The role of germline heterozygous *LZTR1* variants in pediatric cancer predisposition

Major Research Project | MSc Cancer, Stem Cells and Developmental Biology | Utrecht University
Princess Máxima Center – Kuiper Group

Djoeke van den Bosse | 25-10-2022

STUDENT NUMBER: 6455751
EXAMINER: PROF. DR. ROLAND KUIPER
SECOND REVIEWER: DR. JARNO DROST
SUPERVISERS: NIENKE VAN ENGELEN AND JETTE BAKHUIZEN

Abstract

The leucine zipper-like transcription regulator 1 (*LZTR1*) gene encodes for the LZTR1 protein, which is the substrate recognition subunit of the CUL3 E3 ligase. LZTR1 can recognize multiple RAS pathway proteins, resulting in proteasomal degradation. LZTR1 is therefore an inhibitor of the RAS pathway. Recently, multiple cases of pediatric cancer were identified to have a germline heterozygous *LZTR1* variant. The role of germline *LZTR1* variants in the development of pediatric cancer has yet to be investigated, though a broad genotype-phenotype association already exists for *LZTR1*. Germline *LZTR1* variants can cause Noonan syndrome in children, and they give a predisposition to the development of Schwannomatosis in adults. Somatic *LZTR1* variants are found in glioblastomas, where a complete loss of LZTR1 in the tumor occurs. In this project we investigated whether germline heterozygous *LZTR1* variants predispose in children to the development of cancer, and through which mechanism these variants could potentially contribute to tumor development. We have created *LZTR1* knock out, heterozygous mutant, and wild type cell lines using CRISPR-Cas9. These cell lines have been genetically validated using PCR and Sanger sequencing. TIDE analysis software was used to determine the different indels present in the cell lines, and this prediction has been validated using PCR cloning and Sanger sequencing. Functional testing of the cell lines using Western Blot showed that LZTR1 was present in the wild type and heterozygous mutant clones but absent in most of the knock-out clones. RNA sequencing of the cell lines was performed to create expression profiles in the aim to find candidate genes and pathways that have an altered expression in the absence of LZTR1, and to compare these expression profiles to RNA sequencing data of tumors of patients with a germline *LZTR1* variant to find overlapping altered genes and pathways. However, the sample size is too small to draw conclusions on the effect of LZTR1 removal on the RNA expression profiles of cell lines. Increasing the number of samples used in the RNA sequencing analyses should therefore be the focus in continuing this project.

Layman's summary

Cancer is caused by various mutations in the DNA which can occur spontaneously during life, called somatic mutations, or which can already be present in the DNA at birth, called germline mutations. The presence of germline mutations which can cause cancer means that someone has a predisposition to cancer. We can identify whether patients have a cancer predisposition by sequencing the DNA of their healthy cells. In this sequencing data we can look for mutations in well-known cancer predisposition genes, but sometimes interesting mutations appear in genes for which we do not yet know if they result in cancer predisposition. An example of such unknown mutations are mutations in the leucine zipper-like transcription regulator 1 (*LZTR1*) gene. We have recently identified multiple different mutations in *LZTR1* in the germline of children with different types of cancer. We see these germline mutations more often in children with cancer than in healthy people. It is not yet known whether these germline *LZTR1* mutations can result in cancer development in children, and thus whether it is a cancer predisposition gene in children. To investigate this, we have created cell lines where the LZTR1 protein is absent, to mimic the presence of a germline *LZTR1* mutation in a patient, as well as cell lines where the LZTR1 protein is still present. We performed RNA sequencing on these cell lines and compared the results with each other, to see what the effect of the removal of LZTR1 is on the gene expression. We looked at individual genes as well as whole signaling pathways to see which genes and pathways are higher or lower expressed when LZTR1 is absent, and whether we can link these genes and pathways to cancer development. Thus far, we cannot draw any conclusion from this analysis because the sample size is too small to determine significant changes in gene expression. Besides the cell lines, we have also collected data from children with cancer and a germline mutation in *LZTR1*. We compared this patient data to a cohort of patients with the same tumor type but without the germline *LZTR1* mutation, to see which genes and pathways are higher or lower expressed in the tumors due to the *LZTR1* mutation. We have performed this comparison for one tumor sample, which is, again, too small to get significant results from the analysis. Lastly, we wanted to compare the results from the cell lines to the results from the patient data. The aim was to look at the genes that are higher or lower expressed in the cell lines without the LZTR1 protein and see if these specific genes have also changed expression in the tumor samples of the patients. To perform this comparison in the future it is necessary to identify more cell clones with and without the LZTR1 protein and perform RNA sequencing on them, to be able to draw conclusions on significant changes in gene expression caused by the absence of LZTR1.

Table of Contents

Introduction

Whole exome or whole genome sequencing of a tumor is often performed to gain insight into the development of a tumor and to find potential drug targets. Sequencing of a patient's germline DNA can also be performed to see if the patient has a possible predisposition to develop cancer. Besides the known cancer predisposition genes, the analysis can also reveal potentially interesting variants in genes for which the role in cancer development is still unknown. Further research should be conducted to determine whether these genes could be new cancer predisposition genes.

An example of germline variants with an unclear significance for tumor initiation are variants in the leucine zipper-like transcription regulator 1 (*LZTR1*) gene in pediatric cancer. Multiple cases have already been described in the literature (1–8) and a few cases in the Princess Máxima Center have been identified as well. The cases described in the literature show a variety of germline LZTR1 variants, mainly loss-of-function variants, which are not restricted to a certain tumor type or a certain hotspot location on the *LZTR1* gene. Loss-of-function variants in *LZTR1* are present in the general population, but with a relatively low allele frequency. The most frequent loss-of-function variant in the population has an allele frequency of 1.71e-4 (9). The increasing number of pediatric cancer patients identified with a germline *LZTR1* variant raises the question whether these variants might predispose children to the development of cancer. *LZTR1* has previously been suggested to be a potential cancer predisposition gene in children (3), but the exact role of germline *LZTR1* variants in pediatric cancer has not yet been investigated.

The *LZTR1* gene is located on chromosome 22 and it encodes the LZTR1 protein (Figure 1), which is a member of the BTB-Kelch protein family. LZTR1 is the substrate recognition subunit of the Cullin3-RING E3 ligase (10,11), and it can bind RAS pathway proteins, including HRAS, KRAS, MRAS, and NRAS (12). This results in proteasomal degradation of these proteins, causing downregulation of the RAS pathway. Previous research has shown that loss of LZTR1 results in overexpression of the RAS pathway (12–14), which is a well-known cancer driving event (15,16). *LZTR1* therefore functions as a tumor suppressor gene.

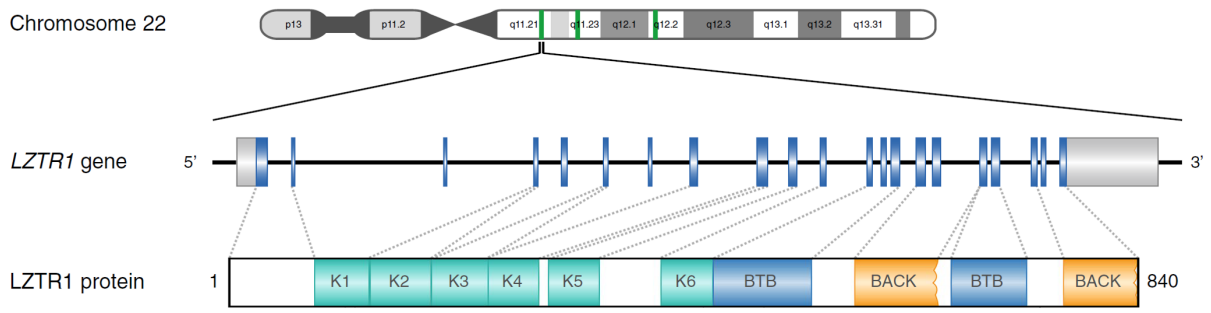

Figure 1: The *LZTR1* gene and protein. The *LZTR1* gene is located on the q arm of chromosome 22 and consist of 21 exons. The LZTR1 protein consist of six Kelch motifs, two BTB domains, and two BACK domains (17).

A link between *LZTR1* mutations and tumor formation in adults has previously been established. Germline *LZTR1* mutations predispose to Schwannomatosis, a benign tumor disorder of the peripheral nervous system (18). For the tumors to develop, a three-step mutation process must occur. The first step is the presence of a heterozygous *LZTR1* mutation in the germline. The second step is a somatic loss of wild type 22q, which contains *LZTR1*, *SMARCB1*, and *NF2*. The last step is a somatic mutation in the remaining *NF2* or *SMARCB1* gene. This process results in a complete loss of LZTR1 as well as additional mutations in genes associated with the development of schwannomas. Somatic *LZTR1*

mutations are associated with glioblastoma (10). Here we see a two-hit inactivation of *LZTR1*, which is common for tumor suppressor genes. In children, germline *LZTR1* mutations are associated with Noonan syndrome. This is a disorder caused by various germline mutations that activate the RAS pathway, and is characterized by congenital heart disease, short stature, facial dysmorphisms, and predisposition to certain types of cancer (19). Noonan syndrome caused by germline *LZTR1* mutations can be inherited in both an autosomal dominant or autosomal recessive manner (8,20,21). Autosomal dominant Noonan syndrome is caused by mutations in the Kelch domain, while autosomal recessive Noonan syndrome can be caused by mutations in the entire *LZTR1* gene (Figure 2).

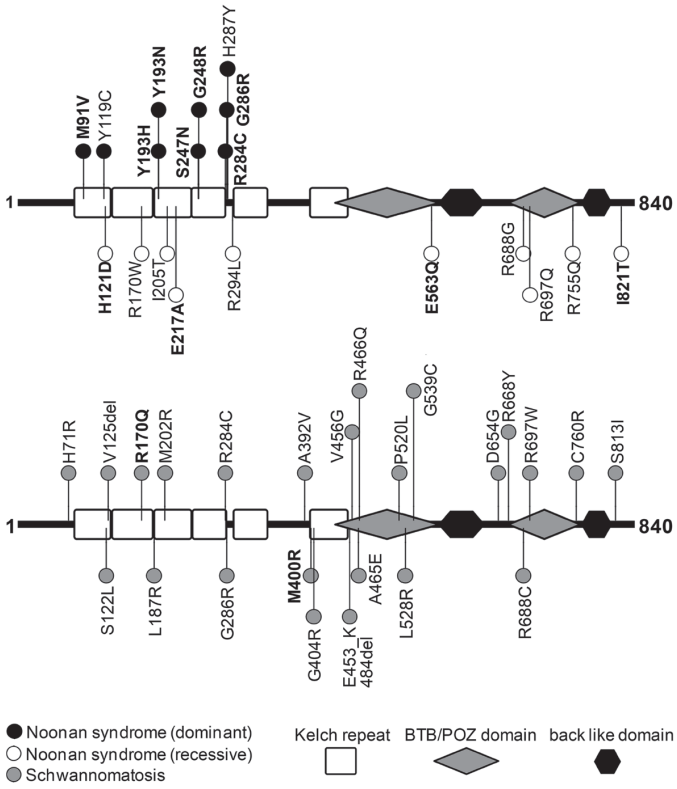

Figure 2: Locations of the different mutations in *LZTR1* causing Noonan syndrome or Schwannomatosis. Autosomal dominant Noonan syndrome is the result of mutations in the Kelch repeat domain of *LZTR1*, while autosomal recessive Noonan syndrome can be caused by mutations in the entire *LZTR1* gene. Mutations resulting in Schwannomatosis are also present in the entire *LZTR1* gene (21).

Taken together, these different examples exhibit a wide spectrum of genotype-phenotype correlations for *LZTR1* mutations. Furthermore, it shows an existing link between germline *LZTR1* mutations and tumor formation, though only in adults. The direct link between germline *LZTR1* mutations and pediatric cancer is not yet clear. We therefore investigated whether germline heterozygous *LZTR1* variants predispose to tumor development in children, and through which mechanism these variants could potentially contribute to tumor formation. Is it a case of haploinsufficiency, like autosomal dominant Noonan syndrome, where the heterozygous mutation is enough to initiate tumor formation? Is a complete loss of *LZTR1* through a somatic mutation necessary for tumor development, like glioblastoma? Or do even more mutations need to occur before tumor formation begins, similar to Schwannomatosis?

To answer these questions, we used a two-way approach (Figure 3). On one side, we collected sequencing data of patients with a germline *LZTR1* mutation to look for a second hit in *LZTR1* in the

tumor and to find mutations in the RAS pathway or other driver mutations which could have initiated the tumor development. On the other side, we aimed to create cell lines with a heterozygous *LZTR1* mutation as well as *LZTR1* knock-out cell lines. After validating the cell lines, we performed RNA sequencing to create expression profiles in the presence and absence of LZTR1. We then aimed to find target genes and pathways that have altered expression in the absence of LZTR1 using these expression profiles. These findings could then be compared to tumor RNA sequencing data of the patients with a germline *LZTR1* variant, to see if the tumors of the patients have an altered expression of the same genes and pathways. This could then be an indication if and through which mechanism the *LZTR1* variants have played a role in the development of the tumor.

