERG1 interactome.

INTRODUCTION

Huntington's disease (HD) is a neurodegenerative disease that causes cell degeneration and death in the striatum and other parts of the brain (1). It is a devastating and fatal disorder for which there is no cure or effective treatment (1). This can lead to motor disturbances, such as chorea and hypokinesia; behavioral changes, such as emotional apathy and depression; and even cognitive deficits, such as dementia (1). After the onset of symptoms, for which the mean age is 45, patients become increasingly dependent on caregivers until death about 18 years later (1). The cause of this disease is a genetic mutation in the *huntingtin* gene (*HTT*) on chromosome 4 (2). *HTT* has a CAG trinucleotide repeat, or a polyglutamine tract in the huntingtin protein (HTT), that usually extends 10–35 times, but in HD patients it repeats 36–120 times (3). The length of this uninterrupted CAG repeat is inversely correlated to the HD age of onset (AO) (4). While clinical interest has been focused on the mutated *HTT* gene for a potential solution or reversal of symptoms, the CAG repeat length only partially explains AO. Although longer CAG repeats strongly correlate with earlier AO, this correlation accounts for only ~40% of the variability in disease onset (5). The remaining ~60% of variation in onset age cannot be explained solely by the CAG repeats, thus indicating the presence of other modifying factors that act prior to onset (5).

A recent genome-wide association study (GWAS) identified several single nucleotide polymorphisms (SNPs) that correlated with differential age of onset compared to CAG-repeat-based model predictions (6). This suggested that these polymorphisms were a CAG-repeat-independent factor in determining HD AO, directing attention towards the genes containing the top SNPs. One of the candidate HD modifier genes is *Transcription Elongation Regulator 1* (*TCERG1* or *CA150*), located on chromosome 5 at the 5BM1 locus (6). While most of the candidate modifier genes proposed in the GWAS were DNA mismatch repair (MMR) genes, *TCERG1* was not (6). Although MMR genes have been implicated in HD onset via somatic expansion of *HTT* CAG repeats, the presence of other candidate modifiers such as *TCERG1* suggests the possibility of multiple genetic mutations that contribute to variability in disease onset through distinct molecular pathways. As its name suggests, *TCERG1* is involved in transcriptional elongation and pre-mRNA splicing through the interaction with multiple components of both the transcription machinery and the spliceosome (7, 8). The *TCERG1* gene has shown a diversity of functions, with studies highlighting its roles in neuron growth and cellular morphology, carcinoma proliferation, and even human immunodeficiency virus gene expression (9, 10, 11).

The possibility of *TCERG1* playing a significant role in HD onset has been known for decades, with evidence of TCERG1 aggregates in HD neurons and the validation of TCERG1 interaction with HTT (12). Although such preliminary evidence pointed to *TCERG1*'s involvement in HD, the GWAS identified the SNP rs79727797 found in *TCERG1* as the potential genetic link between HTT and HD onset (6). The SNP rs79727797 (chr5:146507273-146507273; ~2–3 years delay on HD motor onset) involves an intronic nucleotide substitution from guanine (G) to adenine (A) in chromosome 5 (13), and its presence correlated with a later onset of HD than predicted by a subject's CAG repeat length (14). There is no strong evidence to suggest that SNP rs79727797 directly affects specific protein domains in *TCERG1*; however, its strong correlation to HD onset warrants further investigation.

In this study, we investigated the connection between SNP rs79727797 and HD onset by focusing on the variant's impact on TCERG1 expression and function. We hypothesized two main mechanisms through which SNP rs79727797 could influence HD AO: 1) a change in gene expression by

https://doi.org/10.59720/25-066

influencing DNA regulatory regions and 2) the alteration of protein function through changes in the coding sequence *of TCERG1*, both of which can disrupt normal cellular processes and contribute to disease development (**Figure 1**). Since SNP rs79727797 is located in an intron of *TCERG1* (6) it may affect gene expression, resulting in altered levels of TCERG1 proteins. Such an event would likely transform the overall interactome of TCERG1 through direct or indirect interactions with TCERG1. The likelihood of a missense mutation or change in the amino acid chain of *TCERG1* as a result of this SNP is minimal, but we cannot exclude any sequence alterations associated with SNP rs79727797 until subsequent investigative studies, such as whole genome sequencing, have been conducted. *TCERG1* also undergoes alternative splicing (7), so the presence of SNP rs79727797 may still be linked to deleterious changes in the protein structure and function.

Previous studies have shown that TCERG1 interacts with HTT and accumulates in HD neurons, suggesting a functional relationship between the two proteins (12). TCERG1 overexpression mitigates neurotoxicity and delays striatal cell death (15). Thus, we hypothesized that the presence of SNP rs79727797 impacts TCERG1, including HTT as an interactor, expression, and/or alters protein interactions. Through mechanisms such as alternative splicing, we believed that these changes to TCERG1 could alter HD onset (7). To test our hypothesis, we first cultured two types of lymphoblastoid cell lines (LCLs) derived from HD patients with and without SNP rs79727797. We examined the effects of SNP rs79727797 on overall cell growth and the expression levels of HTT and TCERG1. Lastly, we optimized the co-immunoprecipitation of TCERG1 and its functional interacting proteins to understand how SNP rs79727797 impacts the TCERG1 protein complex or interactome.

