NHLRC2 variants identified in patients with Fibrosis, Neurodegeneration, and Cerebral Angiomatosis (FINCA) – Characterisation of a novel cerebropulmonary disease

Johanna Uusimaa,[1–3,*] Riitta Kaarteenaho,[4,17] Teija Paakkola,[1,3,17] Hannu Tuominen,[5,6,18] Minna K. Karjalainen,[1,2,18] Javad Nadaf,[7,8,18] Teppo Varilo,[9,19] Meri Uusi-Mäkelä,[10,19] Maria Suo-Palosaari,[11] Ilkka Pietilä,[1,3] Anniina E. Hiltunen,[1,3] Lloyd Ruddock,[3,12] Heli Alanen,[3,12] Ekaterina Biterova,[3,12] Ilkka Miinalainen,[3] Annamari Salminen,[1,2] Raija Soininen,[3,12] Aki Manninen,[3,12] Raija Sormunen,[3,5] Mika Kaakinen,[3] Reetta Vuolteenaho,[3] Riitta Herva,[5] Päivi Vieira,[1,2] Teija Dunder,[1,2] Hannaleena Kokkonen,[13,14] Jukka S. Moilanen,[1,15] Heikki Rantala,[1,2] Lawrence M. Nogee,[16] Jacek Majewski,[7] Mika Rämet,[1,2,10] Mikko Hallman,[1,2] and Reetta Hinttala[1–3]

[1]PEDEGO Research Unit and Medical Research Center Oulu, University of Oulu, and Oulu University Hospital, PO Box 5000, FI-90014 Oulu, Finland
[2]Department of Children and Adolescents, Oulu University Hospital, PO Box 23, FI-90029 Oulu, Finland
[3]Biocenter Oulu, University of Oulu, PO Box 5000, FI-90014 Oulu, Finland
[4]Respiratory Research Unit and Medical Research Center Oulu, University of Oulu and Oulu University Hospital; Internal Medicine, Pulmonary Division, University of Oulu and Oulu University Hospital, PO Box 5000, FI-90014 Oulu, Finland
[5]Department of Pathology, Cancer and Translational Medicine Research Unit, University of Oulu, PO Box 5000, FI-90014 Oulu, Finland
[6]Department of Pathology, Oulu University Hospital, PO Box 23, FI-90029 Oulu, Finland
[7]McGill University and Génome Québec Innovation Centre, Montreal, Quebec, QC H3A 0G1, Canada
[8]St Jude Children's Research Hospital (SJCRH), 262 Danny Thomas Place, Memphis, TN 38105, USA
[9]Department of Medical Genetics, University of Helsinki, Haartmaninkatu 8, FI-00251 Helsinki, Finland
[10]BioMediTech Institute and Faculty of Medicine and Life Sciences, University of Tampere, Tampere, Finland
[11]Department of Diagnostic Radiology and Medical Research Center Oulu, Oulu

University Hospital and University of Oulu, PO Box 50, FI-90029 Oulu, Finland
[12]Faculty of Biochemistry and Molecular Medicine, University of Oulu, PO Box 5000, FI-90014 Oulu, Finland
[13]Northern Finland Laboratory Centre NordLab, Oulu University Hospital, PO Box 500, FI-90029 Oulu, Finland
[14] Department of Clinical Chemistry, University Oulu, PO Box 5000, FI-90014 Oulu, Finland
[15]Department of Clinical Genetics, Oulu University Hospital, PO Box 23, FI-90029 Oulu, Finland
[16]Division of Neonatology, Johns Hopkins University School of Medicine, CMSC 6-104A, 600 N. Wolfe St., Baltimore, Maryland 21287, USA

[17]These authors contributed equally to this work.
[18]These authors contributed equally to this work.
[19]These authors contributed equally to this work.

***Corresponding author:**

Johanna Uusimaa, MD, PhD

Professor of Paediatric Neurology

Tel: +358-8-3155819

Fax: +358-8-3155559

E-mail: johanna.uusimaa@oulu.fi

Abstract

A novel multi-organ disease that is fatal in early childhood was identified in three patients from two non-consanguineous families. These children were born asymptomatic but at the age of two months they manifested progressive multi-organ symptoms resembling no previously known disease. The main clinical features included progressive cerebropulmonary symptoms, malabsorption, progressive growth failure, recurrent infections, chronic haemolytic anaemia and transient liver dysfunction. In the affected children, a neuropathology revealed increased angiomatosis-like leptomeningeal, cortical and superficial white matter vascularisation and congestion, vacuolar degeneration and myelin loss in white matter, as well as neuronal degeneration. Interstitial fibrosis and previously undescribed granuloma-like lesions were observed in the lungs. Hepatomegaly, steatosis and collagen accumulation were detected in the liver.

A whole-exome sequencing of the two unrelated families with the affected children revealed the transmission of two heterozygous variants in the NHL repeat-containing protein 2 (NHLRC2); an amino acid substitution p.Asp148Tyr and a frameshift 2-bp deletion p.Arg201GlyfsTer6. NHLRC2 is highly conserved and expressed in multiple organs and its function is unknown. It contains a thioredoxin-like domain, however, an insulin turbidity assay on human recombinant NHLRC2 showed no thioredoxin activity. In patient-derived fibroblasts, NHLRC2 levels were low, and only p.Asp148Tyr was expressed. Therefore, the allele with the frameshift deletion is likely non-functional. Development of the *Nhlrc2* null mouse strain stalled before the morula stage. Morpholino knockdown of *nhlrc2* in zebrafish embryos affected the integrity of cells in the midbrain region.

This is the first description of a fatal, early-onset disease; we have named it FINCA disease based on the combination of pathological features that include fibrosis, neurodegeneration, and cerebral angiomatosis.

Keywords: central nervous system, cerebropulmonary disease, multi-organ disease, interstitial fibrosis, neurodegeneration, brain angiogenesis

Introduction

Recent developments in network-based approaches have led to the discovery of disease-causing genes, the unexpected links between apparently unrelated diseases, diagnostic biomarkers and therapeutic targets for cancer, diabetes and several neurodegenerative diseases including Alzheimer's, Parkinson's, and Huntington's diseases [1,22]. Rare, inherited mutations causing familial neurodegenerative diseases provide the molecular basis for the cellular pathways underlying the pathogenesis of diseases, including the accumulation of aberrant or misfolded proteins, protofibril formation, ubiquitin-proteasome system dysfunction, excitotoxic insult, oxidative and nitrosative stress, mitochondrial injury, synaptic failure, altered metal homeostasis and failure of axonal and dendritic transport [2]. Chronic, progressive pulmonary disorders include interstitial fibrosis, granulomatous changes and diffuse alveolar damage caused by various pathogenetic and genetic mechanisms [25]. However, the current understanding of the molecular basis of degenerative disease–related mechanisms, such as tissue fibrosis, neurodegeneration and angiogenesis, remains incomplete.

The research field focusing on molecular genetic aetiologies and cellular mechanisms of inherited human diseases has undergone a major revolution in the last decade because of modern research technology that includes whole-exome sequencing (WES) and bioinformatic analyses, which provide better means to identify novel proteins, cellular pathways and disease mechanisms. Herein, we characterise a novel, early-onset and progressive multi-organ disease presenting mainly as severe cerebropulmonary manifestations that eventually led to death in infancy. We present the clinical, radiological and histological characterisations of this multi-organ disease, which we named FINCA disease based on the unique histopathological findings of the patients (fibrosis, neurodegeneration, and cerebral angiomatosis). According to our WES analysis, all three affected children with FINCA disease harboured an identical combination of variants in the NHL repeat containing 2 (*NHLRC2*) gene encoding NHLRC2, a protein with a thioredoxin-like domain. However, its function is currently uncharacterized, enabling no direct biochemical assay to address the pathological role of the variants identified. To acquire more information about the function of NHLRC2 *in vivo*, we studied the effects of knocking out or down the expression of *Nhlrc2* in mice and zebrafish models, respectively. Finally, our data suggests that NHLRC2 is an important factor in the maintenance of multi-organ homeostasis, indicating a wider role for NHLRC2 in the pathogenesis of degenerative human diseases.

Materials and methods

Study subjects

As part of the diagnostic protocol, skin and skeletal muscle biopsy samples were collected from three patients with severe, early-onset undefined progressive cerebropulmonary and multi-organ disease. The patients were examined at the Department of Paediatrics of Oulu University Hospital from 2001–2003 (Patients 1 and 2 from Family A) and 2010–2011 (Patient 3 from Family B). Investigations included clinical assessments and radiological, histological, biochemical and molecular genetic analyses. Based on the severe, multi-organ phenotype, a mitochondrial disease was suspected, and muscle biopsies were performed; however, there were no specific abnormal findings. Mitochondrial respiratory chain enzyme activities were normal, there were no deletions in muscle mitochondrial DNA and the sequencing of the coding region of mtDNA revealed no pathogenic variants or exons and exon–intron boundaries of the genes encoding mitochondrial twinkle helicase and polymerase gamma. Furthermore, *CLN1*, *NKX2-1*, *SFTPB*, *ABCA-3* and *SFTPC* genes were sequenced to exclude infantile neuronal ceroid lipofuscinosis, brain–lung–thyroid syndrome and gene defects related to surfactant proteins; no pathogenic variants were found in these genes. The disease course was progressive, leading to the early deaths of all the patients.

Histopathological studies of tissue samples (Electronic Supplementary Material)

Tissue biopsies (lung and skeletal muscle biopsies) and autopsy samples were obtained from all three affected patients. The tissue was fixed in buffered 4% formaldehyde, routinely processed into paraffin blocks and cut into 5.0 μm sections. Hematoxylin-eosin

and luxol fast blue stainings were performed as previously described [8]

Transmission electron microscopy on patient tissue samples and zebrafish *nhlrc2* morphants

Autopsy samples from the lungs and liver from Patient 2 and whole zebrafish embryos at 2 dpf injected with random sequence (5.1 ng) or splice site-blocking morpholino (5.0 ng) were fixed in 1% glutaraldehyde and 4% formaldehyde in 0.1 M phosphate buffer and were then post-fixed with 1% osmium tetroxide (Electron Microscopy Sciences, Hatfield, PA, USA), dehydrated in acetone and embedded in Epon LX112 (#21210; Ladd Research Industries Inc., Williston, VT, USA). Thin sections were cut with a Leica Ultracut UCT microtome (FC6; Leica, Wetzlar, Germany), stained with uranyl acetate (Structure Probe Inc., West Chester, PA, USA) and lead citrate (Laurylab, Brindas, France) and examined with a Tecnai Spirit transmission electron microscope (FEI Company, Eindhoven, Netherlands). Images were captured with a Quemesa CCD camera (Olympus Soft Imaging Solutions, Münster, Germany). Specimens for electron microscopy were processed and analysed in the EM core facility at Biocenter Oulu (Oulu, Finland).

Molecular genetic analyses and genealogy (Electronic Supplementary Material)

Whole exome sequencing (WES) [26] and bioinformatics analysis [31] were performed on samples from both families. Using Sanger sequencing, we confirmed variants and segregation within the families. To investigate the heredity of the mutations, we traced the ancestors of the patients in the Finnish Population Registries and National Archives of Finland [30].

Screening of *NHLRC2* variants in population cohorts

To determine whether the two *NHLRC2* variants were present in the general Finnish population, we used two strategies. First, we examined these variants in a Finnish population of infants born at term (gestational age ≥37 wk, total *n*=306) sampled prospectively at Oulu University Hospital from 2004 to 2007 (*n* = 199) [11] and in 2014 (*n* =107). The individuals born in this hospital mostly originate from northern Finland; therefore, these infants represent a valid population control. Next, we searched the Sequencing Initiative Suomi (SISu) [7] and ExAC data [14] (http://exac.broadinstitute.org/ ;01,2017) for the variants. The SISu data currently covers exonic variants of 10,490 individuals of Finnish origin.

***NHLRC2* expression and function (Electronic Supplementary Material)**

After reverse-transcription PCR (RT-PCR), we used Sanger sequencing to investigate the expression of the mutant alleles. Protein expression in the whole-cell extracts of fibroblasts from all three patients and healthy control subjects, together with the homogenates of control autopsy samples from several types of human tissues, were analysed by immunoblotting. Two constructs containing either full-length human *NHLRC2* or the thioredoxin-like domain alone were expressed in the *E. coli* strain BL21 (DE3). An insulin turbidity assay to test thioredoxin activity was performed, as previously described [10].

Nhlrc2 knockout mice (Electronic Supplementary Material)

Heterozygous C57BL/6N-A^{tm1Brd} Nhlrc2$^{tm1a(KOMP)Wtsi}$/Wtsi mice were obtained from the Infrafrontier-EMMA repository (strain number EM:10219). These mice carry a mutated allele where the function of the *Nhlrc2* gene is inactivated by the insertion of a targeting cassette that contains the expression reporter beta galactosidase gene [27]. Heterozygous mice were cross-bred, pups were genotyped or pregnant females dissected and embryos analysed. The expression of *Nhlrc2* in embryos was studied by staining for the expression of the lacZ reporter. On embryonic day 2.5 (E2.5), morulae were flushed from the uterus in accordance with a previously described protocol [19]. Embryos that were grown overnight in a microculture drop were transferred onto a gelatinised 24-well plate, each into its own well with 300 µl of embryonic stem cell medium, and left to grow for 10–12 days, after which DNA was extracted, as described [23].

Generation of zebrafish morphants (Electronic Supplementary Material)

Two morpholinos against zebrafish *nhlrc2* (ENSDARG00000089581) were designed; the first morpholino was against the ATG-site to prevent translation and the second against the exon-intron boundary in the 3' end of exon 4 (ENSDARE00000919598) to prevent splicing, which would both lead to a frameshift and premature termination of translation. The effect of the splice-site blocking morpholino for *nhlrc2* levels was quantified using PCR. The knockdown efficiency was determined from an agarose gel using ImageJ v1.49 by calculating the ratio of the wild-type (WT, the unmodified band) to the sum of all bands in a single lane (WT and two morpholino modified bands corresponding to exon exclusion and intron inclusion caused by splice site blocking).

Quantification of the affected cells in the midbrain region of zebrafish morphants from transmission electron micrographs

From five zebrafish injected with random sequence control morpholino (5.1 ng) and five zebrafish injected with *nhlrc2* splice site-blocking morpholino (5.0 ng), five random electron micrographs were taken from sagittal thin sections of the midbrain area. All the cells with visible nuclei and that were fully within the imaged area were counted and analysed for their integrity using iTEM software (Olympus Software Imaging Solutions GMBH). Cells with excess vacuolisation or disintegration were classified as affected.

Statistics

Differences were considered statistically significant (*) if the *p* value was 0.01– 0.05, very significant (**) if the *p* value was 0.001–0.01, and highly significant (***) if the *p* value was <0.001. A paired two-tailed student t-test was used to analyse the statistical significance. Zebrafish dechorionation timing differences were calculated with the log rank Mantel-Cox method, and values below 0.05 were considered significant.

Results

Clinical findings

Family A. Two brothers were born after normal consecutive singleton pregnancies to healthy parents in this non-consanguineous Finnish family. Both siblings presented at 2 months of age with feeding problems, muscular hypotonia, shoulder hypertonia and irritability. During the following months, both had delayed development, recurrent

infections, growth failure, respiratory problems and malabsorption. They subsequently developed dystonic tetraplegia, poor visual contact, progressive respiratory difficulties, transient liver dysfunction, and chronic haemolytic anaemia. The elder brother additionally presented with epileptic seizures and kidney dysfunction related to metabolic crisis at 9 months. The disease course was progressive and, despite intensive care, both patients died (at 1 year 9 months and 1 year 1 month, respectively) of progressive respiratory failure.

Family B. In this non-consanguineous Finnish family, the affected son was the first infant born to healthy parents. The mother had gestational hepatosis and toxaemia. At 2 months, the infant presented with irritability, floppiness, haemolytic anaemia, feeding problems, frequent diarrhoea and poor weight gain. At 4 months, he was hospitalised because of a suspected seizure, poor general condition and respiratory problems. During the following months, he had delayed development, dystonia, failure to thrive, strabismus, recurrent infections, and respiratory difficulties. The disease course was progressive, and despite intensive care, the patient passed away at 1 year 2 months due to respiratory failure.