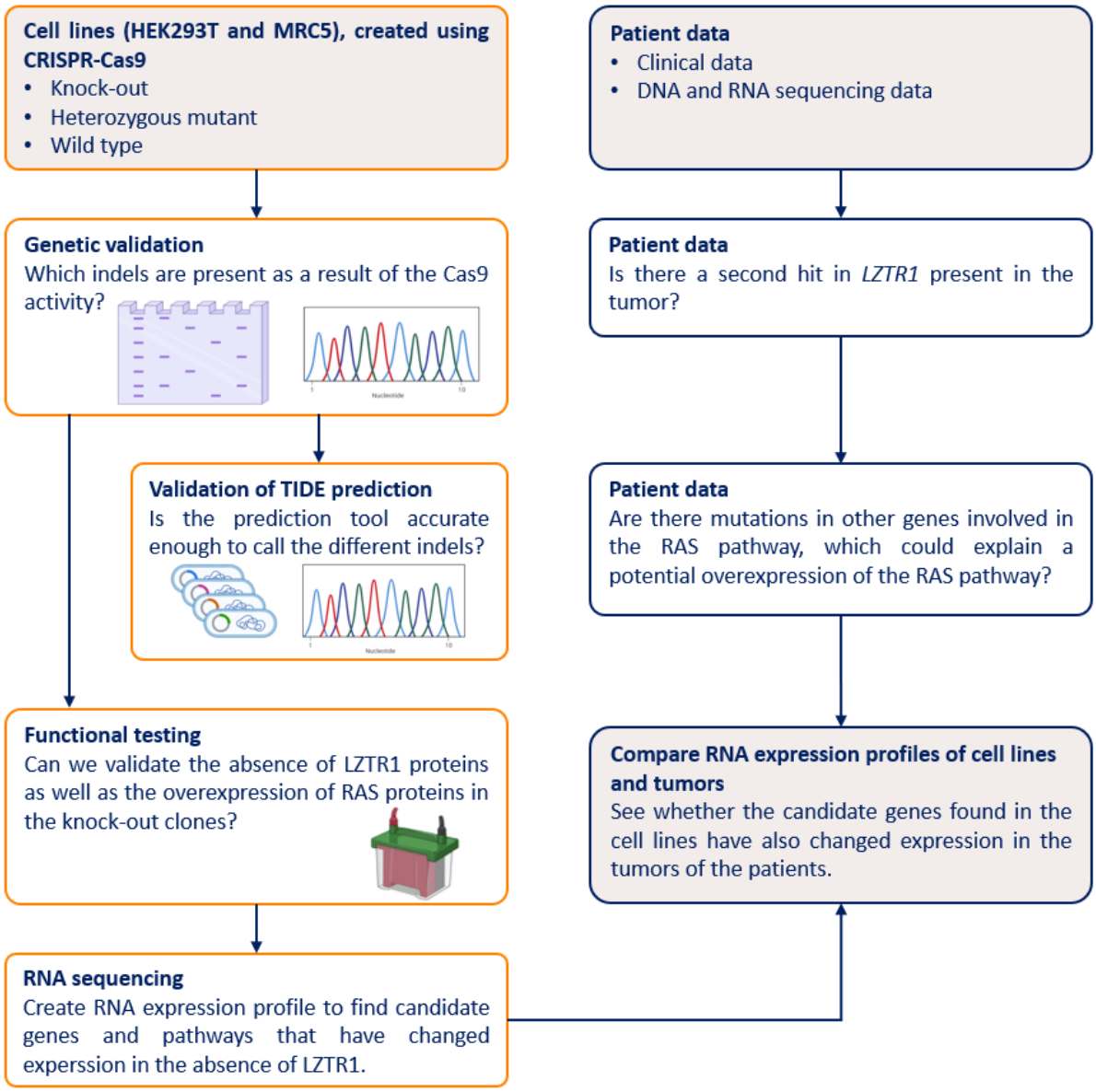

Figure 3: Schematic representation of the research project. The project consisting of two parts: wet lab work to create cell lines and data analysis on patient material. We will create *LZTR1* knock-out and heterozygous mutant cell lines for HEK293T and MRC5 cells, using CRISPR-Cas9 RNP transfections. These cell lines will be genetically validated using PCR, Sanger sequencing, and PCR cloning, and will be functionally tested using Western Blot. We will then perform RNA sequencing to create expression profiles of the cell lines, which will be compared with the RNA expression profiles from the tumors of the patients.

Materials & methods

Cell culture

HEK293T and MRC5 cells were cultured in DMEM (high glucose, GlutaMAX™ supplement, pyruvate (Thermofisher)) supplemented with 10% heat-inactivated FBS, 1% Penicillin-Streptomycin, and 1% Non-Essential Amino Acids. Cells were split twice a week and cells were cultured at 37°C in an atmosphere of 5% CO_2.

RNP transfection

Knock-out and mutant cell lines were created using reverse transfection of RNP complexes as detailed in the IDT protocol (22). In short, a predesigned crRNA (Hs.Cas9.LZTR1.1.AB: ATGGTCGAAGTCCACGCTCG, IDT) was combined in equimolar concentrations with a fluorescently labeled tracrRNA (Alt-R® tracrRNA ATTO™ 550, IDT), heated at 95°C for 5 minutes, and cooled down to room temperature to create a gRNA with a concentration of 1µM targeting exon 1 of *LZTR1*. The gRNA was combined with 1 µM High-fidelity Cas9 enzyme (IDT), Cas9 PLUS™ reagent (from the CRISPRMAX kit, Invitrogen), and Opti-MEM reduced serum medium (Thermofisher) and incubated at room temperature for 5 minutes to create RNP complexes. The RNP complexes were combined with lipofectamine CRISPRMAX Cas9 Transfection Reagent (Invitrogen) and Opti-MEM medium and incubated at room temperature for 20 minutes to form transfection complexes. 200 µl of the transfection complexes was added to each well of a 24-well plate. Next, 1.6×10^6 of either HEK293T or MRC5 cells, suspended in culture medium without antibiotics, were added to the 24-well plate, and the cells were incubated for 48h. Transfection medium was removed from the cells after 48h of incubation and changed with normal culture medium. Microscopic images were taken 24h and 48h after the start of the transfection using a Leica DFC3000 G microscope camera. Images were analyzed in the Leica Application Suite X (LAS X) to determine the transfection efficiency.

DNA isolation

DNA isolation was performed using the QIAamp DNA Blood Mini Kit (Qiagen). A maximum of 5×10^6 cells was centrifuged at 350 *g* for 5 minutes. The cell pellet was washed once with 200 µL PBS and centrifuged again at 350 *g* for 5 minutes. The cell pellet was suspended in 200 µL PBS, 20 µL proteinase K and 200 µL Buffer AL was added, and the solution was mixed by pulse-vortexing for 15 seconds. The cells were incubated at 56°C for a minimum of 10 minutes. Next, 200 µL 100% ethanol was added and mixed by pulse-vortexing for 15 seconds. The solution was moved to a spin column and centrifuged at 6000 *g* for 1 minute. Flow-through was discarded, 500 µL Buffer AW1 was added to the column and centrifuged at 6000 *g* for 1 minute. Flow-through was discarded, 500 µL Buffer AW2 was added to the column and centrifuged at 20,000 *g* for 3 minutes. Flow-through was discarded and the column was centrifuged at full speed for 1 minute. The column was placed in a 1.5 mL Eppendorf tube, 35 µL of MQ was added, and the column was incubated at room temperature for 1 minute. The column was then centrifuged at 6000 *g* for 1 minute to collect the DNA in the Eppendorf tube. DNA concentrations were measured using the Qubit dsDNA Broad Range or High Sensitive Assay kit (Invitrogen).

PCR + Sanger sequencing

The PCR reactions were performed using a 1 U/µl Taq DNA polymerase (Roche). A master mix was composed of 0.25 units of Taq polymerase, 1× reaction buffer containing $MgCl_2$, 0.2 mM dNTPs (Thermofisher), and 0.4 mM of both the forward and the reverse primers (IDT, Table 1). A DNA concentration of 10-25 ng/µl was used for each reaction. The reaction volumes were made up to a total volume of 12.5 µl per reaction using MQ. The PCR program was run according to the manufacturer's protocol (Table 2), and results were visualized by performing gel electrophoresis using a 2% agarose gel.

Table 1: Primers used for PCR and Sanger sequencing.

Primer name	Sequence
LZTR1_AB_FW1	TAGGGACGACACACTGCATT
LZTR1_AB_REV1	ACAGCACCATCCACCTCATT
LZTR1_AB_FW2	AAACGCTCCCCAGAAGGTC
LZTR1_AB_REV2	CCAGCCCAATATCCACAAACA

Table 2: PCR program used in combination with the Taq DNA polymerase.

Step	Temperature	Time	
Initial denaturation	95°C	3 min	
Denaturation	95°C	30 sec	*34 cycles*
Annealing	60°C	30 sec	
Extension	72°C	1 min	
Final extension	72°C	10 min	
Hold	4°C	∞	

The PCR products were purified using the Exo-CIP™ Rapid PCR Cleanup Kit (New England BioLabs). 5 µL of the PCR product was combined with 1 µL of Exo-CIP A and 1 µL of Exo-CIP B. The mixture was incubated at 37°C for 4 minutes and then at 80°C for 1 minute. The purified PCR products were submitted at Macrogen for sequencing, and the sequences were visualized in Geneious Prime.

TIDE analysis

The TIDE analysis was performed by uploading the Sanger sequencing results from our cell clones and from a reference sample in the online program (23). Sanger sequences from the HEK293T and MRC5 parental cell lines were used as reference samples for the TIDE analysis. The various settings were then optimized to get the most reliable analysis by creating the highest possible R^2-value.

Single cell sorting

Transfected cells were single cell sorted into 96-well plates using FACS and cultured in DMEM with 20% heat-inactivated FBS, 1% NEAA, and 0.2% Primocin for two weeks to create single cell clones. Two weeks after sorting, the culture medium was changed to normal culture medium and single cell clones were verified by PCR, Sanger sequencing, and TIDE analysis.

Western Blot

Cells were expanded in a 6-well plate until 80% confluency. Whole cell lysates were created by scraping the cells on ice using 2× Laemmli lysis buffer (0.125M Tris pH 6.8, 4% SDS, 20% Glycerol, 0.005% bromophenol blue, 10% β-Mercaptoethanol). The samples were syringed and boiled at 95°C for 5 minutes to denaturate the DNA. Proteins were resolved on a 4-20% TGX precast gel (Bio-Rad) at 100V for 1 hour, transferred to a PVDF membrane (Bio-Rad) using the Trans-Blot Turbo transfer system, and blocked with 5% (w/v) BSA in PBS for 1 hour. Membranes were incubated with primary antibodies (Table 3) diluted in blocking buffer over night at 4°C. After washing 3 times with 0.05% (v/v) PBS-Tween for 10 minutes, the membranes were incubated with secondary antibodies (Table 3) diluted in blocking buffer for 1 hour at room temperature. Membranes were washed 3 times with PBS-Tween and imaged using the Odyssey Li-Cor imaging system.

Table 3: Antibodies used for Western Blotting.

Antibody	Manufacturer	Dilution
anti-pan-Ras	Cell Signaling Technology (#3965)	1:500
anti-β-Actin	Sigma-Aldrich (A5441)	1:5000
anti-LZTR1	Santa Cruz Biotechnology (sc-390166)	1:100
Goat-anti-Rabbit CF770	Sigma-Aldrich (SAB4600215)	1:10000
Goat-anti-Mouse CF680	Sigma-Aldrich (SAB4600199)	1:10000

PCR cloning

HEK293T cells were harvested under normal conditions at 80% confluency. DNA isolation was performed using the QIAamp DNA Blood Mini Kit (Qiagen) and PCR was performed using Taq DNA polymerase (Roche). The PCR products were cloned in One Shot TOP10 chemically competent cells (Thermofisher) using the pGEM-T vector (Promega). The ligation reactions were performed according to the Promega protocol (24) and left for one hour at room temperature and then overnight at 4°C. The ligation reactions were combined with the competent cells and left on ice for 30 minutes, before heat shocking the cells in a water bath at 42°C for 50 seconds. Cells were returned to ice for 2 minutes, after which LB medium was added and the cells were incubated for 1 hour at 37°C with shaking (150rpm). 500µl of the transformation cultures was plated onto duplicate LA plates with ampicillin, X-Gal, and IPTG. Cultures were grown over night at 37°C, after which the white colonies were picked for colony PCR. The purified PCR products were submitted for Sanger sequencing at Macrogen, and sequences were analyzed using Geneious Prime.

