Establishment and characterization of a DGCR8 monoalleic human pluripotent stem cell line

Ph.D. thesis

Dóra Reé

Molecular Medicine Doctoral School

Semmelweis University

Supervisor: Ágota Apáti, Ph.D.

Official reviewers: Karolina Pircs, Ph.D.

Melinda Pirity, Ph.D.

Head of the Complex Examination Committee:

Miklós Kellenmayer, M.D., D.Sc.

Members of the Complex Examination Committee:

Attila Tordai, Ph.D., D.Sc.

Elen Gócza, Ph.D., D.Sc.

Gábor Földes, M.D., Ph.D.

Budapest

2023

1. Introduction

miRNAs exert their function in the post-transcriptional regulation of gene expression. They have been shown to influence many important biological processes, including cell differentiation, proliferation, apoptosis, and embryonic development. A dysregulated miRNA pattern is often involved in various diseases, suggesting the crucial role of miRNA dosage in pathophysiological processes. The primary emphasis in research has traditionally been on mechanistic investigation of miRNA function. However, it is equally important to understand the importance of alterations in their coordinated regulation in disease pathogenesis. Pri-miRNA processing is one such regulatory step, carried out by the Microprocessor complex consisting of two components: DROSHA and DGCR8 (DiGeorge syndrome critical region 8). DGCR8 is one of the genes affected by DiGeorge syndrome (DGS), a congenital disease caused by hemizygous microdeletions in the 22q11.2 region. Major clinical manifestations in DGS include congenital heart disease, palatal abnormalities, immune deficiency, and central nervous system anomalies, among others. Since 2007 various

DGCR8 knockout model approaches have been established on mice and mESCs. Most of these studies concentrated on conditional- or tissue-specific knockouts. Although these models do indeed provide an ideal platform for studying the DGCR8 protein, non-canonical miRNAs, or non-canonical DGCR8 functions, they have less impact on the interpretation of the partial disruption of DGCR8 presented in DGS. Studies concerning the consequences of monoallelic DGCR8 expression predominantly use mESC-derived neural cultures or mice for studying DGS or SCZ. They often find overlapping phenotype with DGS mouse models, including signs of altered brain miRNA biogenesis. It is noteworthy that transcriptional networks and signaling pathways of mouse and human pluripotent stem cells (PSCs) display considerate divergence due to species differences, making mESCs less reliable as models of human disease. To our knowledge, only two studies have been published with partial or complete loss of DGCR8 in hPSCs. The first reported altered cell cycle and poor self-renewal capacity coupled with spontaneous differentiation in DGCR8$^{-/-}$ hiPSCs. The other reported overlapping phenotypic

alterations when comparing DGS- and DGCR8$^{+/-}$ hiPSC-derived cortical neurons, further suggesting the prominent role of DGCR8 in DGS. Nevertheless, a comprehensive characterization of the DGCR8 haploinsufficient hPSCs remained unexplored within the scope of these earlier investigations.

2. Objectives

The ability to generate human embryonic stem cell (hESC) lines with specific genetic mutations has been a significant advancement in biomedical research. These cell lines serve as invaluable models for investigating early developmental processes, enabling studies that are difficult to perform *in vivo*. This dissertation is dedicated to bridge an important gap in knowledge by establishing and characterizing a hESC line with DGCR8 deficiency. Within the scope of this research, we pursued two primary objectives. Firstly, we sought to determine the feasibility of employing a knock-in/knock-out approach using two tandem CAG-driven selection markers in hESCs, focusing on the DGCR8 gene. Secondly, our aim extended to characterizing and registering the resultant cell line in accordance with the rigorous criteria set for newly

established stem cell lines. Given the profound biological significance of tightly regulated miRNA biogenesis during early embryogenesis, this work also addresses the question of how DGCR8 deficiency impacts the viability, maintenance of pluripotency and differentiation potential of the hESCs, the mRNA and protein levels of the Microprocessor components and the alterations of the functions of the miRNA machinery. In its entirety, this research contributes to the expanding knowledge regarding miRNA regulation and the role of DGCR8 in pathological phenotypes, such as those manifested in DiGeorge syndrome. Furthermore, the cellular model developed herein presents a novel platform for studying miRNA biogenesis within human pluripotent stem cells and their subsequent differentiation.