All three patients appeared healthy at birth, but manifested at 2 months with unique multi-organ symptoms with a progressive disease course that resembled no known disease phenotype. Table 1 summarises the major clinical findings, including progressive cerebropulmonary symptoms, transient liver dysfunction, progressive growth failure and chronic haemolytic anaemia. Radiological investigations revealed over-inflated lungs with perihilar interstitial opacities, diffuse infiltrations and atelectasis (Online Resource Fig. 7) and thin corpus callosum (Online Resource Fig. 8). Additional findings included increased echogenicity of the kidneys (Patient 1) and hepatomegaly during metabolic crisis (Patients 1 and 2), enlarged thymus (Patient 2) (Online Resource Fig. 9), dilated lateral ventricles,

frontal sulci resembling cerebral atrophy (Patients 2 and 3) and increased signal intensity of the globus pallidum (Patient 3) (Online Resource Fig. 8). Visual evoked potentials (VEPs) revealed increased latencies and giant responses, indicating dysfunction of the visual tract (Patient 1). Electroencephalographs (EEGs) showed slow background activity and frontal bilateral rhythmic high-amplitude sharp delta transients, indicating metabolic encephalopathy (Patient 1). The EEG of Patient 3 was normal at 5 months, but at 11 months it was monotonic during sleep and consisted of 4-Hz delta waves without normal sleep spindles and vertex waves. Detailed descriptions of the clinical manifestations are presented in the Electronic Supplementary Material.

Histopathological findings in the brain

Post mortem neuropathological examinations were performed for all three patients. The findings in Patient 1 included brain atrophy (942 g, reference 1050 g); thin corpus callosum; small hippocampal regions and increased angiomatosis-like leptomeningeal, cortical and superficial white matter venous and capillary vascularisation and congestion (Fig. 1a). Superficial vascular changes were observed macroscopically. Widespread cerebral cortical oedema, focal neuronal loss and gliosis were observed (Fig. 1b and c) Symmetrical vacuolar degeneration and myelin loss of the fibres at the level of the crus cerebri in the midbrain were detected, as well as vacuolar degeneration of the white matter of the cerebellum, cerebellar oedema, focal Purkinje cell depletion and gliosis (Fig. 1d). Neuronal depletion of the anterior horns of the spinal cord was also found.

Patient 2 also had brain atrophy (764 g, reference 900 g), a thin corpus callosum and prominent congested leptomeningeal and superficial brain parenchymal vasculature.

Furthermore, the following neurodegenerative findings were detected: vacuolar degeneration and myelin loss of the corpus callosum and central white matter; vacuolar degeneration of the optic tract and the internal capsule; neuronal depletion of the corpora mammillaria, hypothalamic area and amygdaloid nucleus, as well as subthalamic and thalamic nuclei; vacuolar degeneration and myelin loss of the fibres of the crus cerebri and superior cerebellar peduncle and reticular formation in the midbrain, tegmentum and medial lemniscus in the pons (Fig. 1e-h). Other findings included mild vacuolar degeneration, myelin loss of the fibres of the pyramids and reticular formation in the medulla and in the white matter of the cerebellum. In addition, there was patchy neuronal loss in the anterior horns of the spinal cord.

Examinations of samples from Patient 3 revealed brain atrophy (742 g, reference 860 g), but a mostly normal corpus callosum, cerebral cortex and hippocampus. However, vascular congestion and angiectasia were detected. The temporal white matter exhibited vacuolar degeneration. Furthermore, the following findings were detected: vacuolar degeneration and neuronal depletion in the hypothalamic and amygdaloid nuclei and the putamen; vacuolar degeneration and myelin loss in the spinothalamic tract, medial lemniscus, superior cerebellar peduncle and reticular formation; degeneration of the anterior pontine nuclei; and minimal vacuolar degeneration in the white matter and dentate nucleus of the cerebellum.

Histopathological findings in the lungs

The lung biopsy samples from Patients 2 and 3 revealed no normal lung tissue; the lung tissue had been replaced by interstitial fibrosis resembling what is normally associated with

non-specific interstitial pneumonia (NSIP) (Fig. 2a and b). The alveolar epithelial cells appeared hyperplastic, and there were focal loose fibrotic areas positive after Alcian blue/periodic acid–Schiff (AB-PAS) staining. No hyaline membranes or diffuse alveolar damage (DAD) were visible.

Autopsy samples of the lungs were available from all three patients, and all the samples showed varying amounts of interstitial changes and fibrosis (Fig. 2c-h). Interstitial fibrosis in the lungs of Patient 1 was milder than in the lungs of his younger brother (Patient 2) (Fig. 2c). The lungs of Patient 2 revealed advanced interstitial fibrosis as a prominent feature, which was consistent with the findings in the lung biopsy samples; however, the degree of fibrosis in the autopsied lungs was more severe. In the lungs of Patient 3, the fibrosis was mainly moderate or mild. The general histological appearance of the autopsied lungs resembled that of NSIP. Intra-alveolar hyaline, hyaline membranes and squamous epithelial metaplasia reflecting DAD were intermittently observed in the lungs of Patients 1 and 3 but not in Patient 2 (Fig. 2g). Novel histopathological structures that we have termed granuloma-like lesions were observed in all cases, but in varying numbers; multiple lesions in five lung tissue sections in Patient 1, hundreds of lesions in five lung tissue sections in Patient 2 and fewer than 10 lesions in eight lung tissue sections of Patient 3 (Fig. 2c, e and h). Hyaline-resembling fibrotic areas were observed around the arteries in all cases, but the walls of the vessels were otherwise normal (Fig. 2d and e). Some airways were dilated, but no true honeycombing was found.

Most of the granuloma-like lesions were surrounded by hyaline-resembling extracellular material that contained large spindle-shaped cells. These large spindle-shaped cells were positive for alpha-smooth muscle actin (α-SMA) (myofibroblast marker) and

negative for cytokeratin (epithelial cell marker), phosphoglucomutase 1 (PGM1) (macrophage marker), smooth muscle myosin heavy chain (SMMHC) (smooth muscle cell marker) and caldesmon (smooth muscle cell marker) (Fig. 2f). The immunohistochemical phenotype indicated that the cells in the granuloma-like lesions were myofibroblasts.

Histopathological findings in the other organs

All three patients had hepatomegaly (435–485 g, reference weight 330–370 g). An autopsy liver sample from Patient 1 showed widespread microvesicular steatosis and foci of hepatocellular necrosis. Patient 1 also had cardiomegaly (90 g, reference weight 56 g) with slightly increased connective tissue. Autopsy revealed mild muscular atrophy in Patients 1 and 2; previous muscle biopsy samples from these patients revealed no histopathological changes. Furthermore, upon histological examination, the autopsy revealed atrophic thymus and adrenal glands without any specific abnormalities.

Electron microscopic findings in the lungs and liver

We analysed lung and liver autopsy samples from Patient 2 using a transmission EM (TEM). In the lungs, the most abundant cell types were myofibroblasts and smooth muscle and alveolar epithelial cells (Fig. 3a-c). Alveolar epithelial cells were mostly type II cells with intracellular lamellar bodies; only a few type 1 alveolar cells were detected (Fig. 3a and b). In the liver tissue samples, we observed steatosis (Fig. 3d) and collagen accumulation both inside- and outside of the hepatocytes, and a basement membrane was evident in the space of Disse (Fig. 3e and f).

Variants of *NHLRC2* in the two families

WES revealed compound heterozygous variants in *NHLRC2*, which encodes for the poorly characterised NHL repeat containing protein 2 in all three of the affected children (Fig. 4a). The *NHLRC2* gene is located in chromosome 10: 113,854,661-113,917,194 and the protein contains an N-terminal thioredoxin-like domain together with six NHL amino acid sequence repeats (named after three original genes *NCL-1, HT2A* and *LIN-41*) (Fig. 4b) [24]. Previously, mutations in *NHLRC2* have not been associated with any other human disease. The FINCA patients carried the heterozygous variant NM_198514:c.442G>T, which led to amino acid substitution p.Asp148Tyr in exon 3 within the highly conserved thioredoxin-like domain (Fig. 4b and d, Online resource Fig. 10). However, in the classical insulin-reduction activity assay for thioredoxins, neither full-length NHLRC2 nor the isolated thioredoxin-like domain produced in *E. coli* showed significant thioredoxin activity (Fig. 4c).

In addition, all three patients had a heterozygous frameshift 2-bp deletion NM_198514:c.601_602delAG, which caused p.Arg201GlyfsTer6 in exon 3 (Fig. 4a and b, Online Resource Fig. 10). Both the variants were segregated from their healthy parents. Genealogies of the patients, traced back seven to nine generations to the 1750s, revealed that the two families did not have common ancestors. We were unable to confirm the consanguinity of these two core families; their ancestors were born in northern Finland at least 150 miles apart.

Allele frequency of NHLRC2 variants among populations

We did not detect p.Asp148Tyr or p.Arg201GlyfsTer6 variants in the northern Finnish cohort of infants (*n*=306). In the Finnish population represented by Sequencing Initiative Suomi (SISu) data (*n*=10,488), p.Asp148Tyr (rs201701259) was rare, with a frequency of 0.003. In SISu, a frequency of 0.0001 was reported for p.Arg201GlyfsTer6 (rs757267294); however, the variant had not passed through quality control. In the Exome Aggregation Consortium (ExAc) data, the p.Asp148Tyr variant was present in non-Finnish European (*n*=33,351), South Asian (*n*=8,252), and African (*n*=5,202) populations with frequencies of 0.0003, 0.0002 and 0.0001, respectively. The p.Arg201GlyfsTer6 variant had a frequency of 0.0001 in non-Finnish Europeans (*n*=33,239) and was absent from other populations. The two variants were not detected as homozygous occurrences in any of the populations.

According to the SISu data, at least one of the two NHLRC2 variants, p.Asp148Tyr (rs201701259), was enriched in the Finnish population (n=10,488; natural logarithm of odds ratio 1.97), indicating that occurrence of the two variants in a compound heterozygous state in the two core families may be, at least partly, explained by their enrichment in the Finnish population.

Expression of Human *NHLRC2*

We investigated mRNA expression of variants NM_198514:c. 442G>T and c.601_602delAG in the patients' fibroblasts. *NHLRC2* mRNA expressed only the c.442G>T missense variant and not c.601_602delAG, confirming the origin of the variants from separate alleles and the degradation of mRNA with the frameshift 2-bp deletion

(Online Resource Fig. 10). The relative quantity of *NHLRC2* expression was studied by qPCR and showed approximately a two-fold decrease in patient-derived, immortalised fibroblasts compared to the controls (Fig. 4e).

The protein expression of NHLRC2 with the p.Asp148Tyr variant was investigated in the immortalised fibroblasts derived from patients and healthy controls. Immunoblotting revealed approximately a ten-fold decrease in the levels of NHLRC2 in whole-cell extracts from patients compared to control fibroblasts (Fig. 4f and g), further indicating the non-neutral role of the NHLRC2 variants.

Immunoblotting was also performed to investigate the expression of NHLRC2 in tissue homogenates from control autopsy samples (kidney, heart, muscle, liver, lung and brain tissue). NHLRC2 was detected in all human tissues included in the study (Online Resource Fig. 11). According to three public RNA expression databases (mouse) Brain RNA-Seq (https://web.stanford.edu/group/ barres_lab/brain_rnaseq.html) [32], Human Brain Transcriptome (http://hbatlas.org/) and the Atlas of the Developing Human Brain (www.brainspan.org)[13] the expression of mouse *Nhlrc2* is highest in astrocytes, oligodendrocyte progenitor cells, newly formed oligodendrocytes and endothelial cells. Mouse *Nhlrc2* is also expressed in neurons, myelinating oligodendrocytes and microglia. According to the Atlas of the Developing Human Brain (www.brainspan.org)[17], expression of human *NHLRC2* was highest in foetal brains during early pregnancy. The average expression levels of human *NHLRC2* are similar among the cerebellar cortex, mediodorsal nucleus of the thalamus, striatum, amygdala, hippocampus and 11 areas of the neocortex (Online Resource Fig. 12).

Nhlrc2^{tm1a} knockout mice

Nhlrc2 knockout (KO) mice (21 litters, *n*=154) were genotyped from heterozygous breeding pairs (Table 2). The ratio of born mice between WT and heterozygous mice was 1:2.125, which is in line with Mendelian segregation. The heterozygous mice, both female and male, were vital and bred well, indicating normal function of Nhlrc2^{tm1a} KO gametes. We also genotyped 53 embryos from six heterozygous breeding pairs on E10.5 and isolated and genotyped 45 morulae (E2.5) from six heterozygous breeding pairs (Online Resource Fig. 13). We did not detect any homozygous *Nhlrc2* KO mice, embryos or morulae, indicating that the lethality of Nhlrc2^{tm1a} KO homozygotes occurs between the fertilisation and morula (E2.5) stage. *Nhlrc2* expression analysis with LacZ-reporter mice indicated widespread expression throughout the body while WT littermates remained negative (Fig. 5). RNA *in situ* data of the Eurexpress atlas (www.eurexpress.org/ee/) showed strongest *Nhlrc2* expression in the brain, the central nervous system, the alimentary system and in the cardiovascular system of mouse embryos (E14.5) [4].

Zebrafish knockdown and dechorionation

We then decided to study the role of nhlrc2 in early embryogenesis using zebrafish, which is a convenient model organism for developmental studies because of the transparency of the embryos. The morphogenesis of the primary organ systems of a zebrafish is complete in 48 hours and the larvae display food seeking and avoidance behaviour 72 hours after conception [12]. First, we assessed the efficacy of *nhlrc2* splice site-blocking morpholino with PCR. As predicted, morphants had altered mRNA splicing compared to the untreated controls. The effect gradually faded toward 4 dpf (Online Resource Fig. 14). Besides minor

swelling in the abdomen, *nhlrc2* morphants developed normally (Online Resource Fig. 14) without any apparent developmental defects. Both splice site-blocking morpholino (3.6 ng) and translation-blocking morpholino (2.3 ng) were used, with both co-injected with an equal amount of p53 morpholino [21]. A slight delay was observed in the timing of the dechorionation (Online Resource Fig. 15a), but a similar effect was also observed in the random controls (Online Resource Fig. 15b), suggesting both were non-specific effects.

Because there were no apparent changes in the gross morphology of developing *nhlrc2* morphants, we next investigated the central nervous system in more detail using TEM. Because we were not able to ascertain the effect of the translation-blocking morpholinos without a validated antibody, we chose to use the splice site-blocking morpholinos for the TEM analysis. In this experiment, we achieved approximately 95.2% knock-down efficiency at the time point of 2 dpf by using the splice site-blocking *nhlrc2* morpholino (5.0 ng) (Online Resource Fig. 16). Using TEM, the quantitative analysis of the midbrain region from the 2 dpf morphants showed statistically significant increments in the number of affected cells in *nhlrc2* knock-downs compared to the controls injected with a random sequence morpholino (Fig. 6). The data indicate that *nhlrc2* has a key role in maintaining the cellular integrity of the central nervous system in developing zebrafish embryos.

Discussion

Based on clinical and molecular genetic data we have identified a novel cerebropulmonary and multi-organ disease, which we have named FINCA. This disease is characterised by a unique combination of tissue fibrosis, neurodegeneration and cerebral angiomatosis. Using

WES we identified identical compound heterozygous *NHLRC2* variants in three affected children from two unrelated families. In addition, we used a KO mouse model and zebrafish morphants to investigate the consequences of absent or decreased *Nhlrc2* expression; finally, we performed biochemical studies on human recombinant NHLRC2.