RNA isolation and sequencing

Cells were harvested under normal conditions at 80% confluency, at the same time point as the harvesting for the Western Blot samples. RNA isolation was performed using the RNeasy mini kit (Qiagen). A maximum of 5×10^6 cells were centrifuged at 350 *g* for 5 minutes. The cell pellet was washed once with PBS and centrifuged again at 350 *g* for 5 minutes. The PBS was discarded, and the cell pellets were snap frozen using dry ice and 100% ethanol and stored at -80°C. 350 µL of Buffer RLT was added to the frozen cell pellet and mixed by vortexing. The mixture was added to a QIAshredder (Qiagen) and centrifuged at full speed for 2 minutes to homogenize the lysate. 350 µL of 70% ethanol was added to the flow-through and mixed by pipetting. The mixture was added to a spin column and centrifuged at 8,000 *g* for 15 seconds. Flow-through was discarded, 700 µL of Buffer RW1 was added to the column and centrifuged at 8,000 *g* for 15 seconds. Flow-through was discarded, 500 µL Buffer RPE was added to the column and centrifuged at 8,000 *g* for 15 seconds. Flow-through was discarded, 500 µL Buffer RPE was added to the column and centrifuged at 8,000 *g* for 2 minutes. The spin column was placed in a new collection tube and was centrifuged at full speed for 1 minute. The column was placed in a 1.5 mL Eppendorf tube, 35 µL MQ was added, and the column was centrifuged at 8,000 *g* for 1 minute. RNA concentrations were measured using the Qubit RNA Broad Range assay kit (Invitrogen) and RNA quality was determined by measuring the RIN values using a bioanalyzer RNA Nano chip. RNA sequencing was performed by the diagnostics department of the Princess Máxima Center.

RNA sequencing data analysis

The sequencing data was processed by the bioinformatician. In short, read assembly was performed using STAR, variant calling was performed using MUTECT, and variants were annotated using VEP. Read counts were determined in R using Rsubread featureCounts. Counts were normalized for sequencing depth and gene length using Transcripts Per Million (TPM) values. Gene expression of the samples was compared using TPM values and Z-scores. First, we analyzed the *LZTR1* expression of all

the samples for the HEK293T and the MRC5 cell clones separately. Next, we continued with the MRC5 cell line, for which used two wild type and two knock-out clones and performed a manual comparison of their expression profiles to create a list of altered genes in absence of *LZTR1* (Figure 4).

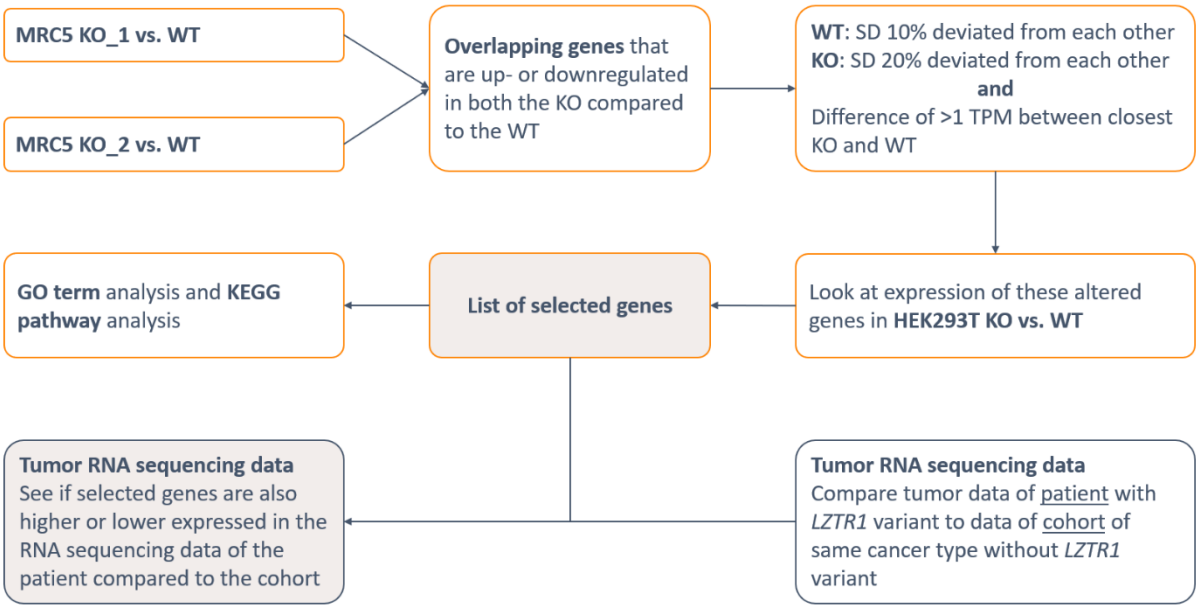

Figure 4: Workflow of the RNA sequencing data analysis. The two knock-out clones were separately compared to both wild type clones, after which the overlapping genes with an altered expression were selected. A filtering step was then performed to select the most interesting genes, after which the list of genes was compared to one wild type and one knock-out clone from the HEK293T cell line. GO term analysis and KEGG pathway analysis was performed on this final list of selected genes. Next, the RNA sequencing data from a tumor sample of a patient was compared to the data of a cohort of patients with the same tumor type but without the germline *LZTR1* variant, after which the list of selected genes was compared with these results.

We first compared each knock-out clone separately to both the wild type clones, after which we selected the overlapping genes that had changed expression in both knock-out clones. We then performed a filtering step to reduce the number of genes in the final list. For this, we selected the genes where the wild type clones had a comparable expression between each other as well as the knock-out clones with each other. For the wild type clones, the standard deviation could not differ more than 10% between the clones and for the knock-out clones not more than 20%. We also focused on the genes where the closest wild type and knock-out clones were at least one TPM away from each other, to increase the chance of finding genes with a significant difference in expression between wild type and knock-out.

We then used one wild type and one knock-out clone from the HEK293T cell line and compared them to results of the MRC5 cell line, creating a final list of selected genes that have an altered expression in both cell lines in the absence of LZTR1. We looked at the individual genes in this list to see if could find anything remarkable, but we also performed GO term analysis and KEGG pathway analysis with the selected genes to see if certain pathways showed up as significantly up- or downregulated.

For the RNA sequencing data of the tumors of patients, we first compared the patient data to a cohort of patients with the same tumor type but without the germline *LZTR1* variant, to see what the effect of the *LZTR1* variant was on the expression profile of the tumor. These results were then compared to the final list of selected genes from the cell lines to see if the changes in the patient data are similar to the cell lines.

Results

During this project we aimed to create *LZTR1* knock-out and heterozygous mutant cell clones for the HEK293T and the MRC5 cell lines by performing CRISPR-Cas9 RNP transfections. We aimed to validate these clones on a DNA level using PCR and Sanger sequencing, and on a protein level using Western Blot. Besides the cell line creation, we also worked on collecting patient data from children with cancer who have a germline heterozygous *LZTR1* variant.

Patient data overview

So far, we have collected 24 cases of pediatric cancer with a germline heterozygous *LZTR1* variant, both in the Princess Máxima Center and in the literature (Table S1). We have identified 20 different germline *LZTR1* variants in these cases, of which 6 have been published previously in association with Noonan syndrome or Schwannomatosis. The variants are mainly loss-of-function variants (Figure 5B) which are present throughout the entire *LZTR1* gene (Figure 5C). The children present with various types of cancer (Figure 5A), and four patients have also been diagnosed with Noonan syndrome. We were able to collect DNA and RNA sequencing data from the tumors of six patients, which revealed that one of the six patients has a somatic mutation in *LZTR1* in the tumor.

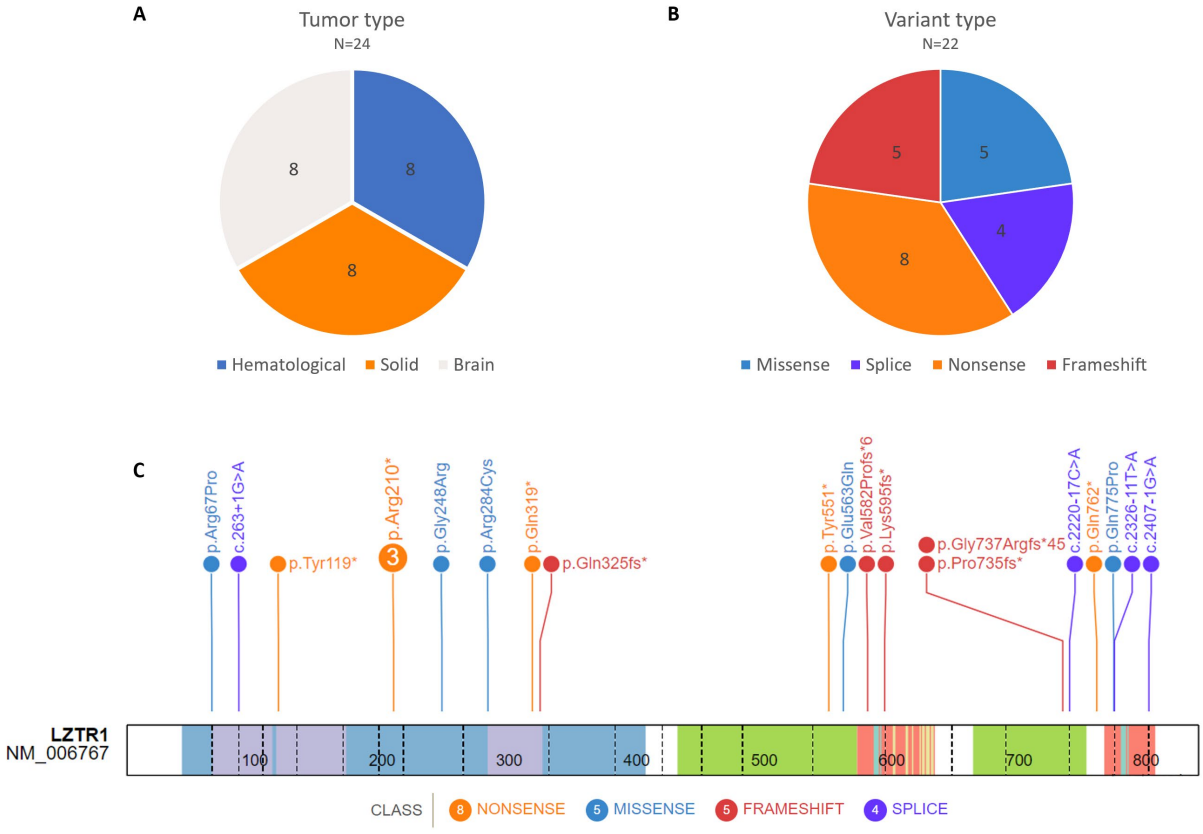

Figure 5: Overview of tumor types and germline *LZTR1* variants present in our database. A: The 24 tumor types are equally distributed over the three main tumor categories. **B:** Most of the variants are nonsense variants, but we have also found frameshift, missense, and splice variants. **C:** A protein paint overview shows that the variants are distributed over the entire *LZTR1* gene, and they are not restricted to a certain domain.

Transfection efficiency

We used a fluorescently labeled tracrRNA and took microscopic images 24 and 48 hours after the start of the transfection to estimate the percentage of cells that have taken up the RNP complexes. The

HEK293T cells did not appear to be in good shape 24 hours after the start of the transfection (Figure 6). The cells in the middle of the well were round, slightly loose, and clumped together instead of spreading out over the bottom of the well. The fluorescent signal of the tracrRNA was present in approximately 80% of the cells. The next day, 48 hours after the start of the transfection, the cells seemed to have recovered slightly. The cells had divided, had a more elongated shape instead of round, were more distributed over the entire well, and had reattached to the bottom of the well. The fluorescent signal of the tracrRNA was now present in approximately 85% of the cells.

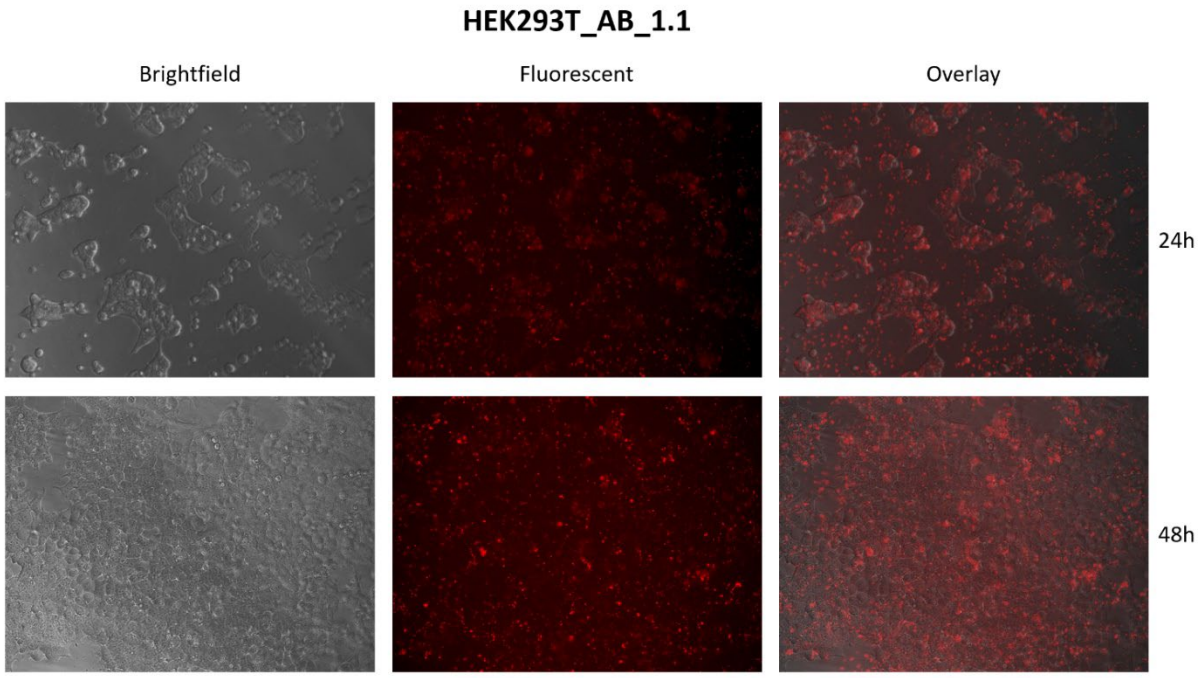

Figure 6: Microscopic images of HEK293T cells 24h and 48h after RNP transfection. Brightfield images at 24h after transfection showed that the HEK293T were not in good condition. The cells were clumped together and round instead of distributed over the bottom of the well. Brightfield images at 48h showed that the cells had recovered and had a more normal shape and equal distribution over the well. The fluorescent images showed that approximately 85% of the cells had taken up the RNP complex at 48h after the start of the transfection.