In this study, we found that SNP rs79727797 had minimal effect on LCL growth or the expression of HTT and TCERG1 proteins. Importantly, our optimized immunoprecipitation protocol enabled a specific capture of TCERG1 and its interacting partners while reducing non-specific binding. Although our results suggest that SNP rs79727797 has limited direct effects on TCERG1 or HTT in LCLs, the optimized IP protocol provides a valuable tool for future studies of the TCERG1 interactome, such as a comparative analysis of the TCERG1 interactome in cells with and without SNP rs79727797. While the connection between *TCERG1* SNP rs79727797 and HD onset remains unclear, such further studies may help uncover the downstream mechanisms through which this SNP contributes to variations in HD onset.

RESULTS

Effects of SNP rs79727797 on Lymphoblastoid Cell Growth

Given the association of SNP rs79727797 with HD onset, it is likely that the SNP impacts HD disease phenotypes (e.g. TCERG1's interaction with HTT). Such change(s) in the TCERG1 interactome, combined with previous studies suggesting TCERG1's role in cell growth, could potentially result in SNP rs79727797 affecting cell growth (9). We used LCLs in this study due to their previously identified genetic information for GWAS signals (6). LCLs have frequently been used in HD studies and for investigating HD modifiers, such as Fanconi-Associated Nuclease 1 (*FAN1*) (16). We cultured

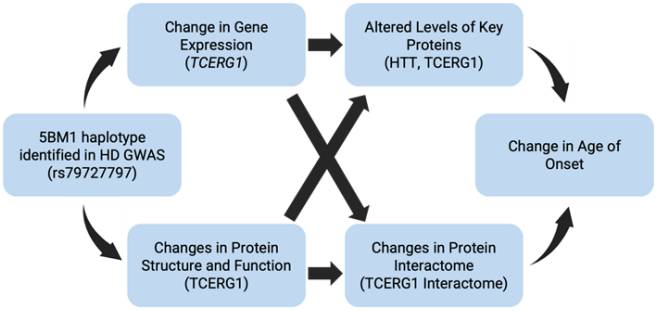

Figure 1: Two potential pathways for GWAS-identified TCERG1 SNPs to influence HD age of onset. Genome-Wide Association Studies (GWAS) have identified multiple Single Nucleotide Polymorphisms (SNPs), including the top SNP rs79727797 which had the strongest association with the age of onset (6). This SNP and other variants in 5BM1 haplotype could alter the expression of Transcription Elongation Regulator (TCERG1) and other related genes. Changes in the expression of these genes could then alter the protein levels of key proteins, such as Huntingtin (HTT) or TCERG1, as well as change TCERG1's protein interactome. Alternatively, a SNP could affect the protein structure and function of protein of interest, such as TCERG1, thus also altering the TCERG1 protein interactome and potentially altering TCERG1 expression. One or both pathways could be the mechanism through which the SNPs directly affect the onset of Huntington's Disease (HD).

four total LCLs derived from HD patients: two control (Con) cells without SNP rs79727797 and two with SNP rs79727797 (SNP).

Over a 21-day period, we measured and compared cell growth. All the cells exhibited a similar latent phase in the first 5 days, followed by logarithmic growth until day 14. Afterwards, the cells often stopped proliferating and some eventually died (**Figure 2A**). This pattern is in accordance with standard culture guidelines of LCL growth (17).

Cell lines LC2 (Con) and LC3 (SNP) had a distinctively faster rate of growth compared to LC1 (Con) and LC4 (SNP). However, the average of two biological replicates in each group showed similar patterns suggesting significant sample-to-sample variability (**Figure 2B**). Unfortunately, we were unable to perform statistical testing due to the small number of biological replicates (n=2). However, the overlapping error bars between the Control and SNP groups suggest no significant difference between the two. This suggests that SNP rs79727797 has no major effect on the growth rate of LCLs. Therefore, despite its significance to *TCERG1* and *HTT*, SNP rs79727797 did not appear to have a major impact on overall cell growth in HD LCLs.

Investigating the Effect of SNP rs79727797 on TCERG1 and HTT Protein Expression

We aimed to determine whether SNP rs79727797 influences TCERG1 and HTT protein expression. SNP rs79727797 and any other GWAS-significant SNPs in 5BM1 are not identified as expression quantitative trait loci (eQTL) for *TCERG1* based on publicly available data (14), suggesting no direct regulatory effect on *TCERG1* mRNA expression. However, it is possible that SNP rs79727797 may affect TCERG1 protein expression through rare deleterious variants (18). Additionally, protein-protein interactions are known to influence protein stability and expression independent of mRNA changes. For instance, knocking down or knocking