NHLRC2 is conserved in eukaryotes, having 84% sequence identity between human and mouse orthologues. The function of NHLRC2 is currently unknown. There are two NHLRC2 transcript variants in humans, both containing six NHL repeats that may form part of a TolB-like six-bladed beta-propeller. Six-bladed beta-propellers occur in a wide range of proteins, including 165 human proteins listed in InterPro [18]. They are often involved in protein–protein interactions. The longer isoform of NHLRC2 also has an N-terminal thioredoxin-like domain with an unusual motif (CCINC), where the active site of thioredoxin (WCGPC for human thioredoxins 1 and 2) is usually located. Interestingly, the missense variant identified from the patients in this study located in the thioredoxin-like domain and led, together with the frameshift variant, to significantly reduced expression of *NHLRC2* at the mRNA and protein level in patient-derived fibroblasts. A classical insulin-reduction activity assay for thioredoxins revealed that neither full-length NHLRC2 nor the isolated thioredoxin-like domain showed significant thioredoxin activity. This may be because of the substrate specificity of NHLRC2, which could be mediated by protein–protein interactions involving the six-bladed beta-propeller. Thus, insulin may be excluded from the active site; however, the thioredoxin fold motif is versatile, with both redox [16] and non-redox functions. An example of a non-redox function occurs in SasA, a circadian clock-associated histidine kinase [29]. This versatility plus the unusual motif in place of the active site indicates that the N-terminal domain has a thioredoxin fold, but no

thioredoxin activity. Conservation of the active-site cysteines indicates that the protein either has or regulates a redox-related function.

Based on large population cohort datasets, the rare variants identified in our patients are not restricted to the Finnish population; they are also present as simple heterozygotes in other populations. Our genotyping analysis of *Nhlrc2* KO mice, foetuses and morulae revealed that homozygosity was lethal whereas heterozygous mice were viable and non-symptomatic. Interestingly, the *nhlrc2* morphant zebrafish embryos had a significant number of affected cells in the midbrain, indicating that *nhlrc2* has an important role in the maintenance of cell survival in the developing brain. In contrast to the postnatal onset of the symptoms in humans, the timing of the effects of the zebrafish knockdown suggests the morphant zebrafish model does not fully reflect the features of the human disease. We are currently working on generating a knockout zebrafish model to better observe the potential symptoms at later stages of development.

Autopsy studies revealed that in FINCA patients, there was increased angiomatosis-like leptomeningeal, cortical and superficial white matter vascularisation and congestion, in addition to white matter degeneration and variable neuronal degeneration. According to currently available RNA expression data in Human Brain Transcriptome (http://hbatlas.org/), the average expression level of human *NHLRC2* is similar among different regions of the human brain. The expression of mouse *Nhlrc2* is the highest in astrocytes, oligodendrocyte progenitor cells, newly formed oligodendrocytes and in the endothelial cells [32]. Altogether, this expression data indicate that *NHLRC2* is widely expressed in the different regions and cell types of the brain. However, based on the current knowledge on the function of NHLRC2 it is difficult to conclude the primary mechanism

behind FINCA disease and to exclude the role of secondary processes such as ischemia as a cause of certain neuropathological findings described in our patients such as depletion of Purkinje cells and pyramidal neurons in Ammon's horn. Lung biopsy and autopsy studies showed severe interstitial fibrosis and previously undescribed granuloma-like lesions. The granuloma-like lesions were enriched with myofibroblasts that accumulate in areas surrounding granulomas in several granulomatous lung diseases such as sarcoidosis, atypical mycobacteriosis and tuberculosis [9]. The predominance of type II alveolar cells and paucity of type I cells were evident. Lung autopsy specimens additionally revealed DAD. Furthermore, all the patients had hepatomegaly. Liver samples for TEM were available only from one patient (Patient 2) revealing steatosis and accumulation of collagen bundles.

From the previous literature, we have found only one other "FINCA-like" cerebropulmonary disease called brain–lung–thyroid syndrome (BLT syndrome, OMIM 118700) or choreoathetosis and congenital hypothyroidism with or without pulmonary dysfunction (CAHTP, MIM #610978) that begins with muscular hypotonia followed by choreoathetosis, dystonia, ataxia, and dysarthria. This is an autosomal dominant infantile-onset disorder caused by mutations in thyroid transcription factor 1 (NKX2-1/TITF1; 14q13.3) [3,28]. The phenotype is less severe than in FINCA and varies between and within families; some patients show neonatal respiratory distress syndrome (RDS), developmental delay, symptoms of possible hypothalamic dysfunction or congenital cardiac septal defects [28]. Neuroimaging in some NKX2-1 related cases has shown structural brain abnormalities including agenesis of the corpus callosum [20], and autopsies have revealed reduced numbers of striatal and neocortical interneurons consistent with a defect in

neuronal migration [13]. In addition to RDS in neonates, pulmonary disease may present with interstitial lung disease in young children, and pulmonary fibrosis in older persons [20]. Histologic studies on patients with NKX2-1 related pulmonary dysfunction have identified interstitial widening and pneumocyte hyperplasia, desquamative interstitial pneumonia, accumulation of foamy alveolar macrophages, and pulmonary alveolar proteinosis [6]. The risk for pulmonary carcinoma is increased in young adults with an NKX2-1 related disorder. In BLT syndrome, thyroid involvement may present as subclinical hypothyroidism with mildly elevated thyroid stimulating hormone (TSH) or hyperthyrotropinemia. Additional symptoms, including failure to thrive, malabsorption, and intellectual deficit, have been reported in patients with deletions on chromosome 14, including NKX2-1 (www.orpha.net). In FINCA disease, in addition to cerebropulmonary symptoms, we also observed failure to thrive, malabsorption, early-onset chronic haemolytic anaemia, and recurrent infections in all three patients. Furthermore, transient liver dysfunction (Patients 1 and 2), transient kidney dysfunction (Patient 1), subclinical hypothyroidism (Patient 3, Electronic Supplementary Material), a transient increase in serum tumour markers and a non-specific decrease in the oxidative activity of monocytes (Patient 2, Electronic Supplementary Material) were noted.

Our studies indicate that NHLRC2 has a significant role in the CNS. However, further research is required to define the complex role of NHLRC2 in the pathogenesis of FINCA disease and, more commonly, its role in tissue fibrosis, neurodegeneration and cerebral angiomatosis. The presentation of FINCA disease and potentially other diseases may vary in response to the genetic defect in *NHLRC2*, environmental factors and other genetic factors. Recently, NHLRC2 has been listed as one of six novel blood-based

biomarkers for Alzheimer´s disease, indicating its yet undefined role in neurodegeneration [15]. Our expression analysis with LacZ-reporter mice indicates widespread expression of *Nhlrc2* throughout the mouse embryo. This, in connection with the severe multi-organ phenotype of FINCA patients, indicates that NHLRC2 has a vital role in early embryogenesis and in maintenance of multi-organ homeostasis after birth. We propose that NHLRC2 variants should be considered in patients with phenotypes that present as a combination of neurological and respiratory symptoms or additional multi-organ manifestations.

Acknowledgements

The authors would like to thank Professors Eric Shoubridge, Kalervo Hiltunen and Christer Betsholtz, Assistant Professor Michael Vanlandewijck, Adjunct Professor Siri Lehtonen and Dr. Riikka Pietilä for their expert advice and support and also Ms Pirjo Keränen, Ms Riitta Vuento, Ms Maarit Haarala, Ms Hanna Seppälä, Ms Kirsi Säkkinen, the Transgenic Core Facility at Biocenter Oulu, and the Laboratory Animal Centre at the University of Oulu for their expert assistance. Biocenter Oulu Electron Microscopy core facility, a member of Biocenter Finland, is acknowledged for their help with EM analysis. The zebrafish work was carried out at University of Tampere core facility, supported by Biocenter Finland. The digital pathology scanner of Northern Finland Biobank Borealis was used in imaging the neuropathological findings. This work was conducted with support from the Research Council for Health of the Academy of Finland (J.U., decision number 138566; R.H., decision numbers 266498, 273790 and 303996; M.H., decision number 1126662; L.R., decision numbers 266457 and 272573); the Sigrid Juselius Foundation

(J.U., R.H. and M.H.); the Foundation for Paediatric Research, Finland (J.U. and M.K.K.); the Alma and KA Snellman Foundation (J.U. and M.K.K.); a Marie Curie International Outgoing Fellowship of the European Union's Seventh Framework Programme (grant agreement number 273669 [BioMit]) (R.H.); Foundation of the Finnish Anti-Tuberculosis Association (R.K.); the Jane and Aatos Erkko Foundation (M.R.); the Competitive State Research Financing of the Expert Responsibility Area of Tampere University Hospital (MR); Special State Grants for Health Research in the Department of Paediatrics and Adolescence at Oulu University Hospital, Finland (J.U.); the National Heart, Lung and Blood Institute of the U.S. National Institutes of Health under award number HL-54703 (L.M.N.) and the Eudowood Foundation (L.M.N.).

Compliance with ethical standards

Study and ethical approval

All procedures performed in the studies involving human participants were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Prior to the study, the guardians of the patients gave written informed consent to participate in the studies, and this was approved by the Ethics Committee of Oulu University Hospital (EETTMK 51/2008). Furthermore, the guardians of the patients in this manuscript have given written informed consent for the publication of their case details.

All procedures performed in the studies involving animals were in accordance with the ethical standards of the institution or practice at which the studies were conducted. The National Animal Experiment Board of Finland approved the study protocol

(ESAVI/5882/04.10.07/2014). Animal care and experimental procedures were conducted in accordance with the national legislation and EU Directive 2010/63/EU. Zebrafish housing and maintenance were done according to facility permission ESAVI/10079/04.10.06/2015. All applicable international, national and/or institutional guidelines for the care and use of animals were followed.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Barabasi AL, Gulbahce N, Loscalzo J (2011) Network medicine: a network-based approach to human disease. Nat Rev Genet 12:56-68. Doi:10.1038/nrg2918

2. Bossy-Wetzel E, Schwarzenbacher R, Lipton SA (2004) Molecular pathways to neurodegeneration. Nat Med 10 Suppl:S2-9. Doi:10.1038/nm1067

3. de Vries BB, Arts WF, Breedveld GJ, Hoogeboom JJ, Niermeijer MF, Heutink P (2000) Benign hereditary chorea of early onset maps to chromosome 14q. Am J Hum Genet 66:136-142. Doi:S0002-9297(07)62240-X

4. Diez-Roux G, Banfi S, Sultan M, Geffers L, Anand S, Rozado D, et al. (2011) A high-resolution anatomical atlas of the transcriptome in the mouse embryo. PLoS Biol 9:e1000582. Doi:10.1371/journal.pbio.1000582

5. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792-1797. Doi:10.1093/nar/gkh340

6. Hamvas A, Deterding RR, Wert SE, White FV, Dishop MK, Alfano DN, et al. (2013) Heterogeneous pulmonary phenotypes associated with mutations in the thyroid transcription factor gene NKX2-1. Chest 144:794-804. Doi:S0012-3692(13)60595-4

7. Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Finland (2016) Sequencing Initiative Suomi project (SISu).http://sisuproject.fi. Accessed 1/10 2017

8. J. D. Bancroft, A. Stevens (1991) Theory and practice of histological techniques. John Wiley & Sons, Churchill Livingstone, Edinburgh.

9. Kaarteenaho-Wiik R, Sademies O, Paakko P, Risteli J, Soini Y (2007) Extracellular matrix proteins and myofibroblasts in granulomas of sarcoidosis, atypical mycobacteriosis, and tuberculosis of the lung. Hum Pathol 38:147-153. Doi:S0046-8177(06)00417-5

10. Karala AR, Ruddock LW (2010) Bacitracin is not a specific inhibitor of protein disulfide isomerase. FEBS J 277:2454-2462. Doi:10.1111/j.1742-4658.2010.07660.x

11. Karjalainen MK, Huusko JM, Ulvila J, Sotkasiira J, Luukkonen A, Teramo K, et al. (2012) A potential novel spontaneous preterm birth gene, AR, identified by linkage and association analysis of X chromosomal markers. PLoS One 7:e51378. Doi:10.1371/journal.pone.0051378

12. Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF (1995) Stages of embryonic development of the zebrafish. Dev Dyn 203:253-310. Doi:10.1002/aja.1002030302

13. Kleiner-Fisman G, Rogaeva E, Halliday W, Houle S, Kawarai T, Sato C, et al. (2003) Benign hereditary chorea: clinical, genetic, and pathological findings. Ann Neurol 54:244-247. Doi:10.1002/ana.10637

14. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. (2016) Analysis of protein-coding genetic variation in 60,706 humans. Nature 536:285-291. Doi:10.1038/nature19057

15. Long J, Pan G, Ifeachor E, Belshaw R, Li X (2016) Discovery of Novel Biomarkers for Alzheimer's Disease from Blood. Dis Markers 2016:4250480. Doi:10.1155/2016/4250480

16. Martin JL (1995) Thioredoxin--a fold for all reasons. Structure 3:245-250. Doi:S0969-2126(01)00154-X

17. Miller JA, Ding SL, Sunkin SM, Smith KA, Ng L, Szafer A, et al. (2014) Transcriptional landscape of the prenatal human brain. Nature 508:199-206. Doi:10.1038/nature13185

18. Mitchell A, Chang HY, Daugherty L, Fraser M, Hunter S, Lopez R, et al. (2015) The InterPro protein families database: the classification resource after 15 years. Nucleic Acids Res 43:D213-21. Doi:10.1093/nar/gku1243

19. Nagy A, Gertsenstein M, Vintersten K, Behringer R (2003) Manipulating the Mouse Embryo, A Laboratory Manual, 3rd ed. p.198-200. Cold Spring Harbor Laboratory Press, New York

20. Patel NJ, Jankovic J (2014) NKX2-1-Related Disorders. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Mefford HC, et al. (eds), GeneReviews(R), University of Washington, Seattle. GeneReviews is a registered trademark of the University of Washington, Seattle. All rights reserved, Seattle (WA).