The MRC5 cells looked normal 24 hours after the start of the transfection, and the fluorescent signal was present in approximately 85% of the cells (Figure 7). At 48 hours after the start of the transfection, the MRC5 cells had divided but the fluorescent signal had not increased significantly.

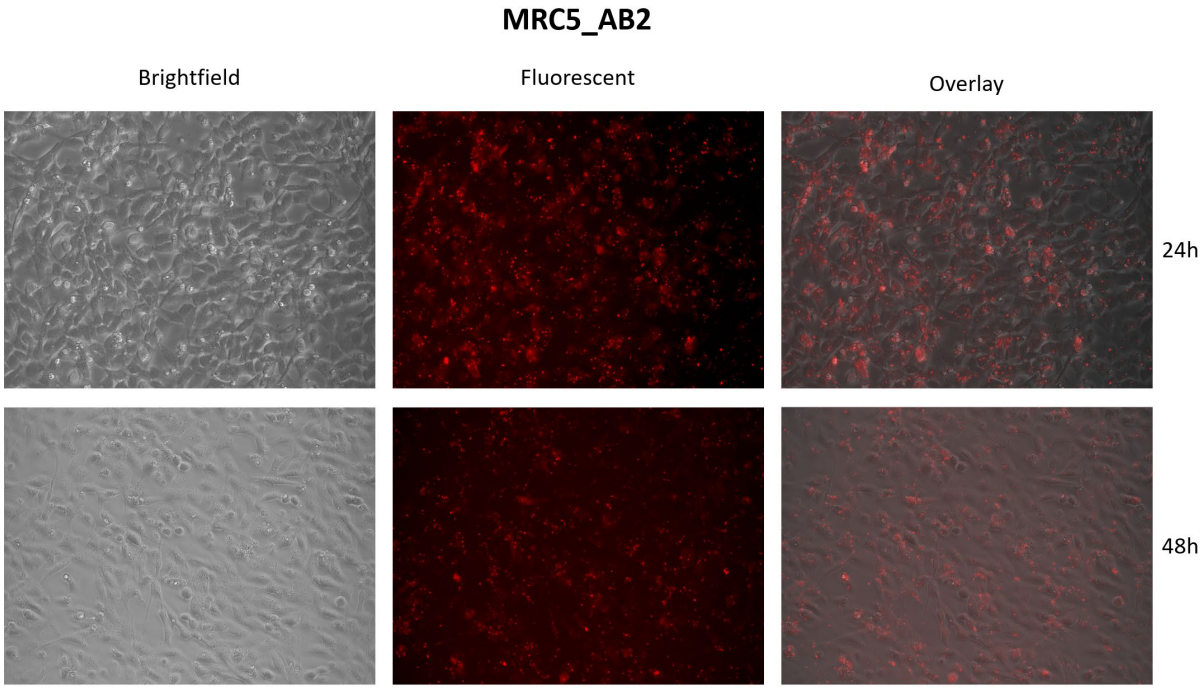

Figure 7: Microscopic images of MRC5 cells 24h and 48h after RNP transfection. Brightfield images at 24h and 48h after the start of the transfection showed that the MRC5 cells looked normal and had divided in this period. The fluorescent images showed that approximately 85% of the cells had taken up the RNP complex, which was similar to the HEK293T cells.

Cleavage efficiency

After the RNP transfection we isolated DNA from a fraction of the transfected cell pool and analyzed this by Sanger sequencing and TIDE analysis to see whether the CRISPR-Cas9 system has cut the DNA at the correct location and to estimate how much of the DNA was edited. The Sanger sequencing for both the HEK293T and MRC5 cell pools shows a change from a clean signal to a mixed signal at the expected break site in *LZTR1* (Figure 8A). The TIDE analysis of the sequences shows a wild-type fraction of 46.6% in the HEK293T cells (Figure 8B) and 26.1% in the MRC5 cells (Figure 8C), indicating that respectively 54% and 74% of the DNA in these cell pools has some type of insertion or deletion (indel).

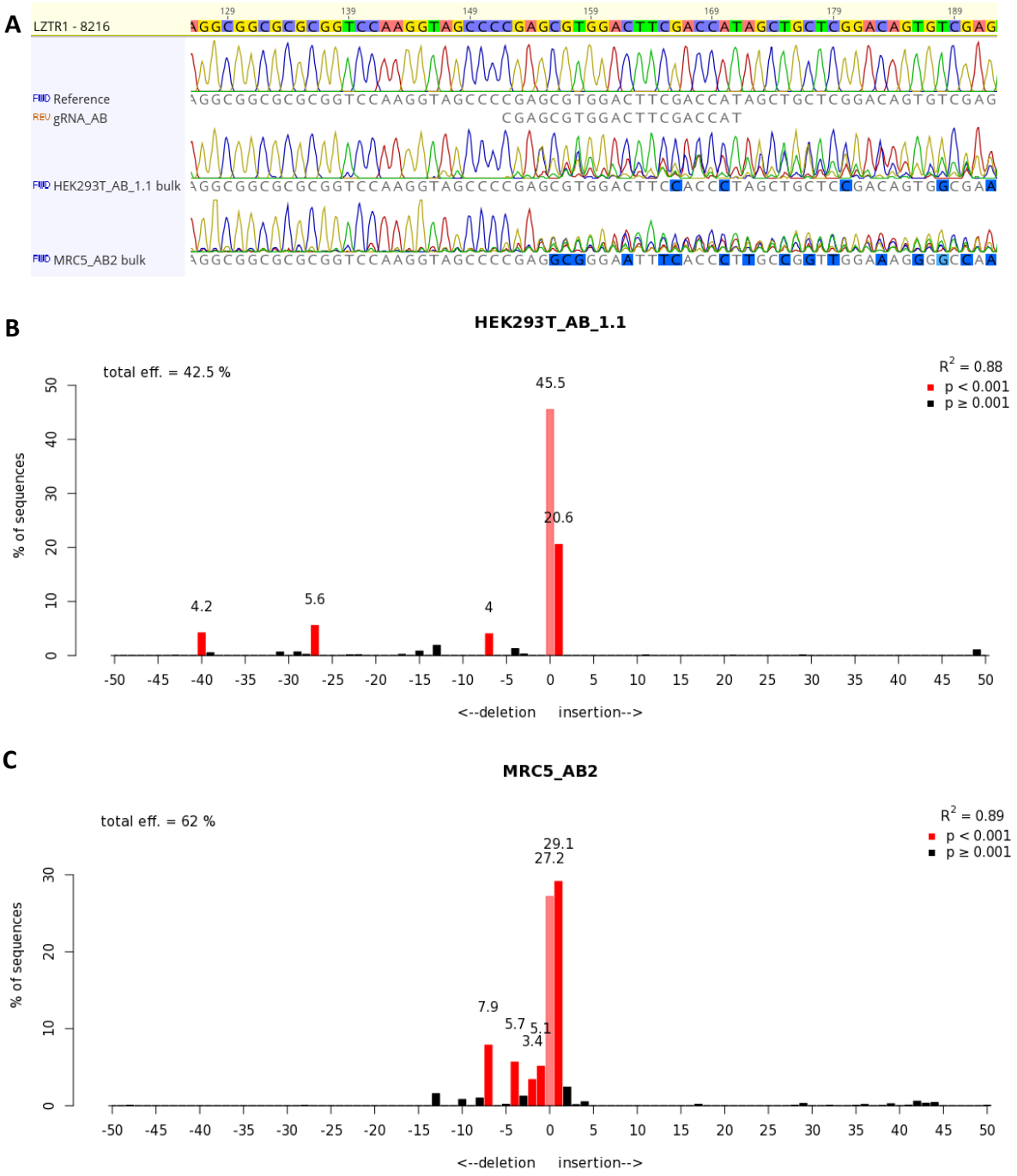

Figure 8: Sanger sequencing and TIDE analysis of a pool of HEK293T cells and a pool of MRC5 cells after transfection with gRNA_AB, compared to a reference sequence. A: Sanger sequences of a reference sample and the HEK293T and MRC5 transfected cell pools showed that for both cell lines, the sequence changes from a clean signal to a mixed signal at the location where the gRNA targets the *LZTR1* gene. **B:** Decomposition of the Sanger sequencing results of the HEK293T cell pool by the TIDE software showed that 45.5% of the DNA is wild type DNA. The rest of the DNA contains various indels. **C:** Decomposition of the Sanger sequencing results of the MRC5 cell pool showed that 27.2% of the DNA is wild type DNA. The rest of the DNA contains various indels.

Genetic validation of single cell clones

After making single cell clones from the transfected cells using FACS, we screened the clones to identify the knock-out, heterozygous mutant, and wild type clones. For the HEK293T cell line, we identified two wild-type clones, two knock-out clones, and three heterozygous mutant clones (Figure

9A, 9B and Table 4). The HEK293T cells have four copies of chromosome 22, so the heterozygous mutant clones were divided into two types: 'HET 3/4' has one wild-type allele and three mutated alleles, 'HET 2/4' has two wild-type alleles and two mutated alleles.

To validate the accuracy of the TIDE analysis for the single cell clones, we performed PCR cloning and Sanger sequencing to create allele specific sequences for six of the HEK293T clones. The results of the PCR cloning showed clean sequences containing only one indel (Figure 9C), instead of the mixed signal visible when sequencing the cell clone (Figure 9A). These indels were the same as the indels predicted by the TIDE analysis for all six of the HEK293T clones, and no additional indels were found.