https://doi.org/10.59720/25-066

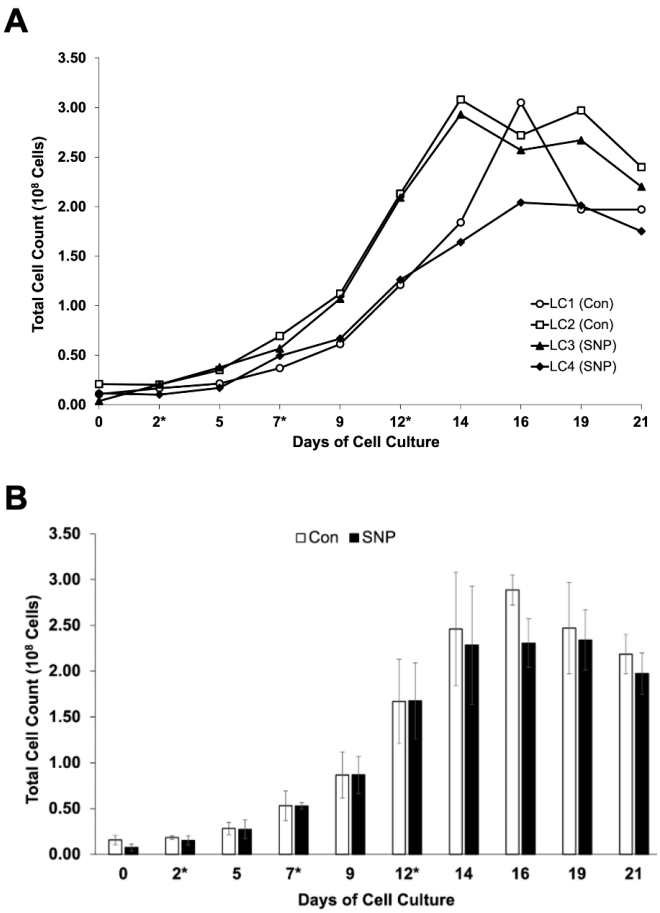

Figure 2: SNP and non-SNP LCL growth over a 21-day period. Four Lymphoblastoid Cell Lines (LCLs) derived from HD patients, two with SNP rs79727797 (SNP, solid symbols) and two without SNP rs79727797 (Con, open symbols), were counted every 2–3 days, before being resuspended in fresh media. LCLs were diluted five-fold on days 2, 7, and 12 (marked with asterisks), but we corrected for the dilution factor when calculating their total cell count. (A) Total cell counts for the four different cell lines. (B) Comparison of the average total cell count between the SNP and non-SNP (Con) cell lines. Error bars represent standard error of mean (SEM) between two biological replicates (n=2).

out HAP40, a well-characterized interactor of HTT, has been shown to reduce HTT levels (19).

To test the potential effect of SNP rs79727797, we conducted western blot (WB) analysis using lysates prepared from LCLs with and without SNP rs79727797. Using specific antibodies, we detected HTT, TCERG1, and glyceraldehyde-3-phosphate dehydrogenase (GAPDH). We detected protein bands corresponding to HTT, TCERG1, and the housekeeping protein GAPDH in both SNP and control cells (**Figure 3A**). Quantification of protein band intensities revealed no significant differences in the expression levels of HTT, TCERG1, or GAPDH between the two groups (HTT, $p = 0.57$; TCERG1, $p = 0.97$; GAPDH, $p = 0.88$; Figure 3). To ensure robust comparisons, we normalized the band intensities of HTT and TCERG1 relative to GAPDH to control for any potential technical variability. Consistent with the initial analysis, the normalized relative expression levels of HTT and TCERG1 did not differ significantly between the SNP and control groups (HTT, $p = 0.46$; TCERG1, $p=0.74$; **Figure 3B**).

In summary, the presence of SNP rs79727797 did not significantly affect the levels of HTT, TCERG1, or GAPDH proteins in the tested HD LCLs. These findings suggest that SNP rs79727797 does not exert a measurable influence on the expression or stability of these proteins under the conditions tested.

Optimization of TCERG1 immunoprecipitation conditions

To investigate how SNP rs79727797 alters the TCERG1 protein and its interactome in HD patient LCLs, we aimed to optimize a specific immunoprecipitation (IP) protocol targeting TCERG1. This optimization would enable the isolation of TCERG1 and its associated proteins with high specificity and clarity, facilitating downstream analyses to understand the functional impact of the SNP on HD mechanisms. In this experiment, we utilized TCERG1 antibodies to immunoprecipitate TCERG1. An IgG antibody was included as a negative control to confirm the specificity