21. Robu ME, Larson JD, Nasevicius A, Beiraghi S, Brenner C, Farber SA, et al. (2007) P53 Activation by Knockdown Technologies. PLoS Genet 3:e78. Doi:06-PLGE-RA-0378R3

22. Santiago JA, Potashkin JA (2013) Integrative network analysis unveils convergent molecular pathways in Parkinson's disease and diabetes. PLoS One 8:e83940. Doi:10.1371/journal.pone.0083940

23. Scavizzi F, Ryder E, Newman S, Raspa M, Gleeson D, Wardle-Jones H, et al. (2015) Blastocyst genotyping for quality control of mouse mutant archives: an ethical and economical approach. Transgenic Res 24:921-927. Doi:10.1007/s11248-015-9897-1

24. Slack F, Ruvkun G (1998) Heterochronic genes in development and evolution. Biol Bull 195:375-376. Doi:10.2307/1543152

25. Spagnolo P, Grunewald J, du Bois RM (2014) Genetic determinants of pulmonary fibrosis: evolving concepts. Lancet Respir Med 2:416-428. Doi:10.1016/S2213-2600(14)70047-5

26. Sulonen AM, Ellonen P, Almusa H, Lepisto M, Eldfors S, Hannula S, et al. (2011) Comparison of solution-based exome capture methods for next generation sequencing. Genome Biol 12:R94-2011-12-9-r94. Doi:10.1186/gb-2011-12-9-r94

27. Testa G, Schaft J, van der Hoeven F, Glaser S, Anastassiadis K, Zhang Y, et al. (2004) A reliable lacZ expression reporter cassette for multipurpose, knockout-first alleles. Genesis 38:151-158. Doi:10.1002/gene.20012

28. Thorwarth A, Schnittert-Hubener S, Schrumpf P, Muller I, Jyrch S, Dame C, et al. (2014) Comprehensive genotyping and clinical characterisation reveal 27 novel NKX2-1 mutations and expand the phenotypic spectrum. J Med Genet 51:375-387. Doi:10.1136/jmedgenet-2013-102248

29. Vakonakis I, Klewer DA, Williams SB, Golden SS, LiWang AC (2004) Structure of the N-terminal domain of the circadian clock-associated histidine kinase SasA. J Mol Biol 342:9-17. Doi:10.1016/j.jmb.2004.07.010

30. Varilo T, Savukoski M, Norio R, Santavuori P, Peltonen L, Jarvela I (1996) The age of human mutation: genealogical and linkage disequilibrium analysis of the CLN5 mutation in the Finnish population. Am J Hum Genet 58:506-512

31. Witkowski L, Carrot-Zhang J, Albrecht S, Fahiminiya S, Hamel N, Tomiak E, et al. (2014) Germline and somatic SMARCA4 mutations characterize small cell carcinoma of the ovary, hypercalcemic type. Nat Genet 46:438-443. Doi:10.1038/ng.2931

32. Zhang Y, Chen K, Sloan SA, Bennett ML, Scholze AR, O'Keeffe S, et al. (2014) An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. J Neurosci 34:11929-11947. Doi:10.1523/JNEUROSCI.1860-14.2014

Figure captions

Fig. 1 Histopathological findings in brain autopsy samples from patients with NHLRC2 variants. **a** Autopsy findings from Patient 1 revealed increased angiomatosis-like leptomeningeal vascularisation. **b** Degeneration and depletion of the pyramidal cells of the hippocampus. **c** Higher magnification of the Ammon's horn showing gradual depletion of pyramidal neurons. **d** Purkinje cells of the cerebellum (between arrowheads). **e-h** Patient 2 had vacuolar degeneration and myelin depletion in the pons. Haematoxylin and eosin stain (**a-f**) and luxol fast blue stain (**g, h**). The images were taken with an Aperio AT2 digital pathology slide scanner (Leica Biosystems, Wetzlar, Germany). Scale bars: 200 µm

Fig. 2 Histological findings in lung samples from patients with NHLRC2 variants. **a, b** Interstitial fibrosis was present in the lung biopsy samples. **c, e, f, h** Granuloma-like lesions (arrows) within regions with variable amounts of necrosis (stars) were present in the autopsy samples from each patient. **d** Perivascular fibrosis was visible around the artery (arrowhead). **f** Spindle-shaped cells in a granuloma-like lesion were positive for alpha–smooth muscle actin (short arrows). **g** Hyaline membranes (short arrows) were present in the alveoli. **a, b** Lung samples collected by lung biopsy (**a, b**) and during autopsy (**c-h**). Samples from Patient 1 (**c**), Patient 2 (**a, d-f**) and from Patient 3 (**b, g, h**). Haematoxylin and eosin stain (**a-e, g**). Scale bars: 80 µm

Fig. 3 Electron microscopy analysis of lung (a-c) and liver (d-f) autopsy samples. **a** Lung alveoli contained type II pneumocytes (P) and the alveolar space was very limited (L, the

lumen of the alveolus). **b** In addition to type II pneumocytes (P), myofibroblasts (M) were present with diverging 'limb-like' projections. **c** Myofibroblasts with ultrastructural features such as abundant endoplasmic reticulum (E) and adherens junctions (A) had produced fibrotic extracellular collagen bundles (C). **d** The liver had many large fat (F) vacuoles in the hepatocytes. **e** Thin and fragmentary basement membrane (B) was detected in the space of Disse adjacent to the endothelial cells surrounding the sinusoidal lumen (L). **f** The accumulation of fat vacuoles (F) and collagen bundles (C) in hepatocytes. Sinusoidal lumen (L). Scale bars: 10 µm (**a, d**), 5 µm (**b**), 2 µm (**c, e, f**)

Fig. 4 Variants p.Asp148Tyr and p.Arg201GlyfsTer6 in NHLRC2 in patients with FINCA disease from two non-consanguineous families. **a** Pedigrees and the segregation of heterozygous variants p.Asp148Tyr and p.Arg201GlyfsTer6 in Families A and B. The probands of the families are indicated with an arrow and the healthy siblings of the patients are indicated with empty symbols. **b** The structural domains of NHLRC2 and location of the variants. **c** Neither full-length NHLRC2 (NHLRC2-FL) nor the isolated thioredoxin-like domain (NHLRC2-trx) produced in *E. coli* showed significant thioredoxin activity in the insulin reduction assay. **d** Conservation of the thioredoxin-like domain region containing p.Asp148Tyr (p.D148Y) in red [5]. **e** QPCR showed a highly significant decrease in the relative quantity of *NHLRC2* mRNA in patient-derived fibroblasts compared to controls according to a two-tailed student t-test. **f** Immunoblotting indicated a significant loss of the p.Asp148Tyr NHLRC2 protein in fibroblasts from Patients 1–3 compared to the levels of GAPDH. **g** Immunoblotting data presented as numerical values measured using Image J and statistical analysis by a two-tailed student t-test.

Abbreviations: D148Y, p.Asp148Tyr; R201GfsX6, p.Arg201GlyfsTer6; Trx, thioredoxin; C, control; P, patient; FL, full length. Differences were considered as statistically significant (*), with *p* value 0.01-0.05 and highly significant (***) with *p* -value < 0.001

Fig. 5 LacZ mouse tissue staining. Heterozygote Nhlrc2^{tm1a} mice had a strong β-galactosidase signal throughout the body. Heart and thymus (**a**); Lungs (**b**); Kidneys, adrenal glands and gonads (**c**); Liver (**d**) and Brain (**e, f**) (**e** WT; **f** heterozygote Nhlrc2^{tm1a} mice). WT littermates were negative for β-galactosidase signal

Fig. 6 Transmission electron microscopy image from the midbrain region of the 2-dpf-zebrafish morphant. **a** A cross section of a zebrafish embryo at 2 dpf with a circled area specifying the exact location from which the electron micrographs were taken. **b** Cells with excess vacuolisation or disintegration were classified as affected (marked with an asterisk in Fig. 6d). In the midbrain 5.8 % of the 810 cells analysed were affected by the *nhlrc2* splice site-blocking morpholinos, whereas only 1% of the 850 cells were affected in MO controls. **c** Midbrain of the random sequence morpholino injected control. **d** Midbrain of the *nhlrc2* splice site-blocking morphant. Abbreviations: SB, splice site-blocking; MO, morpholino. Differences were considered as statistically very significant (**) with *p* value 0.001-0.01

Electronic supplementary material

Fig. 7 Radiological findings of chest X-rays and high-resolution computed tomography of Patients 1–3. **a** Patient 1 at 7 months of age shows flattened hemidiaphragms and hyperlucent lungs demonstrating over-inflated lungs, small air-filled bullae of the right mediobasal lung (short arrows) and perihilar small atelectasis (long arrow). Performed during an acute metabolic crisis at the age of 9 months, the chest X-ray demonstrates perihilar and basal airspace opacities and axial HRCT confirms bilateral consolidations (long arrows) and reticular opacities (short arrows). **b** Patient 2 at 4 months of age demonstrates mediastinal and hilar prominence (short arrows), over-inflated lungs, and diffuse reticular interstitial opacities. At 4 months of age, HRCT reveals widespread bilateral ground-glass opacities (short arrow), interstitial infiltrations (long arrows) and atelectasis (black arrowhead), mildly enlarged thymus (white arrowhead) and left hilar adenopathy (asterisk). At 1 year of age, chest X-rays show the progression of interstitial infiltrations and bilateral dense consolidations (long arrows). **c** Patient 3 at 4 months of age reveals over-inflated lungs and bilateral perihilar interstitial opacities (short arrows). At 13 months of age, the chest X-ray shows diffuse infiltrations with reticulation. HRCT confirms diffuse bilateral ground-glass opacification, peripheral interstitial septal thickening and lobular pleural thickening (long arrows)

Fig. 8 Findings from brain magnetic resonance imaging of Patients 1–3. **a** Sagittal T1-weighted magnetic resonance imaging (MRI) shows the thin corpus callosum of Patient 1 at 7 months of age. **b** Patient 2 at 10 months of age. **c** Patient 3 at 10 months of age (short arrows). The axial T2-weighted and coronal T1-weighted MRIs of Patients 2 (**b**) and 3 (**c**)

demonstrate slightly dilated lateral ventricles and cortical sulci. **c** Increased signal intensity of the globus pallidus on the axial T2-weighted MRI of Patient 3 (arrowheads)

Fig. 9 Findings from abdominal ultrasound and magnetic resonance imaging of internal organs and thymus of Patients 1 and 2. **a** The abdominal ultrasound findings (Patient 1 at 10 months of age) show increased cortical echogenity with the strand-like hypoechogenic outerzone of both kidneys (short arrows). **b** The abdominal T2-weighted axial MRI of Patient 2 at 4 months of age shows hepatomegaly and a non-expansive hypointense lesion of the right liver lobe (long arrows). **c** The lesion is also hypointense on the T1-weighted coronal MRI (long arrows). **d** The axial fat saturation T1-weighted MRI of Patient 2 shows that the lesion (arrowheads) in the liver is less enhanced by the gadolinium contrast agent than the rest of the liver. **e** Normal echogenity of the right liver lobe of Patient 2 at the age of 4 months. **f** A slightly enlarged thymus of Patient 2, which was seen in the coronal T1-weighted MRI (asterisk)

Fig. 10 a Electropherograms showing both heterozygote NM_198514:c.442G>T and c.601_602delAG variants in genomic DNA from the patient. **b** The RNA extracted from the patients' fibroblasts has only c.442G>T. Thus, Sanger sequencing of mRNA confirms that the variants reside in separate haplotypes and have been inherited as compound heterozygotes in all patients. Furthermore, the absence of the c.601_602delAG variant indicates that the transcript with the frameshift deletion is processed through nonsense mediated RNA decay

Fig. 11 Immunoblotting to detect NHLRC2 protein expression in human tissue homogenates from control autopsy samples. NHLRC2 was detected in all studied human organs including the heart (He), kidney (Ki), muscle (Mu), liver (Li), lung (Lu) and brain (Br). GAPDH was used as a loading control

Fig. 12 *Nhlrc2* is expressed in various cell types of the mouse brain and *NHLRC2* expression is highest in the early human life. **a** An RNA-Seq transcriptome and splicing database of the glia, neurons and vascular cells of the cerebral cortex show the highest *Nhlrc2* expression in astrocytes, oligodendrocyte progenitor cells, newly formed oligodendrocytes and endothelial cells. *Nhlrc2* was also expressed in neurons, myelinating oligodendrocytes and microglia. (OPG) oligodendrocyte progenitor cells [32]. Abbreviations: FPKM, fragments per kilobase per million. **b** The average expression levels of *NHLRC2* are similar between the cerebellar cortex (CBC), mediodorsal nucleus of the thalamus (MD), striatum (STR), amygdala (AMY), hippocampus (HIP) and 11 areas of the neocortex (NCX) in the human brain (http://hbatlas.org/). **c** The expression is highest in the early development of the human brain. The number of samples is indicated on the top of the x-axis. Data from www.brainspan.org [17]. Abbreviations: RPKM, reads per kilobase per million; pcw, post conception week; mos, months; yrs, years

Fig. 13 Wild-type (WT) and heterozygous (het) *Nhlrc2* KO morulae and their growth *in vitro*. The morulae were isolated at embryonic day (E) 2.5 and co-cultured in a micro drop culture overnight. On the next day, the E3.5 compacted morulae and early blastocysts were transferred to separate gelatinised wells. At E4.5, the embryos had matured to the blastocyst

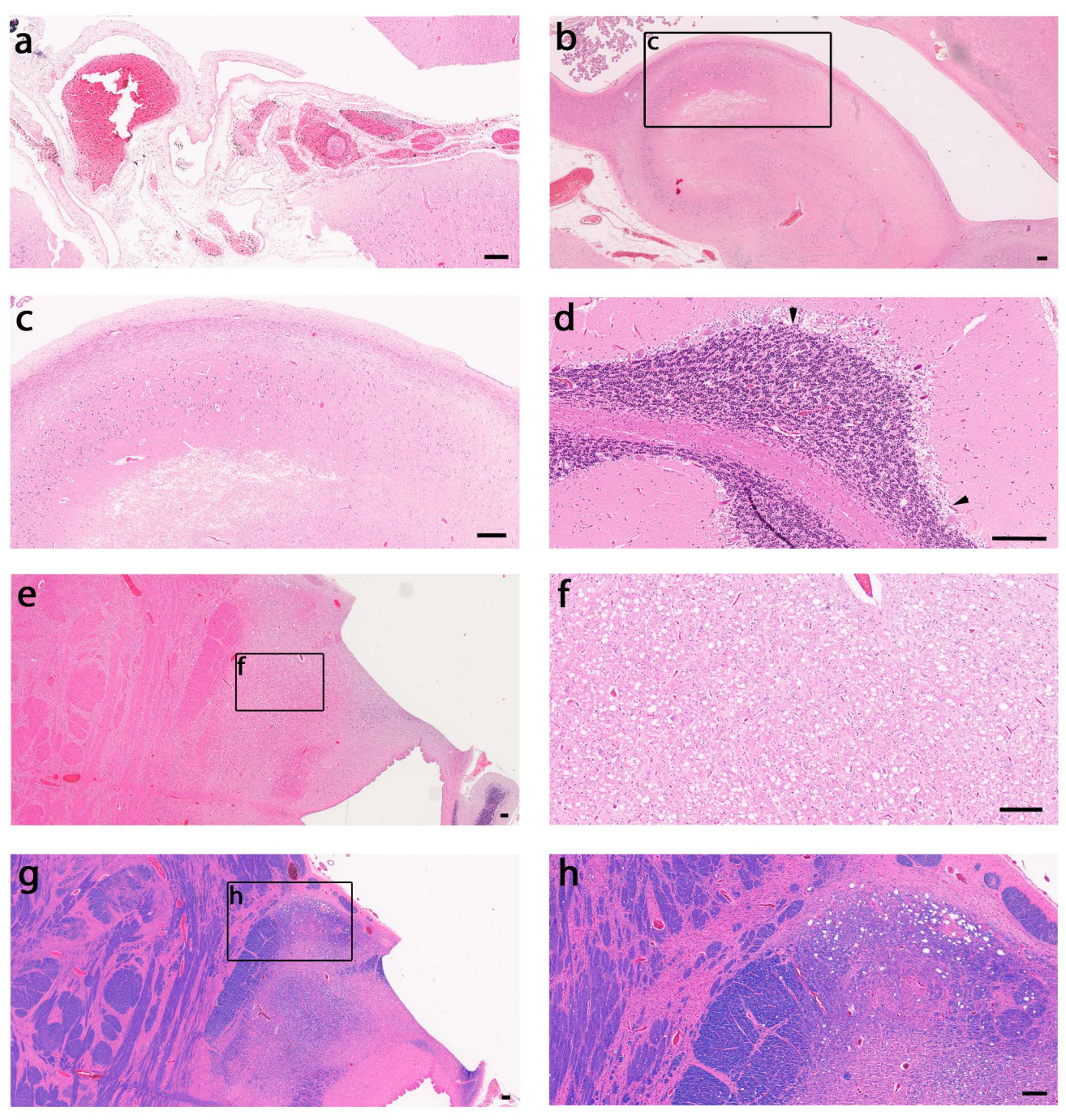
stage and at E5.5 they hatched and attached to the bottom of the plate. Cells were grown for several days to detect possible phenotypes and to obtain enough material for genotyping (E7.5, E10.5 and E13.5). No homozygous morulae were detected