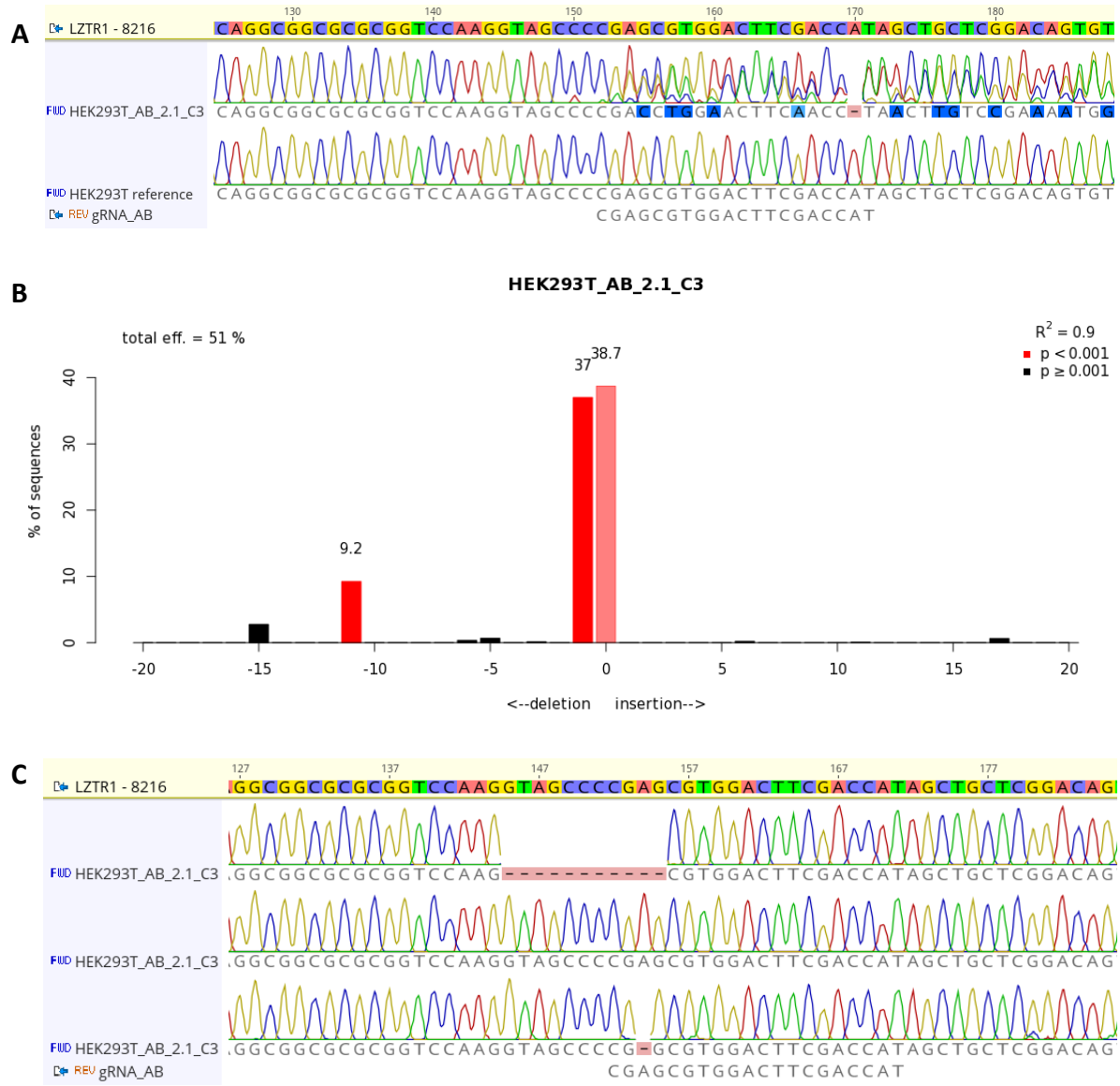

Figure 9: Identification of a HEK293T heterozygous mutant clone using Sanger sequencing, TIDE analysis, and PCR cloning. A: Sanger sequencing of a heterozygous mutant clone showed a change from a clean sequence towards a mixed sequence, at the location where the gRNA targets the *LZTR1* gene. **B:** Decomposition of this sequence by the TIDE software revealed three different alleles present in this clone: a wild type allele, an allele with a 1bp deletion, and an allele with an 11bp deletion. The analysis showed equal amounts of edited and wild type DNA, so the clone was classified as HET 2/4. **C:** To validate the prediction by the TIDE software, we performed PCR cloning, colony PCR, and Sanger sequencing. This enabled us to sequence the individual alleles present in the original cell line. The results showed the same three alleles as predicted by the TIDE software.

For the MRC5 cell line, we identified two wild-type clones, three knock-out clones, and one heterozygous mutant clone (Figure 10 and Table 4). We also identified two clones that could either be knock-out or heterozygous mutant based on the genetic analysis. These clones contain one allele with an indel causing a frameshift, which we confirmed to results in a premature stop codon, and one allele with an indel causing an in-frame deletion of a small part of the LZTR1 protein instead of a frameshift. Whether this deletion of a small part of the LZTR1 protein results in a non-functional protein cannot be predicted by the genetic information only, so further Western Blot analysis should be conducted to determine whether these so-called 'in-frame indels' influence the function of LZTR1.

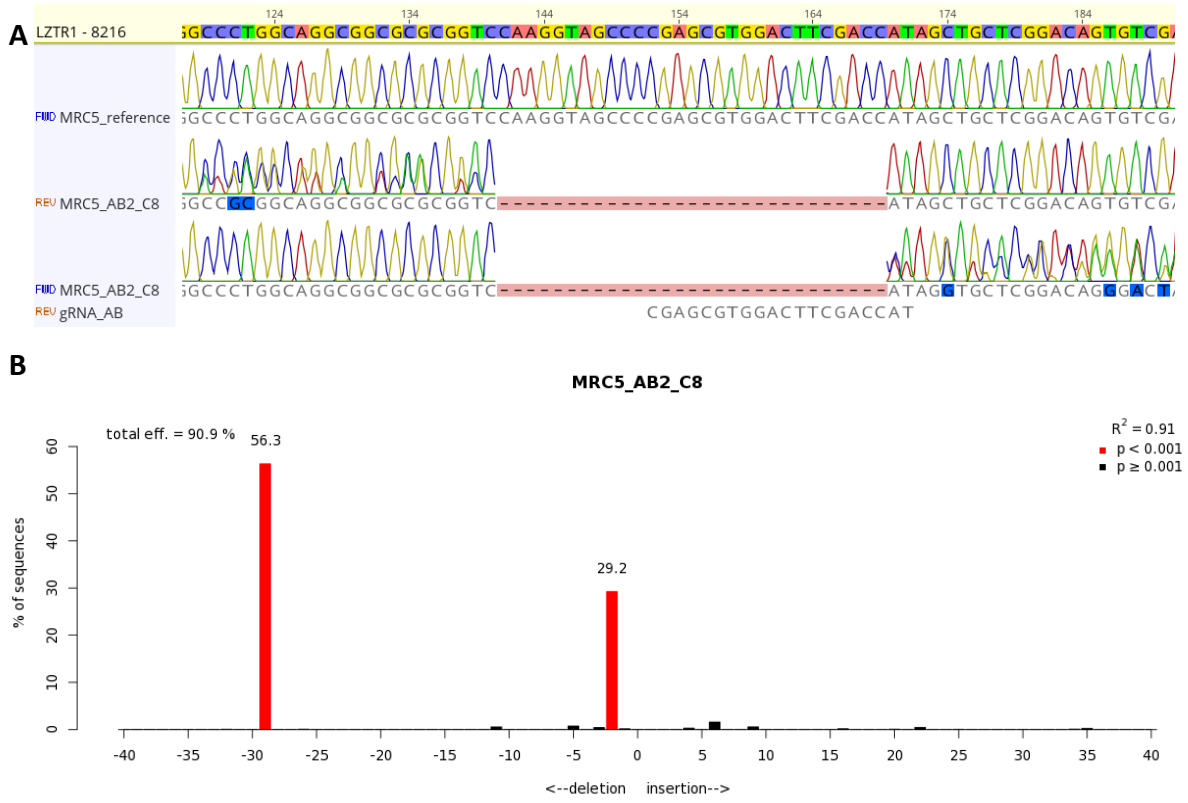

Figure 10: Identification of an MRC5 knock-out clone using Sanger sequencing and TIDE analysis. A: Sanger sequences of an MRC5 knock-out clone showed a large deletion at the expected break site. A clean signal was visible before the deletion and a mixed signal after the deletion, indicating two or more indels present in the cell clone. **B:** Decomposition of the mixed signal by the TIDE software revealed two indels: a 29bp deletion, which was also visible in the Sanger sequences, and a 2bp deletion. Both indels cause a frameshift so the clone was classified as a knock-out.

Table 4: Identified clones and consequences of the indels present in these cell clones.

Clone name	Classification	Indel	cDNA	Amino acid change	RNA sequencing sample identifier
HEK293T_AB_1.1_B5	WT	None	-	-	PMRBM00AYQ
HEK293T_AB_1.1_D4	WT	None	-	-	PMRBM00AYR
HEK293T_AB_1.1_A5	KO	+1bp	c.81dup	p.(Val28Argfs*6)	PMRBM00AYS
		-4bp	c.79_82del	p.(Ser27Trpfs*14)	
		-25bp	c.75_99del	p.(Pro26Alafs*8)	
HEK293T_AB_1.1_D1	KO	+1bp	c.79_80insT	p.(Ser27Metfs*7)	PMRBM00AYT
		-13bp	c.80_92del	p.(Ser27Thrfs*11)	
HEK293T_AB_1.1_B3	HET 3/4	None	-	-	PMRBM00AYU
		-10bp	c.80_89del	p.Ser27fs	

		-28bp	c.64_92del	p.(Ser22Profs*2)	
		-29bp	c.70_98del	p.(Val24Leufs*44)	
HEK293T_AB_1.1_C5	HET 3/4	None	-	-	PMRBM00AYV
		+1bp	c.80_81insG	p.Ser27fs	
		-12bp	c.72_83del	p.(Ala25_Val28del)	
		-13bp	c.80_92del	p.(Ser27Thrfs*11)	
HEK293T_AB_2.1_C3	HET 2/4	None	-	-	PMRBM00AYW
		-1bp	c.79del	p.(Ser27Alafs*15)	
		-11bp	c.70_80del	p.(Val24Argfs*6)	
MRC5_AB2_B8	WT	None	-	-	PMRBM00AZJ
MRC5_AB2_F5	WT	None	-	-	PMRBM00AZK
MRC5_AB2_B7	KO	-7bp	c.78_84del	p.Ser27fs	PMRBM00AZL
		-29bp	c.53_81del	p.Ala20fs	
MRC5_AB2_C8	KO	-2bp	c.80_81del	p.Val28fs	PMRBM00AZM
		-29bp	c.70_98del	p.Val24fs	
MRC5_AB2_F3	KO	-10bp	c.70_79del	p.Val24fs	PMRBM00AZN
		-26bp	c.60_85del	p.Agr21fs	
MRC5_AB2_F7	HET	None	-	-	*NA*
		+1bp	*NA*	*NA*	
MRC5_AB2_C4	HET/KO	-7bp	*NA*	*NA*	*NA*
		-15bp	*NA*	*NA*	
MRC5_AB2_E7	HET/KO	+1bp	*NA*	*NA*	*NA*
		-3bp	*NA*	*NA*	

Functional validation of single cell clones

After identifying the different clones, we used Western Blot to verify the absence of LZTR1 expression and the increase of pan-RAS expression in the knock-out clones compared to the wild type clones. The HEK293T wild type showed a faint band for LZTR1 (Figure 11). The heterozygous mutant clones also showed LZTR1 expression, but background staining in these lanes made it difficult to determine whether the expression is lower than in the wild type clones. A band at the expected height for LZTR1 seemed to be absent in the two knock-out clones. The pan-RAS expression was similar in the knock-out and heterozygous mutant clones but pan-RAS expression seemed very low or even absent in the wild type clones. The bands for β-actin, the housekeeping gene that was used as a loading control, had a comparable intensity for all the samples. Differences in expression for the LZTR1 and RAS proteins can therefore not be explained by a different amount of protein loaded onto the Western Blot.

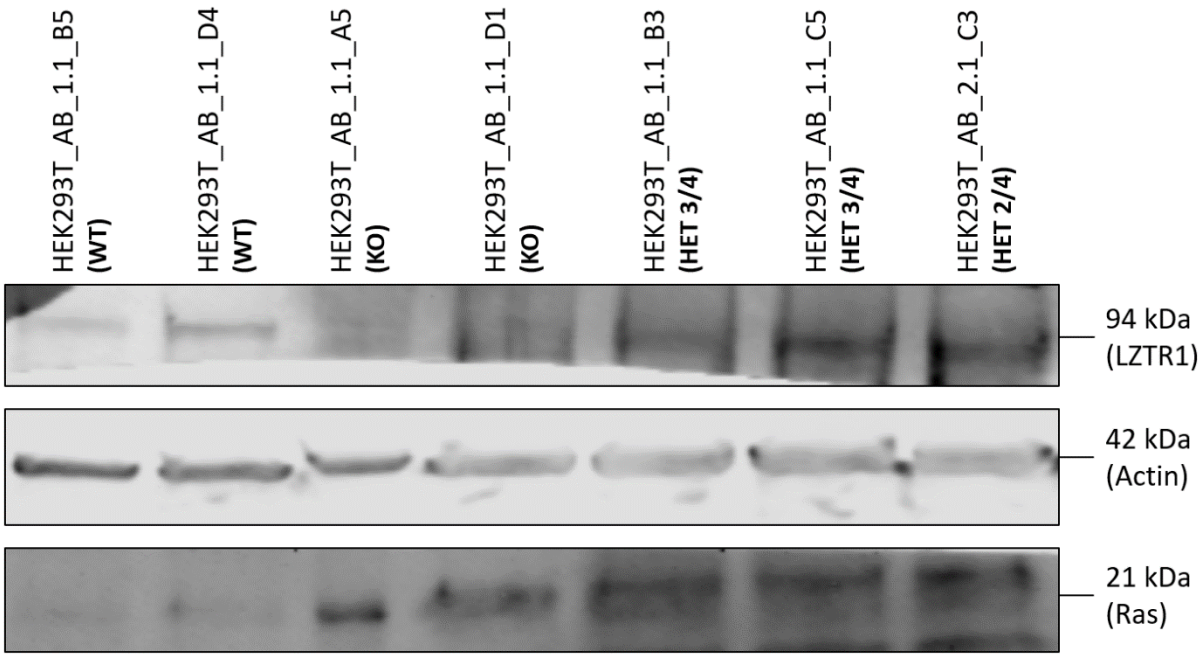

Figure 11: Western Blot for HEK293T selected clones. Top: Staining for LZTR1 showed the expected bands at 94 kDa for the wild type and heterozygous mutant clones, and staining seemed to be absent for both the knock-out clones. **Middle:** β-actin staining was used as a loading control. **Bottom:** Staining for pan-RAS proteins was used to see if we could confirm an increased RAS expression in the LZTR1 knock-out clones. A very faint band is visible for the two wild type clones, while a stronger band is visible for the knock-out and heterozygous mutant clones.
WT = wild type, KO = knock-out, HET 3/4 = heterozygous mutant (3 mutated alleles), HET 2/4 = heterozygous mutant (2 mutated alleles)

The MRC5 wild type clones and heterozygous mutant clone all showed a band for LZTR1 on the Western Blot (Figure 12), but the bands were too faint to distinguish a difference in expression between wild type and heterozygous mutant. LZTR1 expression appeared to be absent in two of the knock-out clones, so these clones can indeed be classified as *LZTR1* knock-out clones. One of the supposed knock-out clones, MRC5_AB2_B7, showed a very faint band at the expected height for LZTR1. This band had the same intensity as a non-specific band visible underneath it, so it could either be residual LZTR1 expression in the knock-out clone or non-specific staining. The clones that could either be knock-out or heterozygous mutant based on the genetic validation all show LZTR1 expression on the Western Blot. The deletion of a small part of the LZTR1 protein still seems to result a detectable protein, so these clones were classified as heterozygous mutant. The pan-RAS staining did not show the expected results. The wild type and knock-out clones all had the same pan-RAS expression, while a stronger band for the knock-out clones would be expected in absence of LZTR1. The three heterozygous mutant clones all had a different pan-RAS expression. The MRC5_AB2_C4 clone showed a pan-RAS overexpression compared to the other samples, MRC5_AB2_E7 showed almost no RAS expression, and MRC5_AB2_F7 was similar in expression to the wild type and knock-out clones. The bands for β-actin had a comparable intensity for all the samples.