3. Methods

Cell culture

hESCs were grown on Matrigel coated plates and kept in mTeSR1 media with or without 0.8 µM puromycin. Cells were passaged as single cells using Accutase and they were then placed in mTeSR1 with 10 µM Y27632.

CRISPR/Cas9 genome editing

The sgRNA was designed with an online tool developed by the Zhang lab (http://crispr.mit.edu/). The custom oligonucleotides were inserted into a pX330 plasmid. The GFP and puromycin resistance genes were cloned into the 'self-cleaving' donor plasmid. hESCs were electroporated with the two vectors using the Amaxa Nucleofector kit for human stem cells.

Single-Cell Cloning

The hESCs were separated using Accutase and the single cells (SSCs) were placed onto 96-well plates using a fluorescent cell sorter in cloning medium (mTeSR1 supplemented with 1/3 MEF-CM). When reaching 80% density, they were transferred to 24- and subsequently 6-well plates. GFP expression was confirmed by FACS.

Trichostatin A Treatment

Cells were grown without puromycin for 3 days, followed by a treatment with 30 nM or 60 nM Trichostatin A (TSA). The next day, GFP expression was measured in treated and untreated control cells using FACS. TSA-induced demethylation was proved by RT-qPCR measurements of the GAGE cancer testis gene.

In vitro spontaneous differentiation

The hESC colonies were separated using collagenase and grown in suspension on ultra-low attachment plates for 6 days. Subsequently, the EBs were moved to 24-well TC plates or confocal chamber slides coated with 0.1% gelatin and allowed to differentiate for another 12 days. The derivatives of the EBs were examined by immunocytochemistry (ICC) and RT-qPCR.

Immunofluorescent staining and flow cytometry

Cells were fixed, blocked, and permeabilized with PFA as described in (Reé et al. 2022). They were then incubated with primary antibodies (60` RT) followed by incubation with secondary antibodies (60` RT). Cell nuclei were stained with DAPI. The SSEA4 FACS analysis was carried out as described in (Reé et al. 2022).

RNA Isolation

Total RNA from hES and differentiated cells was extracted using TRIzol reagent.

miRNA analysis

For pri-miRNA analysis, 1µg of total RNA was reverse transcribed using random oligomers and a high-capacity cDNA reverse transcription kit. For C19MC pri-miRNA,

RT-qPCR was performed using a SYBR Green PCR Master Mix with custom PCR primers. For mature miRNA quantification, the miRCURY LNA™ Universal RT miRNA PCR system was used.

Gene expression analysis

cDNA samples were synthesized from 400 ng of total RNA using the Promega reverse transcription system. mRNA levels for DGCR8, DROSHA, NANOG, AFP, TBXT, and PAX6 were determined using TaqMan gene expression assays. Data were normalized to endogenous control mRNAs: RPLP0 or PolR2A.

Protein Analysis by Western Blotting

Samples were sonicated and then run on 8% acrylamide gels before being electroblotted onto PVDF membranes. The membranes were blocked and incubated with the primary antibody overnight at 4 ° C. Anti-rabbit IgG was used as a secondary antibody. Pierce ECL Western blot substrate was used for signal detection, and the membranes were exposed to Agfa films. An anti-β actin antibody was used to normalize the DGCR8 expression. Expression levels were determined by measuring the density of scanned images using ImageJ.