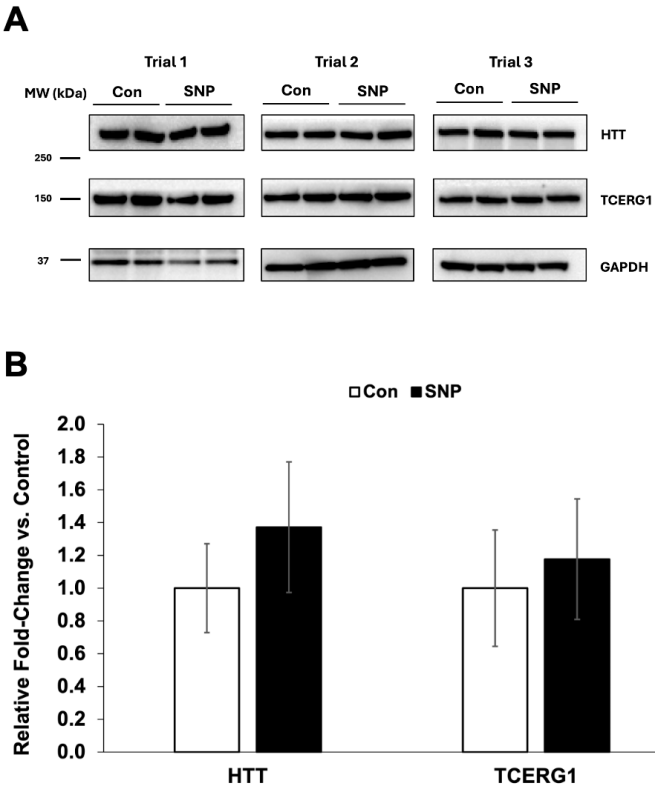

Figure 3: Western blot of cell lysates with and without SNP rs79727797. Four LCLs derived from HD patients – two with the SNP rs79727797 (SNP) and two control cells without it (Con) – were used in each of the three western blot trials. (A) Chemiluminescence was used to image the protein bands following western blotting with HTT, TCERG1, and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) antibodies. (B) Protein band intensities for HTT and TCERG1 were measured and normalized using GAPDH as a reference gene. Averages for the control and SNP were calculated (n=6 per group, each with two biological and three technical replicates). Then, the average GAPDH-normalized intensities for the SNP group were normalized to the average GAPDH-normalized control intensity. A t-test between the control and SNP groups of each protein when normalized for GAPDH ($p=0.46$ for HTT, $p=0.74$ for TCERG1) revealed no statistically significant difference in protein expression between the control and SNP LCLs. Error bars represent standard error of mean (SEM) (n=6).

https://doi.org/10.59720/25-066

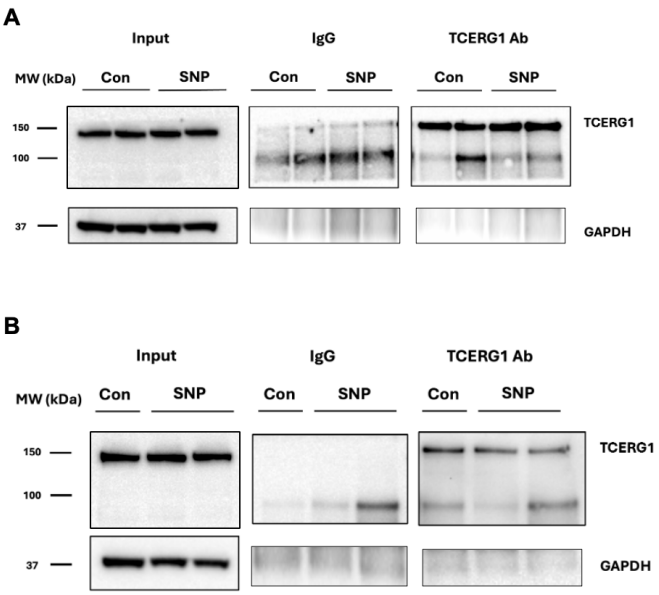

Figure 4: TCERG1 immunoprecipitation with IgG and TCERG1-specific antibodies. Cell lysates of LCLs with and without the SNP rs79727797 derived from HD patients were immunoprecipitated (IP) with Immunoglobulin G (IgG) and TCERG1-specific antibodies, then imaged with chemiluminescence. 2 µg of antibodies were used for every 1 mg of protein lysates and incubated overnight at 4 °C, followed by capture with Protein A magnetic beads. Whole-cell lysates (Input), IgG IP (IgG), and TCERG1 antibody IP (TCERG1 Ab) were subjected to western blotting and probed with TCERG1 and GAPDH antibodies. (A) The first trial used 500 µg of protein lysate in 500 µL final volume. (B) Second trial was performed with BSA-coated magnetic beads and increased lysate concentrations of 1000 µg in 500 µL.

of the TCERG1 IP, as IgG binds non-specifically and does not produce strong signals for any protein, including TCERG1. Initial IP experiments yielded two sets of data. WB analyses demonstrated clear TCERG1 signals when using TCERG1-specific antibodies (**Figure 4A**). As expected, only faint signals were present with the IgG control antibody, corresponding to background staining (**Figure 4A**). These results validate the ability of the IP protocol to specifically capture TCERG1 and its interactome.