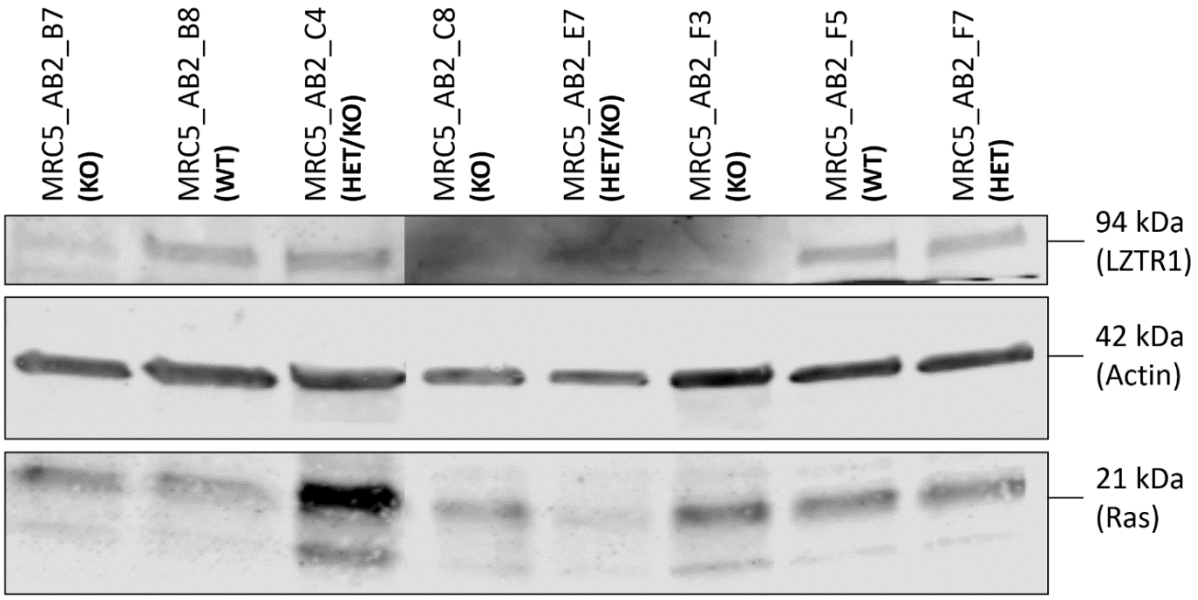

Figure 12: Western Blot for MRC5 selected clones. Top: Staining for LZTR1 showed the expected bands at 94 kDa for the wild type and heterozygous mutant clones. Staining seemed to be absent for the C8 and F3 knock-out clone, but a very faint band is visible for the B7 knock-out clone. A band for LZTR1 was visible for the clones that could either be knock-out or heterozygous mutant, based on the genetic validation. **Middle:** β-actin staining was used as a loading control. **Bottom:** Staining for pan-RAS proteins was used to see if we could confirm an increased RAS expression in the LZTR1 knock-out clones. This staining did not show the expected results. The E7 heterozygous mutant clone showed almost no pan-RAS expression, while the C4 heterozygous mutant clone showed a very high pan-RAS expression. All the other samples showed a similar pan-RAS expression.
WT = wild type, KO = knock-out, HET = heterozygous mutant, HET/KO = heterozygous mutant or knock-out

RNA sequencing

After the RNA sequencing data was processed, we fist looked at the *LZTR1* expression for all the samples. We would expect to see a small decrease of *LZTR1* expression in the knock-out clones compared to the wild type clones. The expression in the HEK293T samples was not in line with this expectation (Figure 13A, correlation between RNA sequencing sample identifier and cell line sample can be found in Table 4). The two wild type clones had a relatively low *LZTR1* expression. One of the knock-out clones had a lower expression than the wild type clones, but the other knock-out clone had a much higher expression than the wild type clones. The three heterozygous mutant clones varied from very low to very high. We then performed hierarchical clustering analysis, which is based in the expression of all genes in the samples. This analysis did not show a clustering of the wild type clones together, the knock-out clones together, and the heterozygous mutant clones together. Instead, the samples were clustered similar to their *LZTR1* expression (Figure 13B).

We did not want to discard all the results from the HEK293T cell line yet, so we decided to use one wild type clone (PMRBM00AYR) and one knock-out clone (PMRBM00AYS) in the further analyses of the RNA sequencing data. The expression results of these two clones were most in line with our expectations.

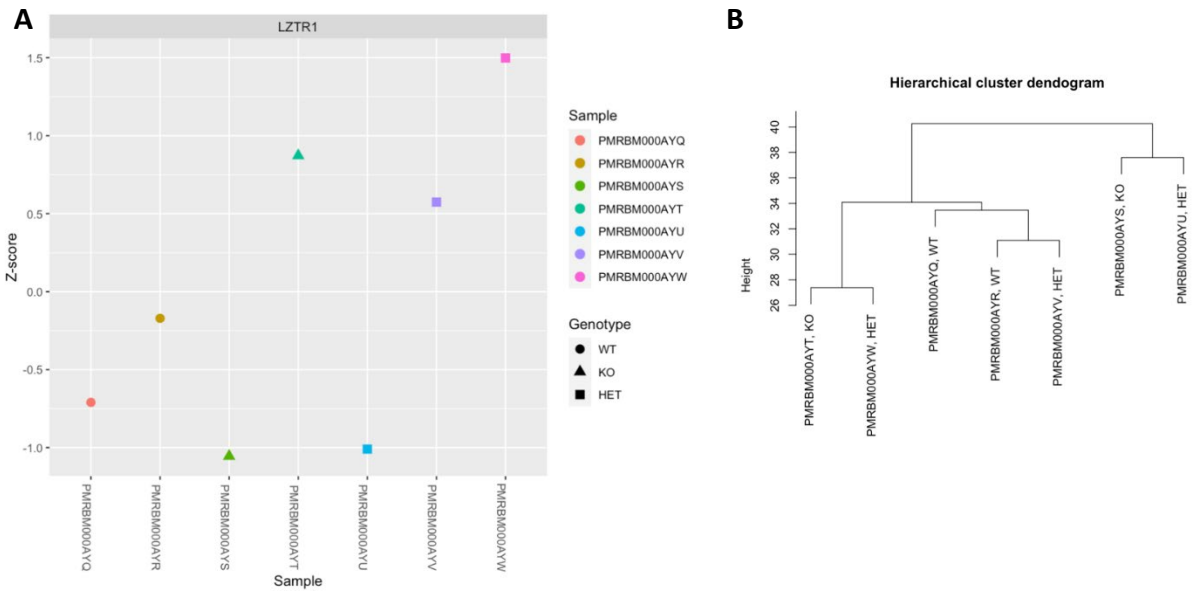

Figure 13: *LZTR1* expression and hierarchical clustering results for the HEK293T clones. A: The two wild type, two knock-out, and three heterozygous mutant clones showed varying levels of *LZTR1* expressions. **B:** Hierarchical clustering analysis of the clones showed that the clones were not clustered together as we would have expected them to be, which would be the wild type clones together, the knock-out clones together, and the heterozygous mutant clones together.
Table 4 shows which cell line sample belongs to each of the RNA sequencing sample identifiers.

The results from the MRC5 clones were more in line with what we had expected. The two wild type clones had a relatively high *LZTR1* expression, and two of the three knock-out clones had a relatively low expression (Figure 14). One of the knock-out clones showed a high *LZTR1* expression, which was the same clone that showed a faint band for LZTR1 in the Western blot. This suggests the presence of residual *LZTR1* expression, so we decided to leave this clone out of future analyses.

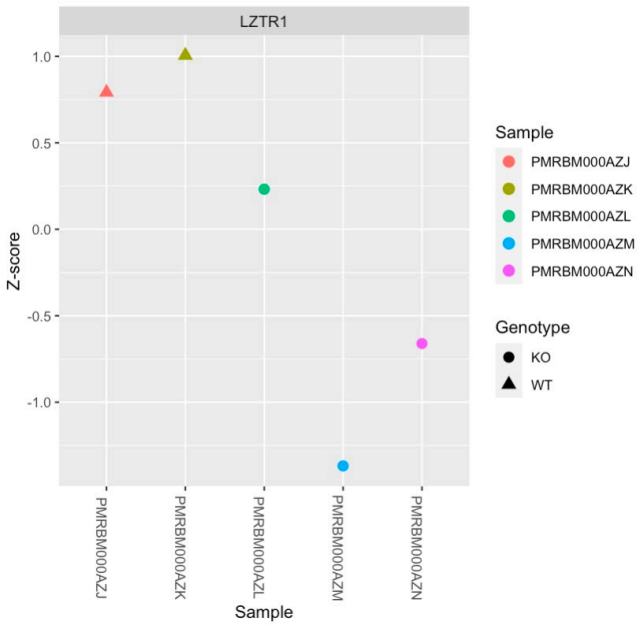

Figure 14: *LZTR1* expression of the MRC5 clones. The two wild type clones showed relatively high expression of LZTR1 while two of the three knock-out clones showed relatively low *LZTR1* expression. The *LZTR1* expression of the last knock-out clone was in between the wild type clones and the other two knock-out clones.

Next, we continued with the manual selection of potentially interesting genes from the MRC5 samples. After comparing the two knock-out clones separately to the wild type clones and selecting the overlapping genes, a list of 6939 upregulated and 2821 downregulated genes was left (Figure 15). After performing the selection of genes with a similar expression between the two wild type clones and between the two knock-out clones, we reduced the list to 1049 upregulated and 200 downregulated genes. We then looked at these specific genes in the HEK293T knock-out clone compared to the HEK293T wild type clone, to see if these genes have also changed expression in the HEK293T cell line. This comparison resulted in a final list of 444 genes that are upregulated and 126 genes that are downregulated in the knock-out clones compared to the wild-type clones for both the HEK293T and the MRC5 cell lines.

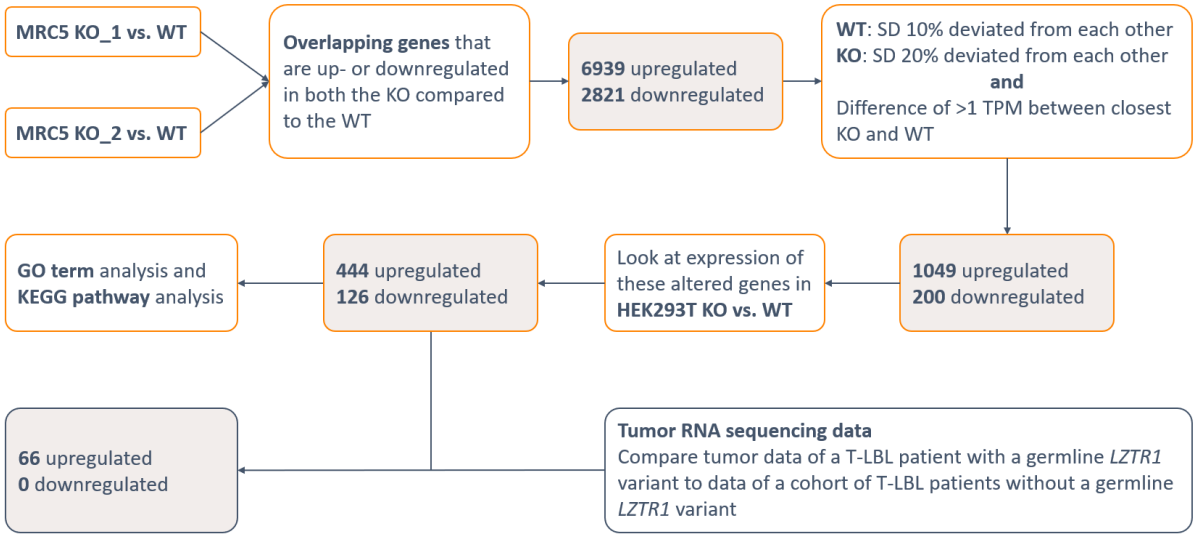

Figure 15: Results of the gene selection from the RNA sequencing data. We performed a manual selection of potentially interesting genes with an altered expression in absence of LZTR1 from the RNA sequencing data of the MRC5 and HEK293T cell lines.