4. Results

Establishment of the DGCR8 deficient hESC line

For the generation of the DGCR8-deficient hESC line a knock-in/knock-out method based on CRISPR/Cas9 targeting was employed. A specific sequence was designed for targeting DGCR8 by identifying unique sequences in Exon 3. These were then cloned into a px330 vector carrying the spCas9 nuclease. Furthermore, a "self-cleaving" donor vector was constructed with two tandem CAG promoter-driven selection marker genes: puromycin resistance and GFP. HUES9 hESCs were electroporated with the two vectors, followed by puromycin selection and GFP based single-cell cloning. 20 out of 100 clones could be successfully expanded and genotyped. We could not detect biallelic insertions, yet one clone out of 20 showed a monoallelic insertion. Unintended cleavage and mutations at untargeted genomic sites were ruled out by sequence analysis of *in silico* predicted potential off-target sites. The resulting cells expressed OCT4 and NANOG pluripotency markers; spontaneous differentiation resulted in the expression of markers of the three lineages. The decreased expression of the DGCR8 protein was proved

by Western blotting. The cell line was published and registered with the unique identifier 'HVRDe009-A-1'.

Gradual loss of transgene expression in DGCR8$^{+/-}$ hESCs

Monoallelic consistent expression of GFP was observed in the HVRDe009-A-1 cells under selection. However, after puromycin deprivation, the expression of GFP gradually decreased. Silencing through promoter hypermethylation was ruled out as a possible cause by treating cells with TSA. qPCR measurements in GFP negative populations showed a decreased copy number. Altogether 25 single cells could be expanded for further genetic characterization without selection: 12 with GFP positive origins, the rest without. On day 30 during clonal expansion, GFP expression in 7 of 12 GFP positive clones decreased by 75%, and completely by day 60 in two cases. Four of the generated clones (two with GFP positive origins, two without) were selected for further genetic characterization. Diagnostic PCR results showed significant portions missing from the expression cassettes, and subsequent sequencing data proved monoallelic disruption of DGCR8 with an intact allele in all selected clones.

DGCR8$^{+/-}$ hESCs maintain pluripotency and trilineage differentiation capacity

The expression of pluripotency markers OCT4 and NANOG was analyzed by ICC and RT-qPCR. The HVRDe009-A-1 clones constantly expressed these markers, indicating that the mutation did not affect the pluripotent state. FACS analysis also demonstrated high expression of SSEA4, further confirming pluripotency. To investigate the differentiation potential of the clones, EB formation was used. Markers of the three germ layers; ectoderm (TUJ1, PAX6), mesoderm (TBXT, SMA), and endoderm (AFP) were detected by ICC and RT-qPCR. The undifferentiated clones displayed enhanced expression of Oct4 and Nanog, which declined upon differentiation. However, 12 days after EB plating, the differentiated offspring presented increased expression in markers representing the three germ layers. These results demonstrated that heterozygous deletion of DGCR8 did not affect hESC pluripotency and spontaneous differentiation capacity.

DGCR8$^{+/-}$ hESCs show decreased DGCR8 expression.

DGCR8 is known to have a complex autoregulatory loop with DROSHA. Therefore, evaluation of DGCR8 mRNA

and protein levels is essential. In our experiments, RT-qPCR results showed slightly fluctuating and not considerably different DGCR8 mRNA levels in the clones compared to the parental HUES9^{wt} line. On the other hand, Western blot detections in each clone resulted in a 40-50% decrease in protein levels relative to the parental line. DROSHA mRNA and protein levels showed clone-specific expression profiles, making comparisons with the HUES9^{wt} line inconclusive. Compared to the parental line, every clone showed indistinguishable protein levels. These results show that in the expression profile of Microprocessor components on the mRNA level may be inconclusive.

DGCR8$^{+/-}$ hESCs show partial disturbance of Microprocessor function.

The Microprocessor complex is essential for the processing of primary miRNA into mature miRNA. As clustered miRNAs share a primary transcript, processing steps can provide an additional layer for their regulation. A previous study has shown that DROSHA depletion by siRNA silencing in hESCs results in a position-dependent, gradual decrease in pri-miRNA processing along the primate-specific C19MC cluster. We addressed the

question whether partial depletion of the other essential component of the Microprocessor complex, DGCR8, may cause a similar phenotype. Therefore, we examined the processing activity in three selected regions along C19MC in our cells. The results indicated clone-specific responses: Two clones showed a modest, gradual decrease toward the 3'end of the cluster, while this decrease was not observed in the other two. Interestingly, our findings did not demonstrate a clear correlation between DGCR8 expression levels and the observed responses. Even so, they do suggest a potential disturbance in the Microprocessor function among cells where the DGCR8 protein level is reduced due to a heterozygous mutation.