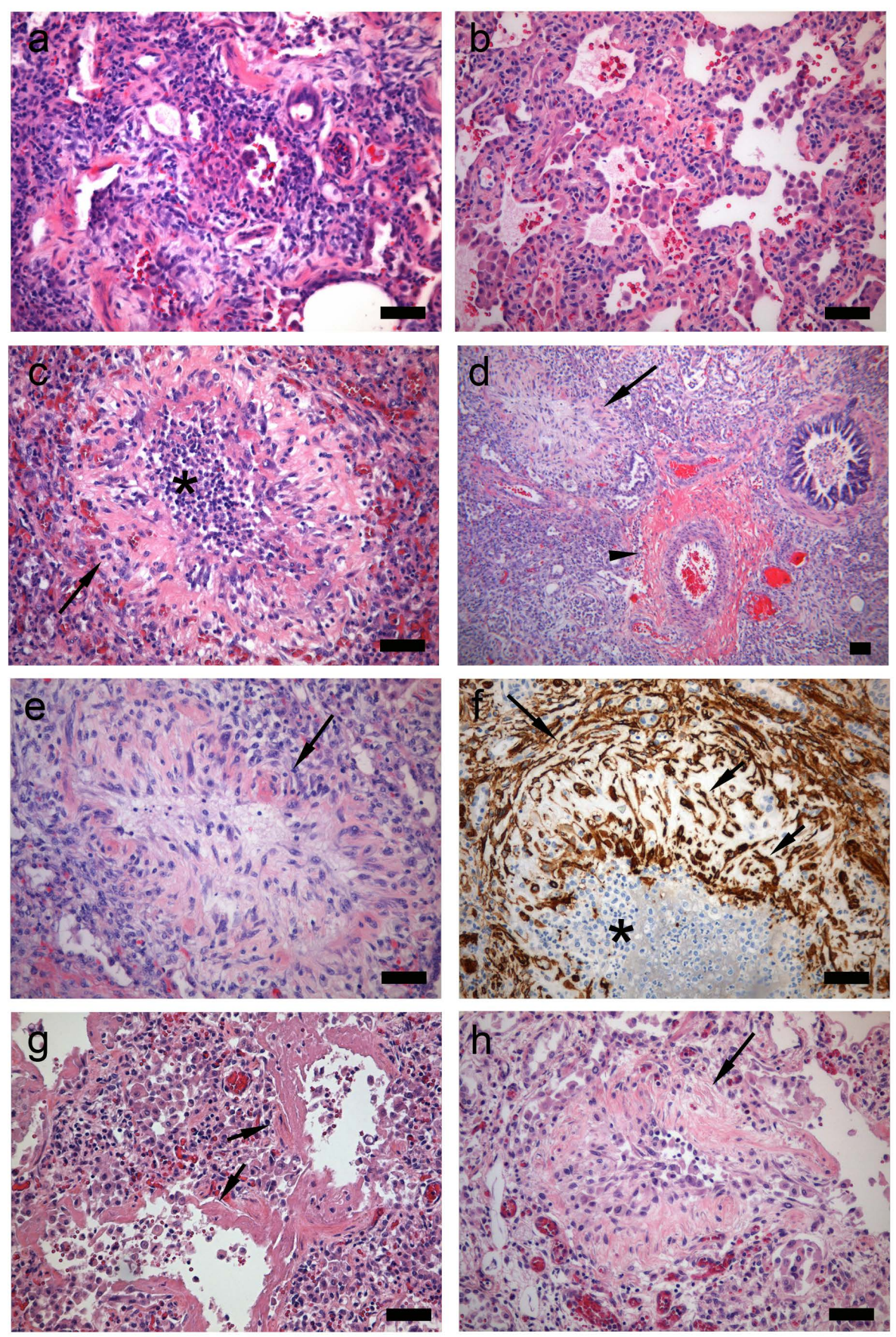
Fig. 14 *nhlrc2* splice site-blocking morpholino alters the mRNA spicing in zebrafish embryos but not the gross morphology of the embryos. Indicated amount of splice site targeting morpholino (together with an equal amount of p53 morpholino) was injected into the yolk sacs of fertilised eggs at the 1-2 cell stage. **a** Arrows on the left show the expected product sizes after the splice site-blocking events (334 bp intron inclusion, 198 bp WT product, 76 bp exon exclusion), and below the image are indicated the times of RNA extraction as days post fertilisation (dpf). **b** Un-injected controls on the left, *nhlrc2* splice site-blocking morpholino (3.6 ng) treated morphants in the middle and *nhlrc2* translation-blocking morpholino (2.3 ng) treated morphants on the right. The embryos appear to have WT morphology; only minor swelling was observed in the abdominal area in some individuals in both morphant groups at 3 dpf. However, this swelling diminished by 4 dpf. The images were taken with Axiovision software version 4.8 using a Zeiss SteREO Lumar V12 fluorescence microscope, equipped with 1.5X camera (Carl Zeiss MicroImaging GmbH, Göttingen, Germany)

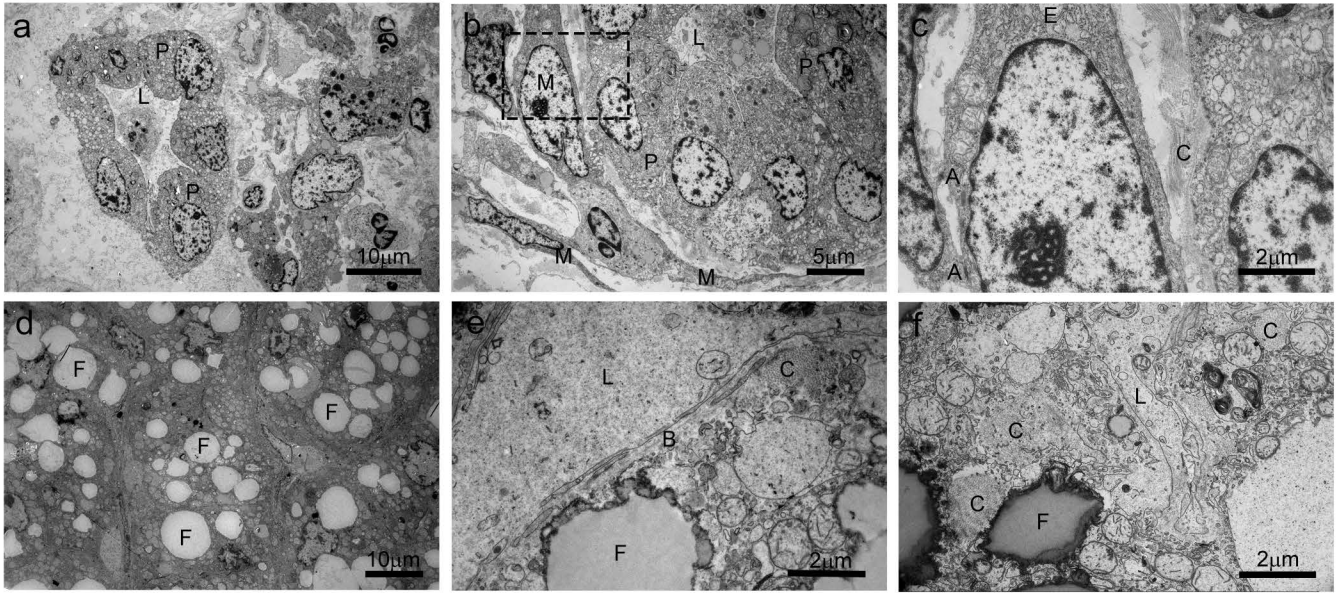
Fig. 15 The timing of dechorionation of the *nhlrc2* morphants. The fraction of dechorionated embryos was recorded for selected hours post fertilisation (hpf). The significances were calculated with a log rank Mantel-Cox test. PBS control N=91,

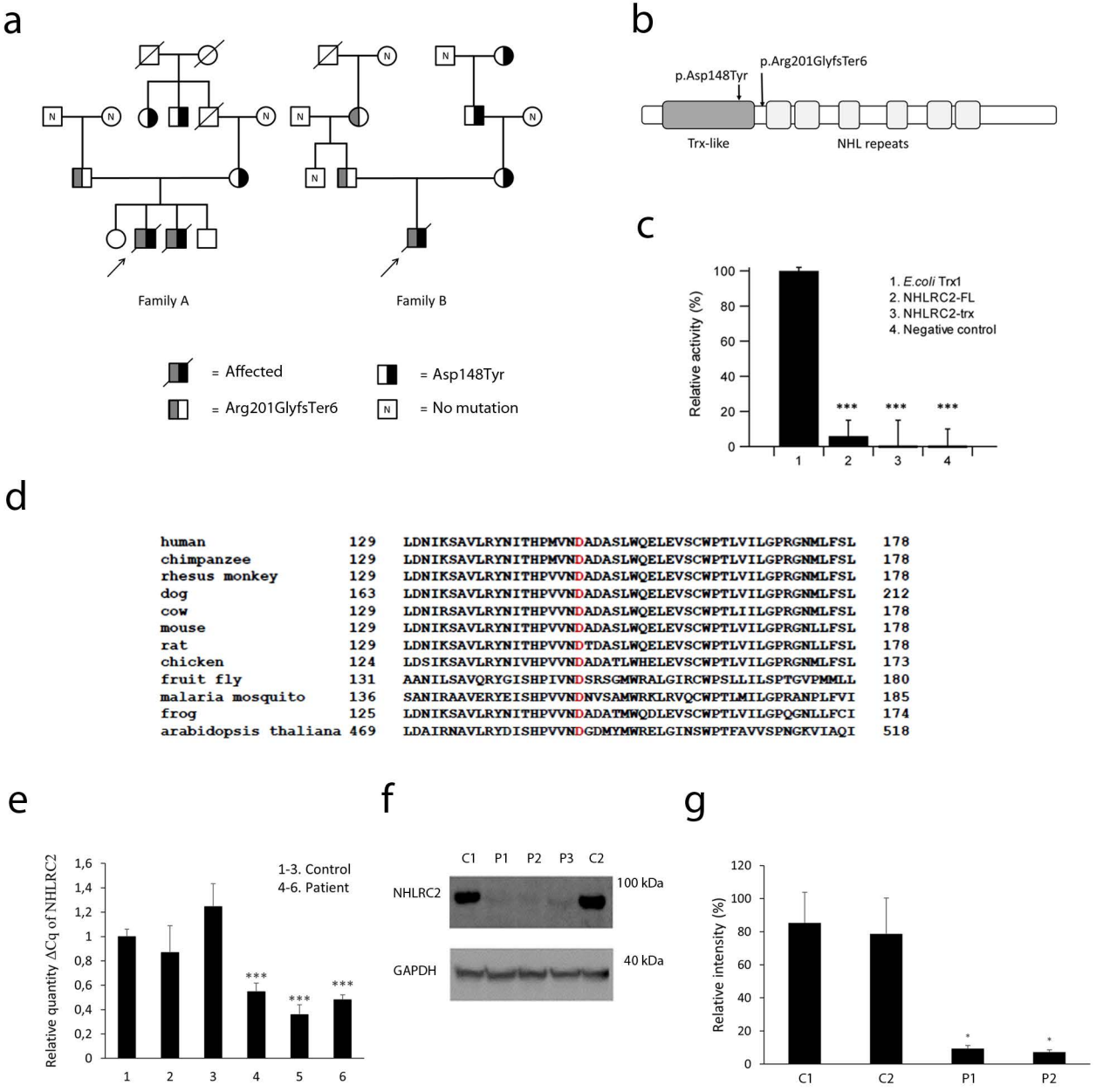
untreated control N=47; p53 MO control N=23; untreated control N=36; random control MO 3.6 ng N=63; untreated control N=34 and splice site-blocking MO 3.6 ng N= 76, untreated control N=34. Graphs were generated with GraphPadPrism 5

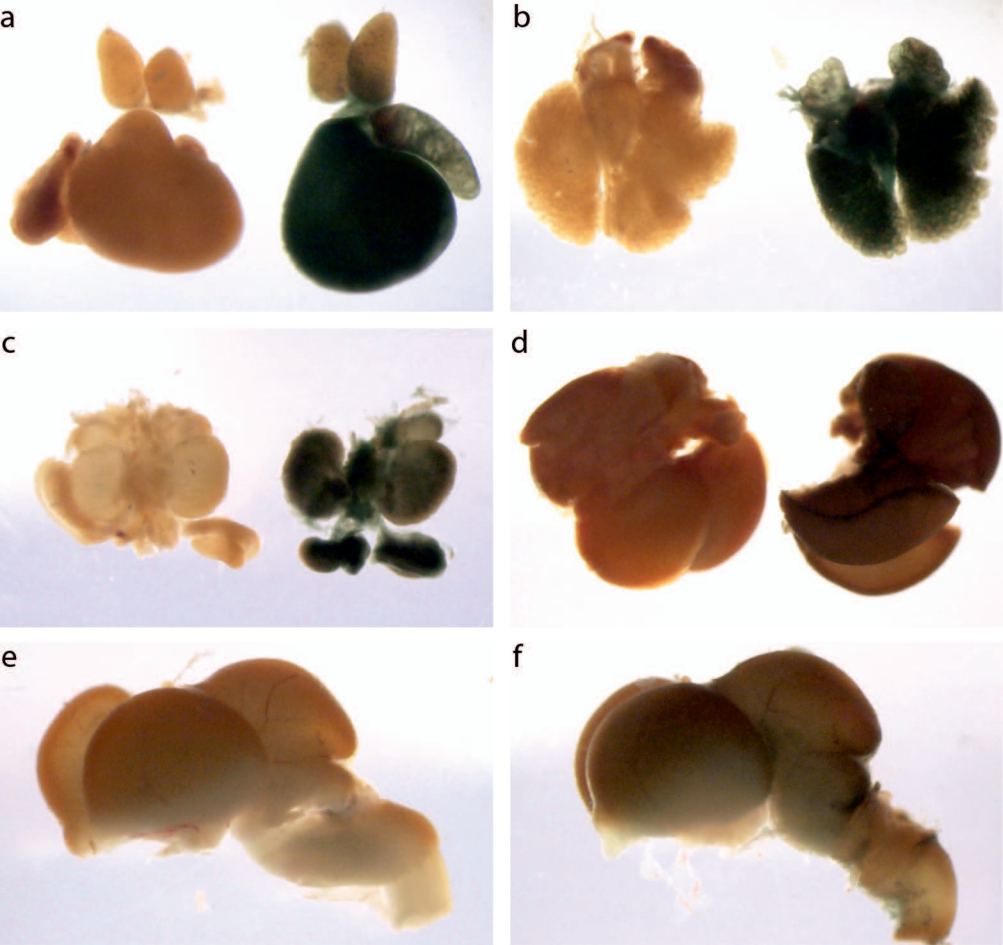
Fig. 16 Knockdown of *nhlrc2* using splice site-blocking morpholino (nhlrc2 SB MO), translation-blocking morpholino (nhlrc2 TB MO), a combination (nhlrc2 SB+TB MO) of the two morpholinos compared to un-injected controls (Control) and random control morpholino (RC MO). **a** *nhlrc2* splice site blocking morpholino alters the mRNA splicing in zebrafish embryos. The indicated amount of splice site-targeting morpholino (together with an equal amount of p53 morpholino) was injected into the yolk sacs of fertilised eggs at the 1-2 cell stage. After the indicated time, RNA was isolated, PCR was performed and PCR product was run on 2% agarose gel. Arrows on the left annotate the expected product sizes after the splice site-blocking events (334 bp intron inclusion, 198 bp WT product, 76 bp exon exclusion), and below the image are the times of RNA extraction as days post fertilisation (dpf). **b** Knockdown efficiencies were calculated using ImageJ 1.49v. The levels of WT transcript are displayed above the bars for samples containing splice site-blocking morpholinos