The KEGG pathway analysis that we performed on the final list of selected genes showed that, for both the upregulated and downregulated genes, most genes belong to metabolic pathways. However, the results from this analysis are not significant because the list of selected genes that was used for this analysis is not based on significant results due to the small sample size. We can therefore not draw any definitive conclusions on certain pathways that might be significantly up- or downregulated in absence of LZTR1.

For the comparison between the cell lines and the patient data we started out with one sample of a patient with a T-LBL without a second hit in *LZTR1* in the tumor. We first compared this patient to a cohort of T-LBL patients without a germline *LZTR1* variant and looked at the *LZTR1* expression. This showed that the patient has a relatively high LZTR1 expression compared to the cohort (Figure 15).

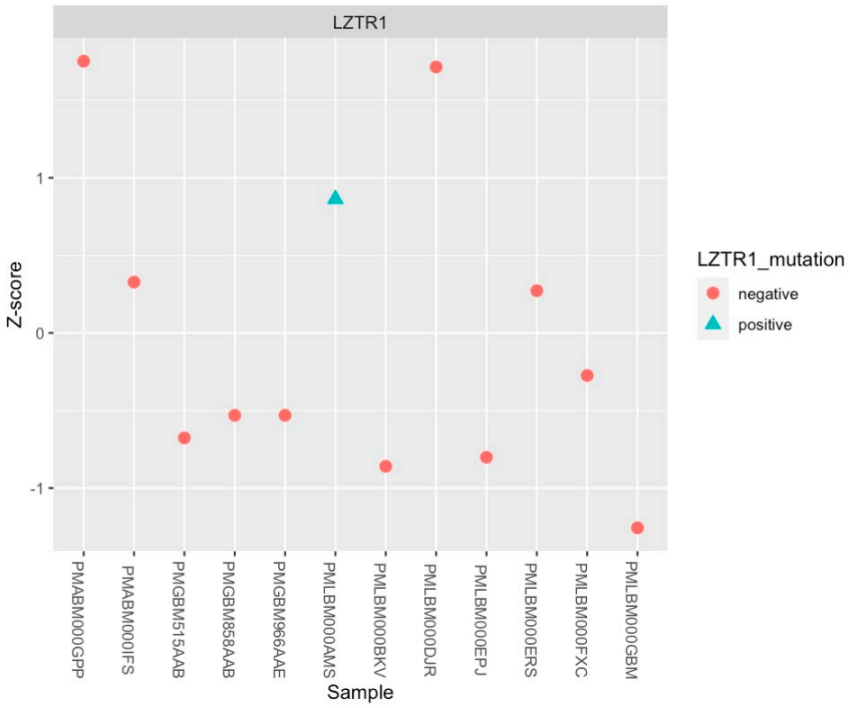

Figure 16: *LZTR1* expression of a T-LBL patient with a germline *LZTR1* variant. This patient has a germline heterozygous *LZTR1* variant but no somatic *LZTR1* mutation in the tumor. The RNA expression of *LZTR1* for the patient was compared to a cohort of T-LBL patients without a germline *LZTR1* variant.

We then compared the list of selected genes from the cell lines to the RNA sequencing data of the T-LBL patient. This showed that 66 of the genes from the selected genes are also upregulated in the patient compared to the cohort, but none of the genes are downregulated in the patient (Figure 15).

Discussion & conclusion

The aim of this project was to investigate whether heterozygous germline *LZTR1* variants, which are known to predispose to Schwannomatosis, could be a potential cancer predisposition gene in children. To do this, we aimed to investigate the role of *LZTR1 in vitro,* by using cells from different tissue origins and create cell lines with various levels of *LZTR1* expression. By performing RNA sequencing on these cell lines, we created expression profiles in the presence and absence of LZTR1. These profiles were compared to RNA sequencing data of patients, to see whether the germline *LZTR1* variant has potentially contributed to the tumor formation.

We successfully transfected both the HEK293T and MRC5 cell lines with the CRISPR-Cas9 RNP complexes. We did encounter a problem with the HEK293T cells during the experiment, because the HEK293T cells were not in good condition the day after the transfection. It appeared that most of the cells had died in the middle of the well. This could be the result of the plate being swirled around to distribute the cells, instead of moving the plate in a T-shape. Fortunately, the cells seemed to be recovering over the coming days, so we were able to continue with these cells instead of having to perform the transfection again. The transfection efficiency and editing efficiency was high for both cell lines, so we could continue with the single cell sorting and subsequent identification of the clones.

We were able to identify multiple wild type and knock-out clones and a few heterozygous mutant clones for both cell lines. The genetic validation of the clones did turn out to be more challenging than expected. When using the TIDE software, we would expect to see four peaks of approximately 25% for the HEK293T cell clones, which would correlate with each of the four *LZTR1* alleles. However, we often observed five peaks indicating more than four alleles, which could be caused by the presence of multiple clones instead of a single cell clone. We sequenced the cell clones at various time points and the number and percentages of the peaks remained the same throughout time, which would make it unlikely that there were multiple clones instead of one. We also observed clones with three or four peaks but with percentages that would not correspond with four times 25%. This made it difficult to identify heterozygous mutant clones, with an equal amount of wild type alleles and edited alleles, for the HEK293T cell line.

We observed the same issue with the MRC5 clones. We would expect to see two clones of approximately 50% each, but we often observed either three peaks or two peaks with unequal heights. It could be that the TIDE software is unable to accurately determine the proportion of a certain indel in the Sanger sequences of single cell clones. During the validation that we performed, we only looked at which indels are present in the analysis but not at the ratio between the different indels. We would need to perform the PCR cloning experiment at a larger scale to be able to decide whether the ratio between the alleles is equal or not. We did not do this because the aim of our PCR cloning experiment was to verify if the different indels determined by the TIDE software were correct, and not to verify if the percentages of the different indels, thereby the ratio of the different alleles, were correct.

The Western Blot results confirmed that, in line with the genetic validation, the LZTR1 proteins was indeed absent from most of the knock-out clones. However, we did encounter a lot of background staining, making the analyses difficult. A few clones that were labeled as knock-out based on the genetic validation showed a faint band for LZTR1 on the Western Blot, which could indicate the presence of detectable, but not necessarily functional, LZTR1 proteins. This could be possible when a new start codon is present after the CRISPR-Cas9 target site, resulting in a truncated but still detectable LZTR1 protein. However, previous research with LZTR1 knock-out clones was performed with a gRNA targeting exon 1 as well, and there were no reports of any remaining LZTR1 proteins in the knock-out clones (12). It might also be possible that the DNA at the CRISPR-Cas9 target site has

changed over time, regaining the function of the LZTR1 protein. However, as mentioned before, we sequenced the single cell clones at multiple time points and no changes in the DNA were detected, making this an unlikely scenario. Furthermore, the RNA sequencing data showed that the exact indels determined by the TIDE software, which would result in a frameshift, are present in the knock-out clones. There were also no wild type reads present in the RNA sequencing data for the knock-out clones, indicating that the presence of residual LZTR1 proteins is unlikely. It is also possible that the faint band for LZTR1 is the results of a technical issue with the Western Blot, and that the band we observe is a result of non-specific staining instead of actual LZTR1 proteins being present. The bands we see have the same intensity as other non-specific bands on the Western Blot, and they are a lot fainter than the LZTR1 staining for the wild type and heterozygous mutant clones. The anti-LZTR1 antibody that we used has also been used in published work before without the background staining or non-specific staining that we observed (12). It might therefore be useful to optimize our Western Blot protocol for this antibody.

The RNA isolation process went well, and the quality and quantity of the RNA was good for all the samples. We confirmed that the same indels detected by the genetic validation were present in the RNA sequencing data that we obtained from the diagnostics department. However, the sample size was too small to perform differential expression analysis, as it requires a minimal sample size of five for each category. We still wanted to see if any interesting changes in gene expression were present in our data, so we performed a manual selection of up- and downregulated genes. We expected to see an increase in RAS expression because it is already known that an absence of LZTR1 increases RAS pathway expression (12). When we looked specifically at the RNA expression of the RAS genes, we did not see a difference between the knock-out and the wild type clones. However, the ultimate effect of *LZTR1* on the RAS pathway is on a protein level. The LZTR1 protein binds the RAS proteins, resulting in proteasomal degradation of the RAS proteins. The RAS proteins accumulate in the absence of LZTR1 causing an overexpression of the downstream pathway, which is a well-known cancer driving event (15,16). It might therefore be likely that the direct effect of the absence of LZTR1 in the knock-out clones is not visible in RNA sequencing data, but only with a Western Blot or proteomics analysis. This is supported by previous research where RNA sequencing and proteomics analysis was performed on *LZTR1* knock-out cardiomyocytes (25). This research shows that there is no significant change in RNA expression for *LZTR1* and *RAS*, but that the RAS proteins are significantly upregulated in the proteomics analysis and on Western Blot. Performing the KEGG pathway analysis also shows that even though the RNA expression of the individual RAS proteins is not increased, the RAS pathway in its entirety is upregulated. This indicates that direct target proteins of LZTR1 may not have an altered RNA expression and that the direct effect of LZTR1 removal is only visible when performing proteomics or Western Blot, but that further downstream effects of the absence of LZTR1 can be visible on RNA sequencing data. However, our Western Blot results did not show an increased expression of RAS proteins in the knock-out clones compared to the wild type clones, even though LZTR1 was absent in the knock-out clones. This might be caused by a technical issue during the Western Blot experiment. The texture of the membrane was visible in the bands for the RAS proteins as well as in the background staining, indicating that something might have gone wrong during one of the steps of the Western Blot protocol. This unusual staining was not visible in the LZTR1 or the β-actin bands. Repeating the Western Blot would be good to improve the quality of our results to see if the overexpression of the RAS proteins is present in our samples after all.

We decided to perform RNA sequencing of our cell lines because we were able to get RNA sequencing data from the tumors of patients with a germline *LZTR1* variant. This made it easy to directly compare *in vitro* results with tumor data from patients, so we could hopefully draw some conclusions on the involvement of the germline *LZTR1* variant in tumor development in the patient.

So far, we don't have any significant results from our RNA sequencing data and from the comparison with patient data because our sample size is too small. For future analyses we should consider that, with the RNA sequencing data, we are looking at downstream effects of an *LZTR1* knock-out and we will most likely not find any direct target genes affected by *LZTR1*.

For the patient data analysis, we so far collected 24 cases of pediatric cancer with a germline heterozygous pathogenic *LZTR1* variant and we obtained sequencing data from 6 of these cases. This sequencing data showed that only one of the six patients has a second hit in *LZTR1* in the tumor. In general, a two-hit inactivation would be expected for a tumor suppressor gene to contribute to tumor development. This is also the case for somatic mutations in *LZTR1* in glioblastomas, where biallelic mutations result in a complete absence of LZTR1, but it does not seem to be necessary for these cases of pediatric cancer. It might be possible that the heterozygous variant in *LZTR1* is sufficient to contribute to tumor development. However, because heterozygous germline *LZTR1* variants are also present in the population in healthy individuals, it does not seem likely that a heterozygous *LZTR1* variant alone is enough to initiate tumor development. This incomplete penetrance might suggest that an additional mutational process is required before a tumor can develop during childhood.

One possibility is that a somatic *LZTR1* mutation in the tumor is necessary after all. The presence of a second hit in *LZTR1* in only one out of six tumors could indicate that not all germline heterozygous *LZTR1* variants have contributed to the tumor development. It might be possible that for some cases it is just a coincidence that we have identified a germline *LZTR1* variant, but that the variant did not actually contribute to tumor development. This does raise the question whether only the germline variants in combination with a second hit in *LZTR1* in the tumor contribute to tumor development, or whether other variants have contributed through a different mechanism. It might be possible that the type of mutation or the location of the mutation is important for the mechanism of disease. This would be similar to Noonan syndrome, where the autosomal dominant variant is caused by mutations in the Kelch domain of *LZTR1* while the autosomal recessive variant can be cause by mutations throughout the entire gene (21).

Another possibility is that, like the development of Schwannomatosis, another more complex mutational mechanism needs to occur in addition to the germline *LZTR1* variant. We have not investigated this yet, but it could be interesting in the future to compare tumor data from the different patients with a germline *LZTR1* variant with each other to find overlapping alterations besides the *LZTR1* variant. This might be a difficult comparison because of the various tumor types, each with unique characteristics and mutations present in the tumors, and low sample size of our data cohort. However, because we are aiming to find a general process that has occurred in all patients regardless of tumor type, it might still be interesting to investigate further.