5. Conclusions

This dissertation aims to focus on the establishment and characterization of a DGCR8 hemizygous hESC line. Our knock-in/knock-out NHEJ approach with two tandem CAG driven selection markers resulted in one GFP expressing monoallelic DGCR8 mutant clone (HVRDe009-A-1), suggesting that our method may be feasible for site-directed mutagenesis in hard-to-edit genomic regions and cells. We observed spontaneous loss

of transgene expression caused by genetic rearrangements and the loss of significant portions of both marker-expressing units in the clonal derives of HVRDe009-A-1. Sanger sequencing provided evidence for the sustained monoallelic disruption of DGCR8 in all investigated single cell clones, enabling their future use in fluorescent assays. miRNAs play a pivotal role in the post-transcriptional regulation of gene expression; disruption of miRNA regulation may lead to differentiational defects. Our results show that haploinsufficiency in DGCR8 is not sufficient to disrupt the pluripotent state or the trilineage differentiational potential of hESCs. Given the overall phenotype, we assessed the question whether monoallelic DGCR8 expression is complemented by regulatory mechanisms, or whether there is a major drop in the mRNA and protein levels. Interestingly, while HVRDe009-A-1 cells and their progeny did not show considerably difference in DGCR8 mRNA expression, protein levels were approx. 50% reduced compared to parental HUES9^{wt} cells, suggesting that mRNA levels can be inconclusive. Finally, a previous study found that the processing of the members of the C19MC cluster shows a

position-dependent profile with decreased activity towards the 3' regions in DROSHA knockdown hESCs. We hypothesized that haploinsufficiency in DGCR8 may cause a similar phenotype. Surprisingly, our results did not demonstrate a clear correlation between DGCR8 protein levels and the observed response, while two clones showed a modest, gradual decrease toward the 3′ end, this could not be observed in the other two clones. More studies are needed to decipher the exact processes behind the incomplete penetration of this phenotype.

6. Bibliography of the candidate's publications

Publications related to the PhD thesis

Reé D, Borsy A, Fóthi Á, Orbán TI, Várady G, Erdei Z, Sarkadi B, Réthelyi JM, Varga N, Apáti Á. Establishing a human embryonic stem cell clone with a heterozygous mutation in the DGCR8 gene. Stem Cell Res. 2020 Dec 22;50:102134.

Reé D, Fóthi Á, Varga N, Kolacsek O, Orbán TI, Apáti Á. Partial Disturbance of Microprocessor Function in Human Stem Cells Carrying a Heterozygous Mutation in the DGCR8 Gene. Genes. 2022 Oct 23;13(11):1925.

Other publications

Szabó E, Juhász F, Hathy E, Reé D, Homolya L, Erdei Z, Réthelyi JM, Apáti Á. Functional Comparison of Blood-Derived Human Neural Progenitor Cells. Int J Mol Sci. 2020 Nov 30;21(23):9118.

Szabó E, Reé D, Jezsó B, Vincze K, Földes G, Molnár AÁ, Réthelyi JM, Apáti Á. Generation of iPSC lines from peripheral blood mononuclear cells of identical twins both suffering from type 2 diabetes mellitus and one of them additionally diagnosed with atherosclerosis. Stem Cell Res. 2020 Dec;49:102051.

Onódi Z, Visnovitz T, Kiss B, Hambalkó S, Koncz A, Ágg B, Váradi B, Tóth VÉ, Nagy RN, Gergely TG, Gergő D, Makkos A, Pelyhe C, Varga N, Reé D, Apáti Á, Leszek P, Kovács T, Nagy N, Ferdinandy P, Buzás E, Görbe A, Giricz Z, Varga ZV. Systematic transcriptomic and phenotypic characterization of human and murine cardiac myocyte cell lines and primary cardiomyocytes reveals serious limitations and low resemblances to adult cardiac phenotype. J Mol Cell Cardiol. 2022 Apr;165:19–30.