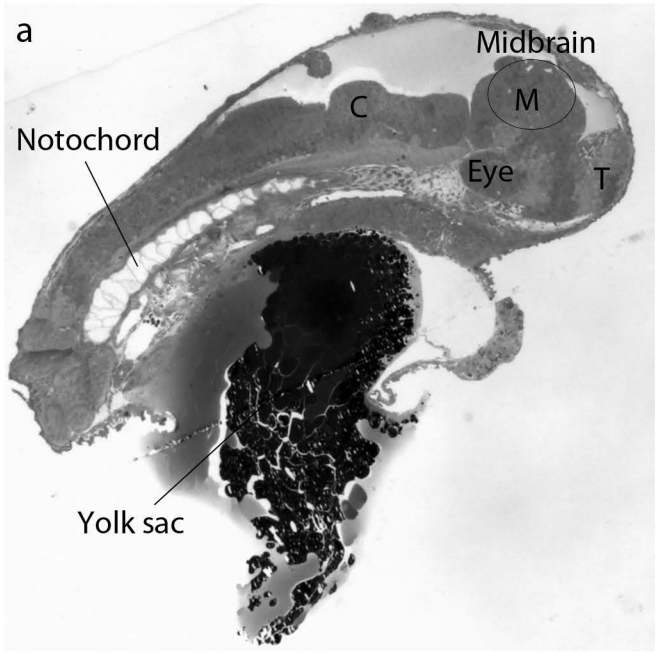

a
b
c
c
d
e
f
f
g
h
h

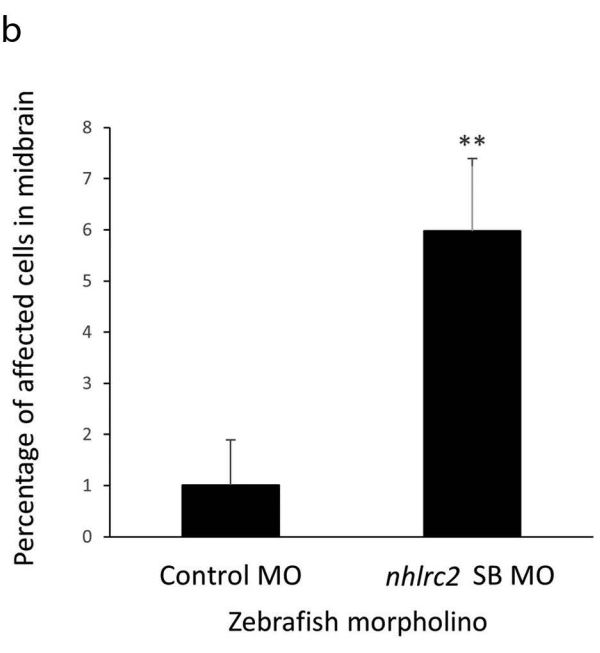

a
b
c
d
e
f
g
h

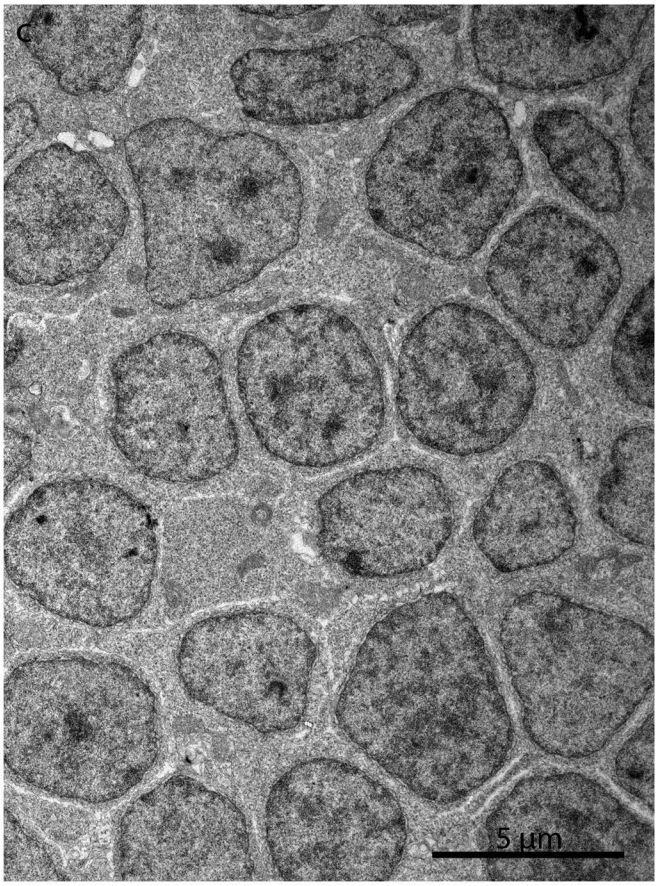

a
P
L
P
10μm
b
M
L
P
P
M
M
5μm
c
E
A
A
C
2μm
d
F
F
F
F
10μm
e
L
C
B
F
2μm
f
C
L
C
C
F
2μm

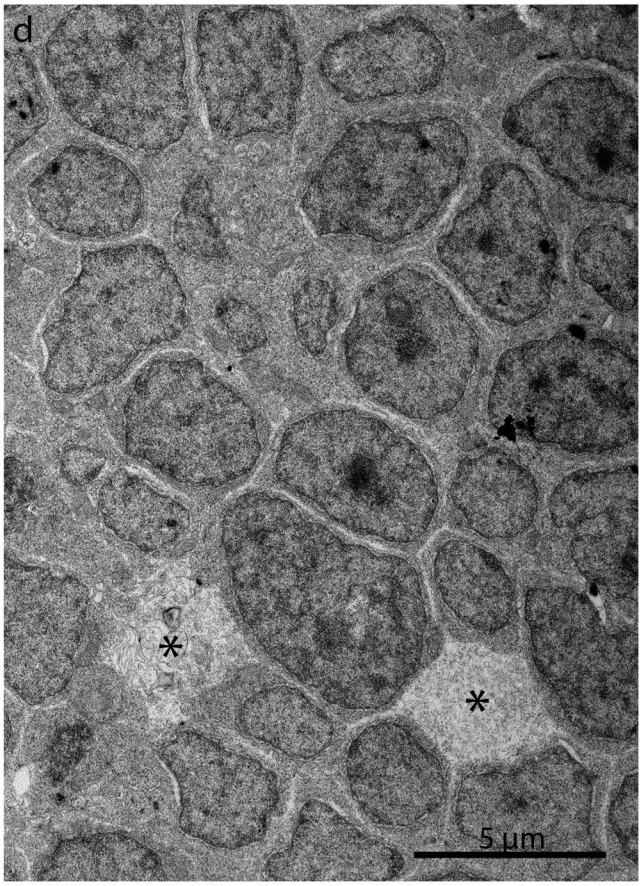

a
Family A
Family B
= Affected
= Arg201GlyfsTer6
= Asp148Tyr
= No mutation
b
p.Asp148Tyr
p.Arg201GlyfsTer6
Trx-like
NHL repeats
c
Relative activity (%)
1. E.coli Trx1
2. NHLRC2-FL
3. NHLRC2-trx
4. Negative control
d
human 129 LDNIKSAVLRYNITHPMVNDADASLWQELEVSCWPTLVILGPRGNMLFSL 178
chimpanzee 129 LDNIKSAVLRYNITHPMVNDADASLWQELEVSCWPTLVILGPRGNMLFSL 178
rhesus monkey 129 LDNIKSAVLRYNITHPVVNDADASLWQELEVSCWPTLVILGPRGNMLFSL 178
dog 163 LDNIKSAVLRYNITHPVVNDADASLWQELEVSCWPTLVILGPRGNMLFSL 212
cow 129 LDNIRSAVLRYNITHPVVNDADASLWQELEVSCWPTLIILGPRGNMLFSL 178
mouse 129 LDNIKSAVLRYNITHPVVNDADASLWQELEVSCWPTLVILGPRGNLLFSL 178
rat 129 LDNIKSAVLRYNITHPVVNDTDASLWQELEVSCWPTLVILGPRGNLLFSL 178
chicken 124 LDSIKSAVLRYNIVHPVVNDADATLWHELEVSCWPTLVILGPRGNMLFSL 173
fruit fly 131 AANILSAVQRYGISHPIVNDSRSGMWRALGIRCWPSLLILSPTGVPMMLL 180
malaria mosquito 136 SANIRAAVERYEISHPVVNDNVSAMWRKLRVQCWPTLMILGPRANPLFVI 185
frog 125 LDNIKSAVLRYNITHPVVNDADATMWQDLEVSCWPTLVILGPQGNLLFCI 174
arabidopsis thaliana 469 LDAIRNAVLRYDISHPVVNDGDMYMWRELGINSWPTFAVVSPNGKVIAQI 518
e
Relative quantity ΔCq of NHLRC2
1-3. Control
4-6. Patient
f
C1 P1 P2 P3 C2
NHLRC2
100 kDa
GAPDH
40 kDa
g
Relative intensity (%)
C1
C2
P1
P2


a
b
c
d
e
f


a
Midbrain
C
M
Notochord
Eye
T
Yolk sac


b
Percentage of affected cells in midbrain
0
1
2
3
4
5
6
7
8
**
Control MO
nhlrc2 SB MO
Zebrafish morpholino


c
5 µm


d
*
*
5 µm

Electronic supplementary material

NHLRC2 variants in patients with Fibrosis, Neurodegeneration, and Cerebral Angiomatosis (FINCA) – Characterisation of a novel cerebropulmonary disease

Johanna Uusimaa,[1–3,*] Riitta Kaarteenaho,[4,17] Teija Paakkola,[1,3,17] Hannu Tuominen,[5,6,18] Minna K. Karjalainen,[1,2,18] Javad Nadaf,[7,8,18] Teppo Varilo,[9,19] Meri Uusi-Mäkelä,[10,19] Maria Suo-Palosaari,[11] Ilkka Pietilä,[1,3] Anniina E. Hiltunen,[1,3] Lloyd Ruddock,[3,12] Heli Alanen,[3,12] Ekaterina Biterova,[3,12] Ilkka Miinalainen,[3] Annamari Salminen,[1,2] Raija Soininen,[3,12] Aki Manninen,[3,12] Raija Sormunen,[3,5] Mika Kaakinen,[3] Reetta Vuolteenaho,[3] Riitta Herva,[5] Päivi Vieira,[1,2] Teija Dunder,[1,2] Hannaleena Kokkonen,[13,14] Jukka S. Moilanen,[1,15] Heikki Rantala,[1,2] Lawrence M. Nogee,[16] Jacek Majewski,[7] Mika Rämet,[1,2,10] Mikko Hallman,[1,2] and Reetta Hinttala[1–3]

***Corresponding author:**
Johanna Uusimaa, MD, PhD
Professor of Paediatric Neurology
Tel: +358-8-3155819
Fax: +358-8-3155559
E-mail: johanna.uusimaa@oulu.fi

Case reports

Family A

Patient 1. Patient 1 was born at 38 weeks after a normal pregnancy and delivery to healthy, unrelated parents. The apgar score was 10/10, birth weight 3690 g, birth height 49 cm and head circumference 35 cm. He manifested with muscular hypotonia, dystonia and irritability at 2 months of age. He had delayed development and increased respiratory rate at the age of 3 months and alternating strabismus and loose faeces with transient diarrhoea at the age of 4 months.

When the patient was clinically examined at the age of 7 months, he was found to have dystonic tetraplegia, increased deep tendon reflexes, increased respiratory rate (48 breaths/min) with a bell-shaped chest but normal breath sounds and normal oxygen saturation values (SaO2) and poor visual contact. He also presented with unexpected unconsciousness and epileptic seizures. Laboratory investigations revealed haemolytic anaemia with decreased blood haemoglobin (75–92 g/l, reference 100–136 g/l) and blood haematocrit (0.21–0.26, reference 0.31–0.42), reticulocytosis (3.90–4.93%, reference 0.8–2.0%), increased plasma lactate dehydrogenase (1048–1698 U/l, reference 135–375 U/l) and plasma haemoglobin (129–414 mg/l, reference 0.29–2.00 g/l). The Coombs test was negative. Blood cell morphology analysis revealed anisocytosis, macrocytosis, poikilocytosis, polychromasia, hypochromia, microcytosis, ovalocytes and red cell fragments. Ophthalmological investigation revealed decreased vision in both eyes (Teller visus 20/800/1200 and 20/600/800, respectively).

From the age of 6 months, the child had breathing difficulties and recurrent respiratory and gastrointestinal symptoms but had negative screening results for common

viruses. Chest X-rays and high-resolution computed tomography (HRCT) at 7 months of age showed flattened hemidiaphragms and hyperlucent lungs, demonstrating over-inflated lungs, small air-filled bullae in the right mediobasal lung and perihilar small atelectasis (Fig. 7a).

During an acute metabolic crisis at the age of 9 months, a chest X-ray showed perihilar and basal airspace opacities and axial HRCT confirmed bilateral consolidations and reticular opacities (Fig. 7a). The patient's metabolic crisis included acute liver dysfunction with icterus and renal failure at the age of 9 months, which was triggered by viral gastroenteritis. During the subsequent 3-month period, he presented with hepatomegaly, increased liver transaminases and bilirubin, as well as icteric sclerae and skin. Extensive metabolic screening was negative, except for the following abnormal liver and kidney function tests taken during the acute metabolic crisis: increased plasma aspartyl aminotransferase (152–218 U/l, reference 10–50 U/l) and alanine aminotransferase (84–86 U/l, reference 10–50 U/l), slightly increased serum bilirubin (51–92 μmol/l, reference ≤25 μmol/l), conjugated bilirubin (62 μmol/l, reference ≤5 μmol/l), increased plasma gamma-glutamyltransferase (124–147 U/l, reference 5–80 U/l) and plasma ammonia (NH_4) (90–110 μmol/l, reference <80 μmol/l), decreased serum prealbumin (0.09–0.15 g/l, reference 0.24–0.42 g/l) and plasma albumin (27 g/l, reference 36–45 g/l), and increased plasma creatine (111–139 μmol/l, reference 15–35 μmol/l) and plasma urea (14.9–27.5 mmol/l, reference 1.7–8.3 mmol/l). Beta-hydroxybutyric acid and octandioic acid were present in urine during the metabolic crisis.

An abdominal ultrasound revealed hepatomegaly and increased cortical echogenity with a strand-like hypoechogenic outer zone in the kidneys (Fig. 9a). A nasogastric tube

was implanted when the boy was 1 year old due to feeding problems, poor weight gain and progressive growth retardation (retardation of growth from 0 standard deviations [SD] to -2 SD by the age of 7 months, decrease in relative weight from +12% to -20% and decrease in head circumference from -0.5 SD to -1.2 SD).

Brain magnetic resonance imaging (MRI) at the age of 7 months revealed a thin corpus callosum (Fig. 8a). Brain auditory evoked potentials and electroretinogram were normal, but visual evoked potentials showed increased latencies and giant responses that resembled the findings of a muscle–eye–brain (MEB) disease and late-infantile neuronal ceroid lipofuscinosis (Jansky–Bielschowsky). Therefore, a rectal biopsy was also performed but did not reveal any vacuolar cells. Electroencephalogram (EEG) revealed slow background activity and frontal bilateral rhythmic high-amplitude sharp delta transients, indicating metabolic encephalopathy.

A second episode of acute liver dysfunction at the age of 1 year 8 months was triggered by viral gastroenteritis and fever. The disease course was progressive, and the patient died at the age of 1 year 9 months due to respiratory failure.

Patient 2. Patient 2 was the younger brother of Patient 1. He was born at 38 + 6 weeks with an Apgar score of 9/9, birth weight 3280 g, height 51 cm and head circumference 34.5 cm. Similar to his older brother, feeding problems and tachypnoea were noticed at the age of 2 months.

At the age of 4 months, the infant manifested with right-sided chest prominence, a relative weight of -32% and muscular hypotonia. In contrast to his older brother, the infant was able to make eye contact and responded with a smile. However, during the follow-up,

he presented with poor weight gain, developmental delay and gastrointestinal symptoms (diarrhoea) by the age of 5 months. At that time, breath sounds and SaO2 were normal.

A chest X-ray at 4 months demonstrated mediastinal and hilar prominence, over-inflated lungs and diffuse reticular interstitial opacities (Fig. 7b). Also at 4 months, HRCT revealed widespread bilateral ground-glass opacities, interstitial infiltrations and atelectasis, mildly enlarged thymus, and left hilar adenopathy (Fig. 7b and 9f). A brain MRI at 10 months of age showed a thin corpus callosum and slightly dilated lateral ventricles and cortical sulci (Fig. 8b). At 1 year of age, chest X-rays showed the progression of interstitial infiltrations and bilateral dense consolidations (Fig. 7b). The child required oxygen support during respiratory failure triggered by a viral infection. An MRI of the liver revealed an abnormal signal (Fig. 9b-d), but the liver ultrasound was normal (Fig. 9e). A nasogastric tube was implanted at the age of 1 year due to feeding problems and progressive growth retardation that led to growth arrest. This patient did not present with seizures, and electrophysiological examinations were not performed.