However, there were faint background bands, mostly near the 100 kDa marker in both the IgG and TCERG1-Ab conditions. Furthermore, there was a faint TCERG1 band at the 150 kDa marker for the IgG control. These signals likely indicate a degree of non-specific binding from non-target proteins to the beads, and in the IgG condition, low levels of TCERG1 to the beads. To mitigate this variability, we implemented a modified IP protocol. We pre-coated magnetic beads with bovine serum albumin (BSA) to block non-specific protein binding to the TCERG1 antibodies, and we increased the concentration of cell lysates to enhance the binding of TCERG1 in the lysate to the magnetic beads. The optimized protocol produced results consistent with the initial experiments: TCERG1 signals were detected in the TCERG1-specific IP samples, with markedly reduced background signals and non-specific binding, especially in the IgG condition (**Figure 4B**). Importantly, the signal strength for TCERG1 was robust, confirming that BSA coating did not impede TCERG1 binding.

These findings demonstrate that the optimized IP protocol reliably captures TCERG1 with high specificity. Notably, the protocol was successful in both conditions with and without SNP rs79727797 and yielded comparable results, suggesting its applicability for investigating the interactome of TCERG1 under either condition. This method provides a reliable platform for future studies exploring the molecular mechanisms by which TCERG1 interactome alterations contribute to HD pathology. Proteins interacting with TCERG1 may co-immunoprecipitate and be detected in WB analyses and/or be identified via liquid chromatography-tandem mass spectrometry (LC-MS/MS) (**Figure 5**).

DISCUSSION

This study aimed to explore the role of the TCERG1 SNP rs79727797 in HD onset by investigating potential changes in cell growth, gene expression, and protein interactions in HD patient LCLs with the SNP. Our findings suggest that SNP rs79727797 does not directly influence cell growth or the expression of HTT or TCERG1. We successfully developed a specific immunoprecipitation protocol to isolate TCERG1 complexes, laying the groundwork for future studies to explore SNP-related changes under alternative conditions or in other cell types with greater relevance to HD pathology.

Our results indicate that SNP rs79727797 has minimal impact on the growth of LCLs. Cell growth curves over a 21-day period showed no significant differences in growth rates or patterns between the control and SNP cell lines. Although there were slight variations in the cell count during the initial plating of the cell lines, these variations ultimately did not greatly affect cell growth; for example, though the cell line LC3 started with the lowest cell count on Day 0, it ended with the second highest count by Day 21. SNP rs79727797 also had minimal effects on protein expression. WB analyses of lysates

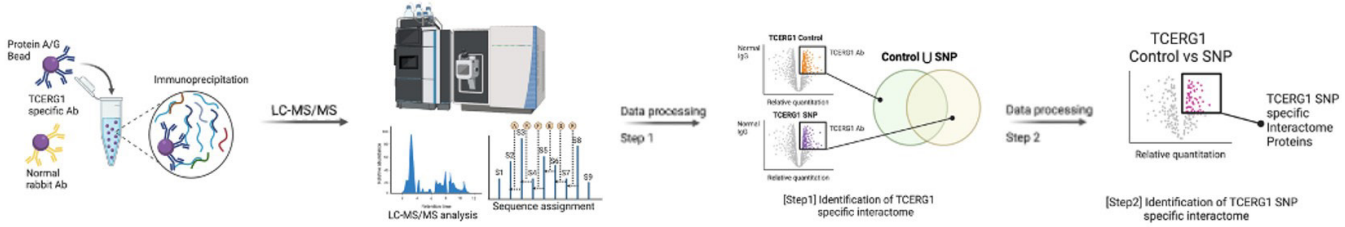

Figure 5: Future implementation of TCERG1 IP protocol. Proposed process to identify the potential SNP-specific (SNP vs. control) interactome of TCERG1 (created with BioRender.com). In future experiments, proteins that co-immunoprecipitate with TCERG1 during the IP experiment could be identified using Liquid Chromatography Tandem Mass Spectrometry (LC-MS/MS). A comparative analysis with proteins co-immunoprecipitated by normal IgG would then help filter out non-specific interactions. By further comparing control and SNP interactomes, we could then isolate SNP-specific interactome proteins, revealing how the SNP may influence TCERG1 protein-protein interactions.

https://doi.org/10.59720/25-066

revealed consistent levels of HTT, TCERG1, and GAPDH proteins across both cells with the SNP and cells without the SNP. These findings suggest that SNP rs79727797 does not directly affect cell growth or the expression of these proteins in LCLs.

A key achievement of this study was the optimization of TCERG1 immunoprecipitation protocols. Through adjustments, including the use of control IgG, BSA-coated beads, and increased lysate concentrations, we minimized non-specific binding and confirmed a TCERG1-specific band. The persistence of non-specific signals in the IgG control is expected, as IgG is inherently non-specific. These data suggest that the observed TCERG1 bands are due to true protein-protein associations rather than background noise, strengthening the validity of the optimized IP protocol. This protocol therefore provides a reliable method for isolating the TCERG1 protein complex and can be used to study its interactome. For example, proteins interacting with TCERG1, such as HTT, could co-immunoprecipitate and be further analyzed via WB analysis or mass spectrometry.