In conclusion, we managed to create *LZTR1* knock-out and wild type cell clones for two different cell lines. We validated these clones on DNA and protein level and performed RNA sequencing on the clones. The sample size for the RNA sequencing data is too low to be able to select target genes or pathways that have significantly changed expression in absence of LZTR1. We can therefore not compare the cell line data with the patient data yet, and thus cannot draw any conclusions on the involvement of the germline *LZTR1* variants in tumor development. However, there are still some interesting questions remaining that could be worth investigating in the future.

Future perspectives

The most important improvement for this project would be to increase the power of the RNA sequencing analysis by increasing the number of samples we have so far. The RNA sequencing analysis has been performed using two wild type and two knock-out clones for the MRC5 cell line, and only one wild type and one knock-out clone for the HEK293T cell line. However, we would need at least five wild type and five knock-out clones to calculate p-values. We need these values to perform differential expression analysis so we can determine which genes are significantly up- or downregulated in the knock-out clones, instead of the manual gene selection that was performed so far. We also need to perform the differential expression analysis to get relevant results from the KEGG pathway analysis to determine if certain pathways are significantly up- or downregulated in the knock-out clones.

Besides increasing the number of wild type and knock-out clones, we also need to identify more heterozygous mutant clones. One of the main things that we were interested in was the effect of a heterozygous *LZTR1* variant and whether these variants result in an expression profile more like the wild type clones or more like the knock-out clones. However, our analysis of the cell lines has thus far only been performed on the wild type and knock-out clones. We first wanted to see whether we could find a significant difference in expression between the wild type and knock-out clones, before comparing the heterozygous mutant clones with these results. We did not analyze the heterozygous mutant clones because we did not have any significant results for the comparison between the wild type and knock-out clones yet, though the analysis of the heterozygous mutant clones would be most interesting to answer our research questions.

A further improvement of the cell line analyses could be to add a clone to each cell line with an LZTR1 overexpression. Ideally, these cells should show the opposite effect compared to a knock-out clone. If a certain gene is upregulated in the *LZTR1* knock-out clone and downregulated in the *LZTR1* overexpressed clone, it is more plausible that the change in expression is the result of the difference in *LZTR1* expression.

Another addition to the project could be to increase the number of different cell lines that we use. The original aim of the study was to use at least three cell lines from various tissue origins because the germline *LZTR1* variants are found in various types of tumors. By using multiple cells lines, we would be able to see the general effect of LZTR1 removal on various cell types, thereby eliminating cell type specific effects of LZTR1. For now, we managed to use two cell lines, the HEK293T cells which are human kidney endothelial cells, and the MRC5 cells which are human lung fibroblasts. It might be interesting to expand the analysis with a blood cell line and neural cells, because of the hematological cancers and brain tumors in patients with a germline *LZTR1* variant.

One thing to consider when selecting new cell lines is to choose diploid cells. Our results from the HEK293T cell line have proven the difficulty in creating a heterozygous mutant cell line when there are more than two alleles of *LZTR1* present. Because the heterozygous mutant cell clones are important for answering our research question, it would be useful to be able to add more of those samples to the analysis.

Lastly, it might be useful to add a proteomics analysis of the cell lines besides the RNA sequencing analysis. Ultimately, the effect of LZTR1 is on a protein level because it can bind certain target proteins which will then be tagged for proteasomal degradation. The direct effect of the absence of LZTR1 might therefore not be visible on RNA sequencing level, and the RNA sequencing data of our cell clones might only show further downstream effects. Adding proteomics to the project could therefore be an interesting addition to see the direct effect of LZTR1 as well.

Acknowledgements

I would first like to thank my supervisor Nienke van Engelen for guiding me through this project for the past year and supporting me throughout the practical work. The open discussions we had together, combined with the given freedom to perform the experiments in my own way, have helped me a lot in becoming an independent researcher. I would also like to thank Jette Bakhuizen for assisting me in the clinical part of this research and providing the clinical data used within the project. Next, I want to thank Roland Kuiper for enabling this research and for welcoming me into his group, and providing a pleasant working environment together with the rest of the Kuiper group. I would also like to thank Mariangela Sabatella for teaching me the practical skills needed to perform the experiments and patiently answering the many questions I had, as well as Michelle Kleisman who processed all the RNA sequencing data for me. Lastly, I would like to thank Laurens van der Meer and Britt Vervoort from the van Leeuwen group for their willingness to support us by providing materials, practical knowledge, and theoretical insights for this project, as well as Willem Cox for operating the FACS machine for us. This, together with the insightful discussions regarding the direction of the project, helped us greatly in starting up this new line of research.

References

1. Jacquinet A, Bonnard A, Capri Y, Martin D, Sadzot B, Bianchi E, et al. Oligo-astrocytoma in LZTR1-related Noonan syndrome. Eur J Med Genet. 2020 Jan;63(1):103617.

2. Gröbner SN, Worst BC, Weischenfeldt J, Buchhalter I, Kleinheinz K, Rudneva VA, et al. The landscape of genomic alterations across childhood cancers. Nature. 2018 Mar 15;555(7696):321–7.

3. Foss-Skiftesvik J, Stoltze UK, van Overeem Hansen T, Ahlborn LB, Sørensen E, Ostrowski SR, et al. Redefining germline predisposition in children with molecularly characterized ependymoma: a population-based 20-year cohort. Acta Neuropathol Commun. 2022 Aug 25;10(1):123.

4. Chinton J, Huckstadt V, Mucciolo M, Lepri F, Novelli A, Gravina LP, et al. Providing more evidence on LZTR1 variants in Noonan syndrome patients. Am J Med Genet Part A. 2020 Feb 11;182(2):409–14.

5. Byrjalsen A, Hansen TVO, Stoltze UK, Mehrjouy MM, Barnkob NM, Hjalgrim LL, et al. Nationwide germline whole genome sequencing of 198 consecutive pediatric cancer patients reveals a high incidence of cancer prone syndromes. Eng C, editor. PLOS Genet. 2020 Dec 17;16(12):e1009231.

6. Wong M, Mayoh C, Lau LMS, Khuong-Quang D-A, Pinese M, Kumar A, et al. Whole genome, transcriptome and methylome profiling enhances actionable target discovery in high-risk pediatric cancer. Nat Med. 2020 Nov 5;26(11):1742–53.

7. van Tilburg CM, Pfaff E, Pajtler KW, Langenberg KPS, Fiesel P, Jones BC, et al. The Pediatric Precision Oncology INFORM Registry: Clinical Outcome and Benefit for Patients with Very High-Evidence Targets. Cancer Discov. 2021 Nov 1;11(11):2764–79.

8. Johnston JJ, van der Smagt JJ, Rosenfeld JA, Pagnamenta AT, Alswaid A, Baker EH, et al. Autosomal recessive Noonan syndrome associated with biallelic LZTR1 variants. Genet Med. 2018 Oct;20(10):1175–85.

9. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020 May 28;581(7809):434–43.

10. Frattini V, Trifonov V, Chan JM, Castano A, Lia M, Abate F, et al. The integrated landscape of driver genomic alterations in glioblastoma. Nat Genet. 2013 Oct 5;45(10):1141–9.

11. Nacak TG, Leptien K, Fellner D, Augustin HG, Kroll J. The BTB-kelch protein LZTR-1 is a novel Golgi protein that is degraded upon induction of apoptosis. J Biol Chem. 2006 Feb;281(8):5065–71.

12. Abe T, Umeki I, Kanno S, Inoue S, Niihori T, Aoki Y. LZTR1 facilitates polyubiquitination and degradation of RAS-GTPases. Cell Death Differ. 2020 Mar 23;27(3):1023–35.

13. Steklov M, Pandolfi S, Baietti MF, Batiuk A, Carai P, Najm P, et al. Mutations in LZTR1 drive human disease by dysregulating RAS ubiquitination. Science (80-). 2018;362(6419):1177–82.

14. Bigenzahn JW, Collu GM, Kartnig F, Pieraks M, Vladimer GI, Heinz LX, et al. LZTR1 is a regulator of RAS ubiquitination and signaling. Science (80-). 2018;362(6419):1171–7.

15. Ney GM, McKay L, Koschmann C, Mody R, Li Q. The Emerging Role of Ras Pathway Signaling in Pediatric Cancer. Cancer Res. 2020 Dec 1;80(23):5155–63.

16. Simanshu DK, Nissley D V., McCormick F. RAS Proteins and Their Regulators in Human Disease. Cell. 2017 Jun;170(1):17–33.

17. Hutter S, Piro RM, Reuss DE, Hovestadt V, Sahm F, Farschtschi S, et al. Whole exome sequencing reveals that the majority of schwannomatosis cases remain unexplained after excluding SMARCB1 and LZTR1 germline variants. Acta Neuropathol. 2014 Sep;128(3):449–52.

18. Piotrowski A, Xie J, Liu YF, Poplawski AB, Gomes AR, Madanecki P, et al. Germline loss-of-function mutations in LZTR1 predispose to an inherited disorder of multiple schwannomas. Nat Genet. 2014 Feb 22;46(2):182–7.

19. Roberts AE, Allanson JE, Tartaglia M, Gelb BD. Noonan syndrome. Lancet. 2013 Jan;381(9863):333–42.

20. Yamamoto GL, Aguena M, Gos M, Hung C, Pilch J, Fahiminiya S, et al. Rare variants in SOS2 and LZTR1 are associated with Noonan syndrome. J Med Genet. 2015 Jun;52(6):413–21.

21. Motta M, Fidan M, Bellacchio E, Pantaleoni F, Schneider-Heieck K, Coppola S, et al. Dominant Noonan syndrome-causing LZTR1 mutations specifically affect the Kelch domain substrate-recognition surface and enhance RAS-MAPK signaling. Hum Mol Genet. 2019 Mar 15;28(6):1007–22.

22. Alt-R ® CRISPR-Cas9 System: Cationic lipid delivery of CRISPR ribonucleoprotein complex into mammalian cells. Integrated DNA Technologies. 2019.

23. Brinkman EK, Chen T, Amendola M, Van Steensel B. Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Res. 2014;42(22):1–8.

24. pGEM-T and pGEM-T Easy Vector Systems; Instructions for use of products A1360, A1380, A3600 and A3610. Promega Corporation. 2013.

25. Hanses U, Kleinsorge M, Roos L, Yigit G, Li Y, Barbarics B, et al. Intronic CRISPR Repair in a Preclinical Model of Noonan Syndrome–Associated Cardiomyopathy. Circulation. 2020 Sep 15;142(11):1059–76.

Table S1: Overview of the collected cases. BCP-ALL = B-cell Precursor Acute Lymphoblastic Leukemia; T-LBL = T-cell Lymphoblastic Lymphoma; PTLD = Post-Transplant Lymphoproliferative Disease; ATRT = Atypical Teratoid Rhabdoid Tumor; CNS = central nervous system; AML = Acute Myeloid Leukemia; NA = Not available; AR = autosomal recessive

Case	Malignancy	Germline variant	Germline protein change	Germline effect	Tumor variant	Tumor protein change	Noonan Syndrome
1	BCP-ALL	c.200G>C	p.Arg67Pro	Missense	No		
2	T-LBL	c.2326-11T>A	-	Splice variant	No		
3	T-LBL	c.1135C>T	p.Gln379*	Nonsense	No		
4	PTLD	*NA*	*NA*	*NA*	*NA*		
5	Neuroblastoma		p.Pro735fs*	Frameshift	Yes, c.644G>T	p.Trp215Leu	
6	Nephroblastoma	c.2324A>C	p.Gln775Pro	Missense	No		
7	Diffuse midline glioma	c.2407-1G>A	-	Splice variant	No		
8	ATRT (CNS)	c.1653C>A	p.Tyr551*	Nonsense	No		
9	Papillary Thyroid Carcinoma	c.974delA	p.Gln325fs*	Frameshift	No		
10	ALL	c.742G>A	p.Gly248Arg	Missense			Yes
11	Neuroblastoma		p.Tyr119*	Nonsense			
12	Retinoblastoma	c.628C>T	p.Arg210*	Nonsense			
13	ALL	c.628C>T	p.Arg210*	Nonsense			
14	High grade glioma		p.Lys595fs*	Frameshift			
15	B-ALL	c.1687G>C	p.Glu563Gln	Missense			Yes, AR
16	ALL and AML	c.628C>T	p. Arg210*	Nonsense			Yes, AR
16		c.2220-17C>A	-	Splice variant			Yes, AR
17	Sarcoma		p.Val582Profs*6	Frameshift			
18	Osteosarcoma	c.955C>T	p.Gln319*	Nonsense			
19	Oligoastrocytoma	c.850C>T	p.Arg284Cys	Missense			Yes
20	Brain tumor	*NA*	*NA*	*NA*			
21	Brain tumor	*NA*	*NA*	*NA*			
22	Neuroblastoma	c.2205delinsTC	p.Gly737Argfs*45	Frameshift			
23	ATRT	c.263+1G>A	-	Splice varaint			
24	Ependymoma	c.2284C>T	p.Gln762*	Nonsense			