Laboratory investigations revealed chronic haemolytic anaemia: decreased plasma haemoglobin (83–95 g/l, reference 100–136 g/l) and blood haematocrit (0.24–0.26, reference 0.31–0.42), reticulocytosis (3.46–3.99%, reference 0.8–2.0%), increased plasma lactate dehydrogenase (404–707 U/l, reference 150–450 U/l) and a negative Coombs test. The blood cell morphology analysis revealed anisocytosis, microcytosis, macrocytosis and polychromasia. Plasma alanine aminotransferase was elevated (184–257 U/l, reference 10–50 U/l), plasma aspartyl aminotransferase was within normal limits (48 U/l, reference ≤ 45 U/l) and both serum bilirubin (31–36 μmol/l, reference 5–25 μmol/l) and conjugated bilirubin (9 μmol/l, reference ≤5μmol/l) were slightly increased. Serum prealbumin (0.24–

0.27 g/l, reference 0.24–0.42 g/l) and plasma haemoglobin (63 mg/l, reference <100 mg/l) were within normal limits while serum haptoglobin (<0.08 g/l, reference 0.29–2.00 g/l) was at a lower level. Plasma ammonia (NH_4) was normal (37 μmol/l, reference <80 μmol/l). Furthermore, transient slightly elevated blood lactate (1.61 mmol/l, reference 0.33–1.33 mmol/l) and slightly elevated blood pyruvate (105 μmol/l, reference 30–80 μmol/l) were noted, but the lactate/pyruvate ratio was normal (15, reference <25). Serum carcinoembryonic antigen (S-CEA) was elevated at 5 months (13.2–22.1 μg/l, reference <3 μg/l), and serum neuron-specific enolase (S-NSE) levels were slightly elevated at the age of 6 months (20.1–20.4 μg/l, reference <12.5 μg/l). Because of the recurrent infections and granulomatosis that occurred in the patient's older brother, a nitroblue tetrazolium (NBT) assay was performed along with other immunological tests. NBT showed normal oxidative activity in neutrophils (95.5% versus reference 99.1%) but decreased activity in the monocytes (57% versus reference 90.9%). Urinary organic acid analysis revealed that ethylmalonic acid levels were slightly elevated.

At the age of 1 year, the boy's subcutis was thin, muscle mass was decreased, and muscle tone varied from hypotonia to hypertonia and dystonia. He presented with visual problems, alternating strabismus, and variable eye contact. The disease course was progressive, and the patient died due to respiratory failure at the age of 1 year 1 month after treatment at the intensive care unit (ICU) for 7 weeks.

Family B

Patient 3 was born at 36 + 1 gestational weeks by emergency caesarean section because of a failing heart rate. He was the first child. His mother had undergone ursodeoxycholic acid treatment because of gestational hepatosis and toxaemia. The infant's birth weight was

2910 g, and the Apgar score was 10/10. During the first week of life, he was treated twice for newborn jaundice with blue light therapy for a total duration of 4 days. ABO immunisation was suspected but not proven (serum bilirubin had increased up to 349 μmol/l at 8 days).

At 2 months of age, the patient presented with irritability, floppiness and feeding problems. Laboratory tests revealed haemolytic anaemia, with decreased plasma haemoglobin (69–84 g/l, reference 100–136 g/l) and plasma haematocrit (0.19–0.23, reference 0.31–0.42), reticulocytosis (4.0–5.1%, reference 0.8–2.0%), increased plasma lactate dehydrogenase (404–643 U/l, reference 150–450 U/l), undetectable plasma haptoglobin levels and a negative Coombs test. Blood cell morphology analysis revealed polychromasia, anisocytosis, macrocytosis, poikilocytosis and red cell fragments. The patient had frequent diarrhoea and poor weight gain. A milk allergy was suspected because of the constant diarrhoea, but changes in the formula only slightly improved his diarrhoea.

At 4 months, the infant presented with a suspected seizure during an influenza B infection, and he was hospitalised because of poor general condition, breathing problems (gasping for air), and low Sa02 values (70%). Chest X-ray at 4 months of age revealed over-inflated lungs and bilateral perihilar interstitial opacities (Fig. 7c).

At 13 months of age, chest X-rays showed diffuse infiltrates with reticulation (Fig. 7c). Thoracic HRCT confirmed diffuse bilateral ground-glass opacification, peripheral interstitial septal thickening and lobular pleural thickening (Fig. 7c). EEG was normal during wakefulness but unusually monotonic during sleep, during which it consisted of 4-Hertz delta waves without normal sleep spindles and vertex waves.

The child continued to be irritable with failure to thrive and, at age 7.5 months, he was referred to a neurologist because of slow development. Physiotherapy was started, and he was referred to an ophthalmologist because strabismus was noticed. His development did not progress.

At the age of 8 months, the child suffered almost daily from tachypnoea and cough. However, by auscultation, lung sounds were normal. At 10 months of age, he followed with his eyes but did not actively grasp toys or turn onto his stomach. He had brisk deep tendon reflexes and an indifferent plantar reflex. His breathing was tenuous, and tachypnoea was noted. The size of the child's liver was at the upper limit. Brain MRI showed slightly enlarged lateral ventricles and cortical sulci, as well as a thin corpus callosum. T2-weighted axial MRI revealed a mildly increased signal intensity of the globus pallidi (Fig. 8c).

The child underwent gastrostomy because of weight loss. Blood lactate was intermittently elevated (1.685–3.93 mmol/l, reference 0.33–1.33 mmol/l), and the plasma lactate/pyruvate ratio was at an increased level (35, reference <25). Serum thyroid-stimulating hormone (TSH) was slightly decreased (0.5–0.87 kU/l, reference 1.3–8.5 kU/l), but serum thyroid hormone levels (free T4 and T3) were normal. Immunological tests revealed that serum immunoglobulin G was slightly diminished (1.1–1.6 g/l, reference 2.7–10.1 g/l). In one instance, urinary organic acid analysis indicated small amounts of dicarboxylic acids without ketone bodies. Otherwise the extensive metabolic screening was normal. The child's breathing difficulties became progressively worse, and he died due to respiratory failure at 1 year 2 months of age after being treated at the ICU for 5 weeks.

Copy number variants and large regions of homozygosity were studied by HumanCytoSNP-12 (v2.1) (Illumina, San Diego, California, USA). The results were normal and no pathogenic variation was detected in Patient 3.

Methods

Histopathological studies of tissue samples

Tissue biopsies (lung and skeletal muscle biopsies) and autopsy samples were obtained from all three patients. The tissue was fixed in buffered 4% formaldehyde, routinely processed into paraffin blocks and cut into 5.0 μm sections. For immunohistochemistry, the sections were deparaffinised in xylene, rehydrated through graded ethanol, and stained with the following antibodies from Dako/Agilent Technologies (Santa Clara, CA, USA): CD68 (clone PG-M1, 1:100 dilution, Dako code M0876), Ck-PAN (Clone MNF-116, 1:700 dilution, Dako code M0821), caldesmon (clone h-CD, dilution 1:1000, Dako code M3557), alpha-smooth muscle actin (clone 1A4, dilution 1:500, Dako code M0851), and smooth muscle myosin heavy chain (SMMHC) (clone SMMS-1, dilution 1:100, Dako code M3558).

DNA extraction

Total genomic DNA from the blood samples of the patients and their family members, as well as from the patients' fibroblasts, was extracted with the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany). Saliva samples were collected from the family members of the affected children using Oragene DNA Collection kits (DNA Genotek, Ottawa, Canada) and the prepIT L2P kit (DNA Genotek, Ottawa, Canada) was used for DNA extraction. For

the Northern Finnish population controls, DNA was extracted from umbilical cord blood with the UltraClean DNA Blood Isolation Kit (MO BIO Laboratories, Carlsbad, CA, USA). For a subset of the samples, whole-genome amplification was performed, as previously described [4].

Whole-exome sequencing and bioinformatics

Next-generation sequencing library preparation, enrichment, and sequencing were performed using the Nimblegen V2 Exome (http://www.nimblegen.com/products/seqcap/ez/v2/) at the Technology Centre of the Institute for Molecular Medicine Finland (FIMM), University of Helsinki, as previously described [13]. DNA from the affected patients and their parents was also analysed by exome sequencing. Bioinformatics analysis was performed, as previously described [16]. Briefly, the Burrows-Wheeler Aligner (BWA v. 0.5.9) was used to map all reads to UCSC hg19. PCR duplicates were removed from alignments using Picard version 1.87 (http://broadinstitute.github.io/picard/). Indels were realigned with the Genome Analysis Toolkit (GATK 1.0.5083). SAMtools version 0.1.18 was used to call the variants. All variants were annotated with ANNOVAR and in-house scripts, and the most likely protein damage variants (nonsense, splice-site, frame shift, indel and missense) were considered for further analysis. The variants were also annotated for allele frequency in public databases such as Exome Aggregation Consortium (ExAC) (http://exac.broadinstitute.org/); the National Heart, Lung, and Blood Institute (NHLBI) exomes (V.0.0.14, http://evs.gs.washington.edu/EVS/); the 1000 Genomes Project database (http://www.internationalgenome.org) and our in-house database of >1,000

exomes. To remove common variants and sequencing artefacts, variants with minor allele frequencies of >0.01 in any of the aforementioned databases were removed from further analysis. The only candidate gene that was common to both families was *NHLRC2*.

Variant analysis and genealogy of the two families

Using Sanger sequencing, we confirmed variants and segregation within the families. The region encompassing the two *NHLRC2* variants was amplified by PCR with the primers 5′-TGG TGT TCA CTC GGC TAA GT-3′ and 5′-GTC CAT TCT TCC AAA CGA CCA-3′ and then sequenced. To investigate the descent of the variants, we performed a genealogical study in accordance with published criteria [14]. We traced the ancestors of the patients from the Finnish Population Registries and scrutinised the microfiche copies available in the National Archives of Finland.

Fibroblast cultures

Fibroblasts from healthy individuals and patients were grown at 37°C in an atmosphere of 5% CO_2 in high-glucose Dulbecco's Modified Eagle's Medium (DMEM) (Corning, Manassas, VA, USA) supplemented with 10% foetal bovine serum. The patient and control cell lines were transduced with a retroviral vector expressing the E7 gene of type-16 human papilloma virus and a retroviral vector expressing the protein component of human telomerase to immortalise the cells [17].

RNA extraction from fibroblasts

We used trypsin (Lonza, Verviers, Belgium) to harvest 1×10^7 cultured fibroblasts from

healthy individuals and patients. Total RNA was extracted with the RNeasy Plus Mini kit (Qiagen, Hilden, Germany). After reverse-transcription PCR (RT-PCR), we used Sanger sequencing to investigate the expression of the mutant alleles. The OneStep RT-PCR kit (Qiagen, Hilden, Germany) was used for the patients' cDNA analysis by qPCR.

qPCR to measure the expression of NHLRC2 mRNA

qPCR was performed according to the manufacturer's instructions (IQ™ SYBR Green Supermix, Bio-Rad, Hercules, CA, USA) using a CFX Connect™ Real-Time System (Bio-Rad, Hercules, CA, USA). Melting temperature (T_m) of the analysis was adjusted to 61°C, and TATA-Box Binding Protein (*TBP*) and Glucuronidase Beta (*GUSB*) were chosen as endogenous control genes. Primers 5´-CAGACCTCGCTAGAGTACGC-3´ and 5´-GTTCAGCCATTCTAATCCTTCCG-3´ were used for *NHLRC2*, 5´-GTCTGCGGCATTTTGTCGG-3´ and 5´-CACACGATGGCATAGGAATGG-3´ for *GUSB* and 5´-CCACTCACAGACTCTCACAAC-3´ and 5´-CTGCGGTACAATCCCAGAACT-3´ for *TBP* (Primer bank IDs 304307770c1, 268834191c1 and 285026518c1) [11,12,15].

Immunoblotting

Protein expression in the whole-cell extracts of the fibroblasts from all three patients and healthy control subjects, together with the homogenates of control autopsy samples from several types of human tissues, were analysed by immunoblotting. Whole-protein extracts from the fibroblasts were prepared in 1.5% n-dodecyl maltoside (Sigma-Aldrich, St. Louis, MO, USA) in phosphate-buffered saline (PBS). Proteins from the whole-cell extracts and

tissue homogenates were precipitated with ethanol and a total of 20–100 μg of protein per sample was used for Tris-glycine SDS-PAGE (4–20% Precise Protein Gel; ThermoFisher Scientific, Waltham, MA, USA). Proteins were transferred to a nitrocellulose membrane (Amersham Hybond ECL; GE Healthcare, Buckinghamshire, UK). Mouse polyclonal antibodies against full-length human NHLRC2 (SAB1400871, Sigma-Aldrich, St. Louis, MO, USA; NBP1-85019, Minneapolis, MN, USA; ab88725, Abcam, Cambridge, UK) and against GAPDH (GeneTex, Irvine, CA, USA) were used, together with secondary antibodies (ThermoFisher Scientific, Waltham, MA, USA) to visualize the proteins by enhanced chemiluminescence (ECL) (GE Healthcare, Buckinghamshire, UK). Signal intensities were quantified (ImageJ) and the level of NHLRC2 was compared to the level of GAPDH [10].

Expression of human NHLRC2 in tissue

Autopsy tissue specimens from human heart, kidney, skeletal muscle, liver, lung and brain were homogenised to examine the overall expression of NHLRC2 in tissues. Tissue (5 mg) was dissected on ice, 'snap frozen' in liquid nitrogen and homogenised in 300 μl of ice-cold 20 mM Tris-HCl (pH 7.5) with Pierce protease inhibitors (ThermoFisher Scientific, Waltham, MA, USA). Homogenates were maintained with constant agitation for 2 hours at 4°C and centrifuged for 20 minutes at 13,500 × *g* at 4°C in a microcentrifuge (Eppendorf, Hamburg, Germany). Protein concentrations were measured from supernatants with the Coomassie Protein Assay kit (ThermoFisher Scientific, Waltham, MA, USA). Due to rapid autolysis of the lung and brain samples, the amount of total protein obtained from these autopsy specimens was low and required precipitation with ethanol prior to

immunoblotting.

Cloning of human *NHLRC2* in *E. coli*

A gene that encodes full-length human *NHLRC2* and is a codon optimised for *E. coli* expression was purchased from GenScript (Piscataway, NJ, USA). The domain organisation of human *NHLRC2* (M1-F726) is defined in InterPro [6] as having an N-terminal thioredoxin-like domain (Q43-D200) and a six bladed beta-propeller (G210-P592). Based on this information plus secondary structure prediction [1], structural homology searches and sequence analysis, we predicted that the protein contained both domains, with the thioredoxin-like domain having an α-helical N-terminal extension and ending around G202 and the beta-propeller domain starting around P222. The region between the two could belong to either or neither domain. Accordingly, we made a series of domain constructs to express the full-length protein and fragments in *E.coli*. The DNA fragments encoding these constructs were amplified by PCR with primers that were designed to include NdeI and BamHI restriction sites on the 5′ and 3′ ends, respectively. The PCR fragments were digested with the corresponding enzymes and ligated into a modified variant of pET23 with the T7 promoter replaced by Ptac [2]. All plasmids generated were sequenced to ensure there were no errors in the cloned genes.

Expression and purification of full-length human NHLRC2 protein and thioredoxin-like domain of NHLRC2

Sixteen *NHLRC2* constructs were expressed in *E. coli* strain BL21 (DE3) with EnPresso®B medium (BioSilta, Cambridgeshire, UK) in accordance with the manufacturer's

instructions. From the analysis of the constructs made, the predicted domain boundaries were correct, and the linker region was better behaved when part of the thioredoxin domain than when only part of the beta-propeller domain. Accordingly, two constructs were used for larger scale protein production and analysis. The first construct contained the gene encoding the full-length human NHLRC2 (M1-F726) protein while the second contained DNA corresponding to the NHLRC2 thioredoxin-like domain (M1-S221). Briefly, BL21 (DE3) cells were transformed with the appropriate construct and plated onto Luria-Bertani (LB) agar plates containing 100 μg/ml ampicillin. Plates were incubated overnight at 37°C. The next morning, several colonies were transferred into 5 ml of LB medium supplemented with 100 μg/ml ampicillin. Cultures were grown at 30°C with shaking at 200 rounds per minute (rpm) for 6 hours. Five millilitre aliquots of preculture were used to seed 500 ml of EnPresso®B with 100 μg/ml ampicillin in 5 L Erlenmeyer flasks. Cultures were incubated at 30°C with shaking at 250 rpm for approximately 16 hours.