Future experiments could leverage this protocol to conduct comparative analyses of the TCERG1 interactome between control and SNP cell lines. By comparing the proteins from the IgG and the TCERG1-specific IP, we can identify and eliminate the non-specific proteins. A second comparison of the control and SNP interactomes can identify the proteins that interacted only with the SNP-altered TCERG1. By further investigation of these proteins, these studies have the potential to reveal the effects of SNP rs79727797 on TCERG1 interactions and its downstream impact on HD onset.

Although we do not observe a direct effect of the SNP rs79727797 on cell growth or protein expression, there may be an alternative indirect mechanism underlying its potential significance in HD onset. Recent studies suggest a relationship between HD onset and the Quasi-Tandem Repeat (QTR) region of *TCERG1*, a tandem hexamer repeat encoding a glutamine-alanine (QA) tract located in exon 4, not far from SNP rs79727797 (14). Specifically, the number of QTR repeats in both alleles was shown to be a genetic modifier of HD that correlates to an earlier age of onset (14). Among multiple variants of this QTR region, the A2 allele, which was the second most prevalent, had a 99% correlation with SNP rs79727797, suggesting a possible indirect mechanism through which the SNP could influence HD onset. Further research into the precise mechanisms linking the SNP and QTR regions to HD AO could provide crucial insights.

It is important to note that results may vary depending on the cell type used. TCERG1 is known to play a role in gene regulation and RNA processing in a cell-type-specific manner (20). LCLs, created using the Epstein-Barr virus (EBV), may have altered gene expression profiles compared to other cell types (21). Moreover, these B-cells are derived from blood, whereas HD pathology primarily affects neurons, particularly medium spiny neurons. Differences in nuclear fraction, morphology, and cellular functions across cell types may influence the effects of SNP rs79727797 (22). Future studies should incorporate neuronal models to better understand the SNP's potential impacts on HD.

In conclusion, this study provides an optimized and validated IP protocol for TCERG1 in HD-related LCLs. While the *TCERG1* SNP rs79727797 appears to have a limited direct effect on cell growth and protein expression in LCLs, its potential role in HD onset, particularly through indirect mechanisms such as its association with QTR alleles, warrants further investigation. The tools and findings from this study lay the groundwork for future research into the TCERG1 interactome and its implications for HD pathology.

MATERIALS AND METHODS

LCL Culture and Counting

The panel of EBV-transformed lymphoblastoid cell lines (LCLs), derived from four previously collected and genotyped HD individuals in the Mass General Brigham Institutional Review Board-approved Center for Human Genetic Research Neurodegenerative Repository, comprised two LCLs naturally carrying the SNP rs79727797 (SNP) and two lacking the variant (Con). The *HTT* CAG sizes of the LCLs were 45 and 16, 46 and 17, 40 and 17, and 41 and 18 repeats for the two alleles, respectively. LCLs with and without SNP rs79727797 were cultured in RPMI 1640 media supplemented with 10% Fetal Bovine Serum and Penicillin. Cells were suspended in 10 mL media and maintained in T25 cell culture flasks. Flasks were incubated until the next passage at 37°C. When counting, cell pellets were first isolated through centrifugation, followed by washing with 5 mL media. After the washing media was discarded, the cells were resuspended in 1 mL media. If needed, a portion of the cells were taken and diluted in a 1:5 ratio with PBS (Gibco). The cells were mixed in a 1:1 ratio with Tryptophan blue and placed on Countess Cell Counting Chamber Slides (Fisher). Remaining cells were then resuspended in 10 mL new media for a new passage. Since LCL growth plateaus when cell densities reach 10^6 cells/mL, cultures were diluted five-fold on days 2, 7, and 12, when cell densities approached or exceeded this threshold (17). By transferring a fifth of the culture into fresh media, cells were maintained in continuous growth and prevented from saturation. Total cell counts were calculated based on initial dilutions to monitor growth over a 21-day period.

Cell Lysis and Western Blot Analysis

Cell pellets were isolated through centrifugation from media. 1 mL IP Lysis Buffer (Pierce) mixed in a 1:100 ratio with 10x Halt Protease Inhibitor Cocktail (Fisher) was added to resuspend the cell pellets, and the cells were then spun for 15 mins at 4°C. Cell lysates were then isolated through centrifugation (15000 x g, 15 mins) and collection of the supernatant. Concentrations of the lysates were calculated with a bicinchoninic acid (BCA) assay (Pierce). 25 µg of cell lysates were mixed with 4x LDS Sample Buffer (Fisher) and 10x Sample Reducing Agent (Fisher) as well as nuclease-free water to normalize the concentration. The samples were then heated at 70°C for 10 mins.

A 4–12% Bis-Tris NuPage Gel (Invitrogen) was loaded with 5 µL of Protein Standards and 20 µL of the inputs and IP samples. Gels were run at 140 V for 1 hr. PVDF membrane transfers were conducted on ice at 100 V for 1 hr. After transfer, the membrane was blocked using 5% nonfat milk diluted with Tris-Buffered Saline with 1.25% Tween-20 (TBST) for 1 hr on a rocker at room temperature (RT). Next, the membrane was cut into pieces according to the different protein sizes, specifically along the 250 kDa and 75 kDa markers. Appropriate primary antibodies were then added to the membranes, which were then incubated overnight in 4°C on a rocker. The primary antibodies were made by mixing 10

https://doi.org/10.59720/25-066

mL 3% nonfat milk diluted with TBST in a 1:1000 ratio with either mouse anti-HTT (1:1000, Sigma), rabbit anti-TCERG1 (1:1000, Novus), or rabbit anti-GAPDH (1:1000, Invitrogen). GAPDH served as a housekeeping protein and internal control to account for technical variation. After the overnight incubation, the membrane was washed five times with 10 mL TBST for 5 mins each time on a rocker at RT. Then, 2 µL of secondary antibody, either goat anti-mouse (1:5000, Invitrogen) or goat anti-rabbit (1:5000, Invitrogen), was added with 10 mL TBST. The membranes were placed on a rocker for 1 hr at RT. Afterwards, the secondary antibody was discarded, and the membrane was washed with 10 mL TBST five times on a rocker at RT for 15 mins each time. Membranes were dipped in ECL Chemiluminescence reagents (Fisher) and imaged using the iBright ChemiBlot software to measure chemiluminescence for each protein quantification.

TCERG1 Immunoprecipitation

LCLs were harvested and lysed as described above. Initial experiments used 500 µg of cell lysates. Dynabead magnetic beads (Fisher) were washed with IP Lysis Buffer (Pierce) mixed with 5 M NaCl in a 1:20 ratio three times. Magnetic racks were used to separate the beads from the supernatant. Then, the beads were placed on a rocker at RT for 1 hr with 2% BSA in PBS. Since BSA coating can reduce non-specific binding but also reduce the recovery of the target protein, the lysate input was increased to 1000 µg in the optimized protocol to ensure proper recovery of TCERG1. 30 µL of beads were used for each sample. Cell lysate concentrations were equalized through appropriate dilution with IP Lysis Buffer. Either IgG or TCERG1 antibodies were added to the lysates and rotated at 4°C for 1 hr. For every 1 mg of lysate, 2 µg of antibody was used. After the incubation, the lysates were mixed with the preincubated beads and rotated at 4°C for 1 hr. Following this, the beads were washed five times with IP Lysis Buffer mixed with 5 M NaCl in a 1:20 ratio. Magnetic racks were used to separate the beads from the supernatant. Finally, the beads were boiled for 10 mins at 70°C to create sample inputs. The beads were taken out from the sample using magnetic racks, isolating the supernatant for WB analysis.

Statistical Analysis

For the lysate protein quantification, we compared the six total replicates (two biological replicates with three technical replicates each) of the non-SNP control versus the six replicates with the SNP using a paired Student's t-test, two tailed.

ACKNOWLEDGMENTS

We would like to express our sincerest gratitude to the Center for Genomic Medicine at Massachusetts General Hospital for providing the invaluable environment and resources to enable this research. We would especially like to thank the Seong Lab and its members for their unwavering support and technical expertise to aid the progression of this project.

Received: March 22, 2025
Accepted: July 22, 2025
Published: September 19, 2025

REFERENCES

1. Ajitkumar, Anitha, et al. "Huntington Disease." *PubMed*, StatPearls Publishing, 2023, www.ncbi.nlm.nih.gov/books/NBK559166/.
2. MacDonald, Marcy E., et al. "A Novel Gene Containing a Trinucleotide Repeat That Is Expanded and Unstable on Huntington's Disease Chromosomes." *Cell*, vol. 72, no. 6, Mar. 1993, pp. 971–983, https://doi.org/10.1016/0092-8674(93)90585-E.
3. Hendricks, Audrey E., et al. "Estimating the Probability of de Novo HD Cases from Transmissions of Expanded Penetrant CAG Alleles in the Huntington Disease Gene from Male Carriers of High Normal Alleles (27-35 CAG)." *American Journal of Medical Genetics Part A*, vol. 149A, no. 7, July 2009, pp. 1375–1381, https://doi.org/10.1002/ajmg.a.32901.
4. Langbehn, Douglas R., et al. "CAG-Repeat Length and the Age of Onset in Huntington Disease (HD): A Review and Validation Study of Statistical Approaches." *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, vol. 153B, no. 2, 22 June 2009, pp. 397-408, https://doi.org/10.1002/ajmg.b.30992.
5. Lee, J.-M., et al. "CAG Repeat Expansion in Huntington Disease Determines Age at Onset in a Fully Dominant Fashion." *Neurology*, vol. 78, no. 10, 8 Feb. 2012, pp. 690–695, https://doi.org/10.1212/WNL.0b013e318249f683.
6. Genetic Modifiers of Huntington's Disease (GeM-HD) Consortium. "CAG Repeat Not Polyglutamine Length Determines Timing of Huntington's Disease Onset." *Cell*, vol. 178, no. 4, 8 August 2019, pp. 887-900, https://doi.org/10.1016/j.cell.2019.06.036.
7. "TCERG1 Transcription Elongation Regulator 1 [Homo Sapiens (Human)]." *NCBI*, http://www.ncbi.nlm.nih.gov/gene/10915. Accessed 30 Aug. 2024.
8. Naftelberg, Shiran, et al. "Regulation of alternative splicing through coupling with transcription and chromatin structure." *Annual review of biochemistry*, vol. 84, June 2015, pp. 165–98. https://doi.org/10.1146/annurev-biochem-060614-034242.
9. Muñoz-Cobo, Juan Pablo, et al. "Transcriptional Elongation Regulator 1 Affects Transcription and Splicing of Genes Associated with Cellular Morphology and Cytoskeleton Dynamics and Is Required for Neurite Outgrowth in Neuroblastoma Cells and Primary Neuronal Cultures." *Molecular Neurobiology*, vol. 54, 14 Nov. 2016, pp. 7808–7823, https://doi.org/10.1007/s12035-016-0284-6.
10. Yang, Pan, et al. "Overexpression of TCERG1 as a Prognostic Marker in Hepatocellular Carcinoma: A TCGA Data-Based Analysis." *Frontiers in Genetics*, vol. 13, 10 Oct. 2022, https://doi.org/10.3389/fgene.2022.959832.
11. Suñé, C, et al. "CA150, a Nuclear Protein Associated with the RNA Polymerase II Holoenzyme, Is Involved in Tat-Activated Human Immunodeficiency Virus Type 1 Transcription." *Molecular and Cellular Biology*, vol. 17, no. 10, Oct. 1997, pp. 6029–39, https://doi.org/10.1128/MCB.17.10.6029.
12. Holbert, S., et al. "The Gln-Ala Repeat Transcriptional Activator CA150 Interacts with Huntingtin: Neuropathologic and Genetic Evidence for a Role in Huntington's Disease Pathogenesis." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 4,

https://doi.org/10.59720/25-066

30 Jan. 2001, pp. 1811–1816, https://doi.org/10.1073/pnas.98.4.1811.
13. "SNP - NCBI." *NIH*. www.ncbi.nlm.nih.gov/snp/?term=rs79727797. Accessed 16 June 2025.
14. Lobanov, Sergey V., et al. "Huntington's Disease Age at Motor Onset Is Modified by the Tandem Hexamer Repeat in TCERG1." *NPJ Genomic Medicine*, vol. 7, no. 53, 5 Sept. 2022, https://doi.org/10.1038/s41525-022-00317-w.
15. Arango, Margarita, et al. "CA150 Expression Delays Striatal Cell Death in Overexpression and Knock-in Conditions for Mutant Huntingtin Neurotoxicity." *Journal of Neuroscience*, vol. 26, no. 17, *Society for Neuroscience*, Apr. 2006, pp. 4649–59, https://doi.org/10.1523/jneurosci.5409-05.2006.
16. Kim, Kyung-Hee, et al. "Genetic and Functional Analyses Point to FAN1 as the Source of Multiple Huntington Disease Modifier Effects." *American Journal of Human Genetics*, vol. 107, no. 1, July 2020, pp. 96–110, https://doi.org/10.1016/j.ajhg.2020.05.012.
17. Coriell Institute for Medical Research. *Lymphoblast Culture FAQ*. Coriell Institute for Medical Research, n.d. Accessed 2 May 2024. https://www.coriell.org/0/PDF/Lymphoblast_Culture_FAQ.pdf.
18. Nicolae, Dan L., et al. "Trait-Associated SNPs Are More Likely to Be EQTLs: Annotation to Enhance Discovery from GWAS." *PLOS Genetics*, vol. 6, no. 4, 1 Apr. 2010, https://doi.org/10.1371/journal.pgen.1000888.
19. Xu, Shiyu, et al. "HAP40 Is a Conserved Central Regulator of Huntingtin and a Potential Modulator of Huntington's Disease Pathogenesis." *PLOS Genetics*, vol. 18, no. 7, 19 July 2022, https://doi.org/10.1371/journal.pgen.1010302.
20. Pearson, James L, et al. "Identification of the Cellular Targets of the Transcription Factor TCERG1 Reveals a Prevalent Role in mRNA Processing." *Journal of Biological Chemistry*, vol. 283, no. 12, 1 Mar. 2008, pp. 7949–7961, https://doi.org/10.1074/jbc.M709402200.
21. Mrozek-Gorska, Paulina, et al. "Epstein–Barr Virus Reprograms Human B Lymphocytes Immediately in the Prelatent Phase of Infection." *Proceedings of the National Academy of Sciences*, vol. 116, no. 32, 24 July 2019, pp. 16046–16055, https://doi.org/10.1073/pnas.1901314116.
22. Ulloa, Romina, et al. "B Cells Adapt Their Nuclear Morphology to Organize the Immune Synapse and Facilitate Antigen Extraction." *Frontiers in Immunology*, vol. 12, article no. 801164, 9 Feb. 2022, https://doi.org/10.3389/fimmu.2021.801164.