Protein production was induced by the addition of 250 μM Isopropyl β-D-1-thiogalactopyranoside (IPTG), and booster media was added. Cells were harvested by centrifugation (3,220 x g, 20 min) after 24 hours of incubation at 30°C. Cells were resuspended in a lysis buffer (50 mM sodium phosphate [pH 7.4], 500 mM NaCl, 5% glycerol, and 10 mM imidazole) supplemented with PierceTM Protease Inhibitors EDTA-free (ThermoFisher Scientific, Waltham, MA, USA), 0.1 mg/ml lysozyme and 20 μg/ml DNAse. Cells were lysed using three cycles of freeze-thawing. Cell debris was removed by centrifugation, and the soluble fraction was loaded onto a 5 ml HiTrap HP column (GE Healthcare, Buckinghamshire, UK). Specifically, bound proteins were eluted from the column with increasing concentrations of imidazole. All the collected fractions were

analysed by SDS-PAGE. Fractions containing the protein of interest were pooled together and dialysed against 20 mM Tris (pH 8.0), 25 mM NaCl and 2 mM Dithiothreitol (DTT). Dialysed protein solution was loaded into a Q HiTrap (GE Healthcare, Buckinghamshire, UK), and the protein of interest was eluted with increasing concentrations of NaCl. Full-length NHLRC2 was additionally purified by size-exclusion chromatography on a Superdex 200 column (GE Healthcare, Buckinghamshire, UK). Samples containing full-length NHLRC2 or the thioredoxin-like domain were concentrated to 9.2 and 7.5 mg/ml, respectively, flash-frozen in liquid nitrogen, and stored at -70°C.

Insulin assay

The insulin turbidity assay to test for thioredoxin activity was performed, as previously described [3]. Briefly, the assay was done in 200 µl of 0.1 M sodium phosphate buffer (pH 7.0) with 1 mM EDTA and 10 mM DTT, as well as 1µM enzymes when appropriate. The reaction was started by the addition of 1 mg/ml insulin and monitored at 540 nm for 30 minutes. Measurements were taken every 30 seconds, and the lag phase for precipitation was monitored. *E. coli* thioredoxin was used as a positive control in the assay and was produced as previously described [8].

Culturing of *Nhlrc2*^{tm1a} mouse morulae

The mating of heterozygous female and male mice was confirmed by the presence of a vaginal plug. On embryonic day 2.5 (E2.5), morulae were flushed from the uterus in accordance with a previously described protocol [7]. The uterus was cut near the cervix, and a 27 G needle attached to a 3 ml syringe was inserted into the oviduct. Both uterine

horns were flushed with PrimeQ™ M2 Mouse Embryo Culture Medium (Amsbio, Abingdon, UK). The morulae were transferred into a micro-culture drop of PrimeQ™ KSOM Mouse Embryo Medium (Amsbio, Abingdon, UK), covered with embryo-tested light mineral oil (Merck-Sigma, Temecula, CA, USA) and equilibrated in an organ culture incubator (37°C, 5% CO_2). Embryos grown overnight in the microculture drop were transferred onto a gelatinized 24-well plate, each into its own well with 300 µl of embryonic stem cell medium, and left to grow for 10–12 days, after which DNA was extracted, as described [9].

Genotyping of *Nhlrc2*^{tm1a} knockout mouse offspring and embryos

Primers CAS_R1_Term, Nhlrc2_119354_F, Nhlrc2_119354_R, LacZ_probe_F, and LacZ_probe_R (https://www.infrafrontier.eu/search?keyword= Nhlrc2) were used with Phire Hot Start II Polymerase (ThermoFisher Scientific, Waltham, MA, USA) for genotyping.

X-gal staining

To demonstrate LacZ expression in the embryos, E14.5 *Nhlrc2*^{tm1a} heterozygote mutant embryos were dissected in ice-cold PBS and fixed at room temperature for 30 min in a solution containing 0.2% glutaraldehyde, 2 mM $MgCl_2$, 5 mM EGTA and 0.1 M kanolinite phenylphosphonate (KPP) (pH 7.3) in water. They were then washed three times for 10 min in 0.02% Nonidet P-40, 0.01% Na-deoxycholate, 5 mM EGTA, 2 mM $MgCl_2$ and 0.1 M KPP pH 7.3 in water and stained by incubating in a 5-bromo-4-chloro- 3-indolyl β-D-galactopyranoside staining solution (X-gal, #R0402; ThermoFisher, Waltham, MA, USA)

composed of 10 mM $K_3Fe(CN)_6$, 10 mM $K_4Fe(CN)_6$ and 1 mg/ml X-gal in DMSO at room temperature overnight. The embryos were thereafter fixed with 4% paraformaldehyde.

Generation of zebrafish morphants

Two morpholinos against the zebrafish *nhlrc2* (ENSDARG00000089581) were designed; the first morpholino was against the ATG-site to prevent translation, and the second one was against the exon-intron boundary at the 3' end of exon 4 (ENSDARE00000919598) to prevent splicing, which both would lead to a frameshift and premature termination of translation. The regions were sequenced for polymorphism before morpholino design. Morpholino sequences were 5'-ACCCGATTCTGCTGATTTACCTTTC-3' (splice site blocking) and 5'-CTGTCAGCTTACTGTAAGACGCCAT-3' (translation blocking) (GeneTools, LLC, Philomath, OR, USA).

The morpholinos were heated to 65ºC before use to achieve complete suspension. Then, 1 nl of each morpholino (1.0– 5.0 ng) in a solution containing PBS and 0.2% rhodamine B (Sigma-Aldrich, St. Louis, MO, USA) was injected into a 1-2 cell stage embryo using a microinjector (PV830 Pneumatic PicoPump, World Precision Instruments, Sarasota, FL, USA). The morphant embryos were screened for fluorescence (570/590 nm excitation/emission) 4 hours post fertilization (hpf), and images were taken at 24, 48, 72 and 96 hpf with a fluorescence microscope (Zeiss Lumar V12, Carl Zeiss, Oberkochen, Germany). A standard random control morpholino and a p53 morpholino (GeneTools, LLC, Philomath, OR, USA) were used to control the non-specific effects of the morpholinos. WT AB zebrafish embryos were maintained according to standard protocols.

Quantitation of knockdown

The effect of the splice site-blocking morpholino for *nhlrc2* levels was quantified using PCR. Primers were designed on exon 3 and exon 5 to detect the loss of exon 4 (forward primer 5'-GTCTCCAATACTGGGCAGGTG-3' and reverse primer 5'-GCCTGCTAATGTTGAGACTTTTCC-3'). Because an effective antibody against zebrafish Nhlrc2 was not available, the effect of the translation-blocking morpholino could not be evaluated.

RNA extractions were done from four to seven pooled embryos using Tri reagent (Molecular research centre, OH, USA), and 70-300 ng of the RNA was used for cDNA synthesis using a SensiFAST cDNA synthesis kit (Bioline, Taunton, MA, USA) according to the manufacturer's instructions. The obtained RNA was treated with DNase I (Thermo Scientific, Waltham, MA, USA) according to the manufacturer's instructions. Then, 70-300 ng DNase-free RNA was used for cDNA synthesis using SensiFAST cDNA synthesis kit (Bioline, Taunton, MA, USA) according to the manufacturer's instructions. The cycling conditions for PCR (SsoFast Evagreen Supermix, Bio-Rad, Hercules, CA, USA) were 95ºC 3:00, 95ºC 0:02, 59ºC 0:05 and 59ºC for 0:10, and a melt curve with an increment of 0.5ºC to 95ºC. 5 µl of the product was run on 2% agarose gel with 1:10 000 SybrSafe DNA gel stain (Thermo Fischer Scientific, Waltham, MA, USA). The knockdown efficiency was determined from an agarose gel using ImageJ v1.49 by calculating the ratio of the WT, unmodified band to the sum of all the bands on a single lane (WT and two morpholino modified bands corresponding to exon exclusion and intron inclusion caused by splice site blocking).

Timing of Dechorionation

Dechorionation of the embryos was recorded every 3 hours between 36 hpf and 53 hpf and then at 60 hpf and 70 hpf.

Supplementary figures

Fig. 7 Radiological findings of chest X-rays and high-resolution computed tomography of Patients 1–3. **a** Patient 1 at 7 months of age shows flattened hemidiaphragms and hyperlucent lungs demonstrating over-inflated lungs, small air-filled bullae of the right mediobasal lung (short arrows) and perihilar small atelectasis (long arrow). Performed during an acute metabolic crisis at the age of 9 months, the chest X-ray demonstrates perihilar and basal airspace opacities and axial HRCT confirms bilateral consolidations (long arrows) and reticular opacities (short arrows). **b** Patient 2 at 4 months of age demonstrates mediastinal and hilar prominence (short arrows), over-inflated lungs, and diffuse reticular interstitial opacities. At 4 months of age, HRCT reveals widespread bilateral ground-glass opacities (short arrow), interstitial infiltrations (long arrows) and atelectasis (black arrowhead), mildly enlarged thymus (white arrowhead) and left hilar adenopathy (asterisk). At 1 year of age, chest X-rays show the progression of interstitial infiltrations and bilateral dense consolidations (long arrows). **c** Patient 3 at 4 months of age reveals over-inflated lungs and bilateral perihilar interstitial opacities (short arrows). At 13 months of age, the chest X-ray shows diffuse infiltrations with reticulation. HRCT confirms diffuse bilateral ground-glass opacification, peripheral interstitial septal thickening and lobular pleural thickening (long arrows)

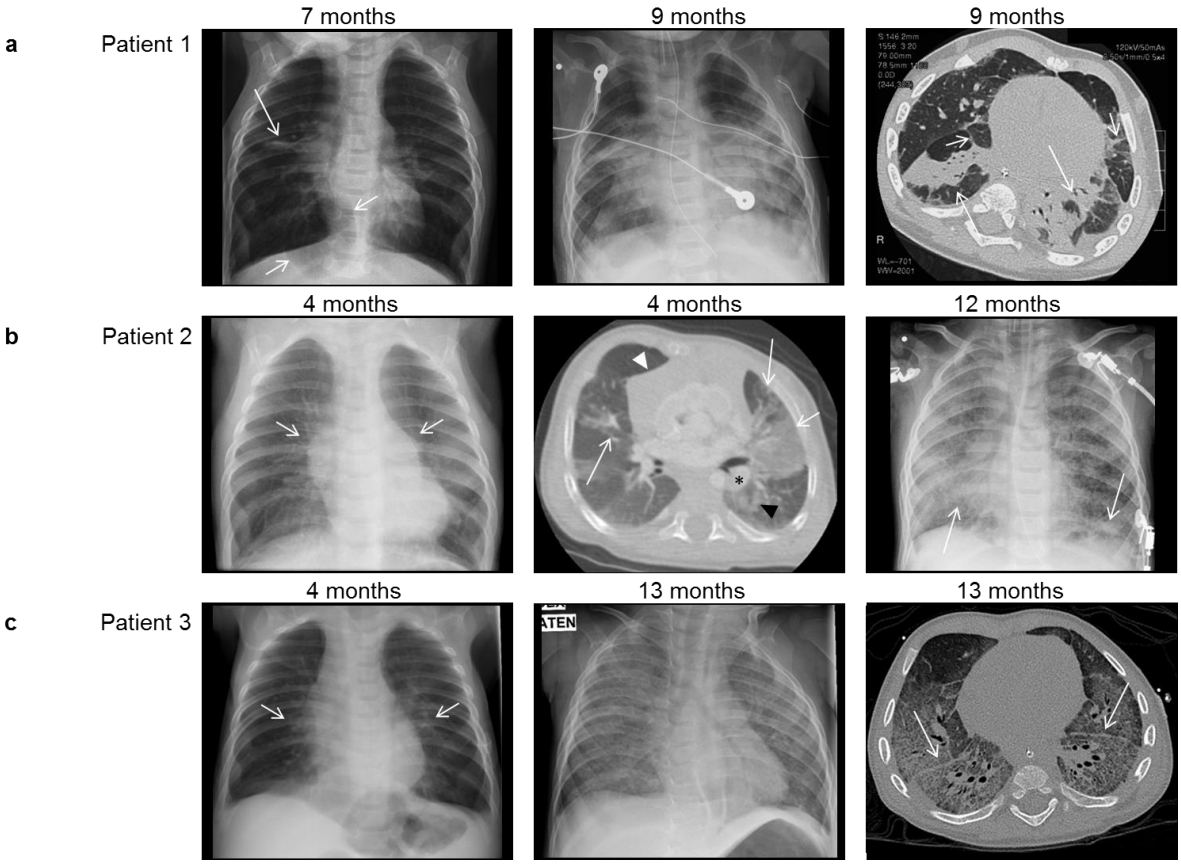

a
Patient 1
7 months
9 months
9 months
b
Patient 2
4 months
4 months
12 months
c
Patient 3
4 months
13 months
13 months

Fig. 8 Findings from brain magnetic resonance imaging of Patients 1–3. **a** Sagittal T1-weighted magnetic resonance imaging (MRI) shows the thin corpus callosum of Patient 1 at 7 months of age. **b** Patient 2 at 10 months of age. **c** Patient 3 at 10 months of age (short arrows). The axial T2-weighted and coronal T1-weighted MRIs of Patients 2 (**b**) and 3 (**c**) demonstrate slightly dilated lateral ventricles and cortical sulci. **c** Increased signal intensity of the globus pallidus on the axial T2-weighted MRI of Patient 3 (arrowheads)

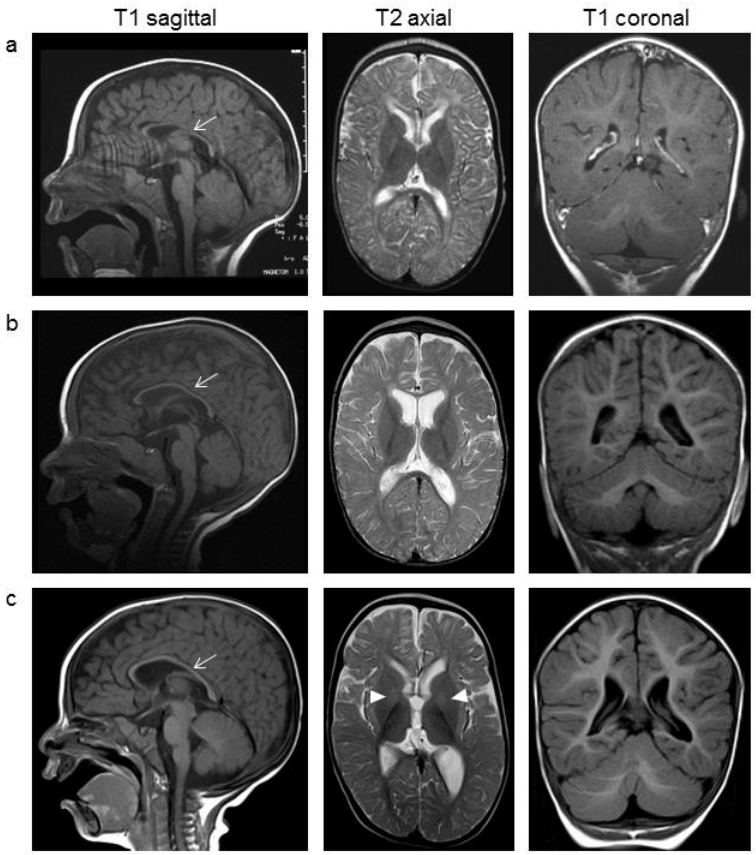


Fig. 9 Findings from abdominal ultrasound and magnetic resonance imaging of internal organs and thymus of Patients 1 and 2. **a** The abdominal ultrasound findings (Patient 1 at 10 months of age) show increased cortical echogenity with the strand-like hypoechogenic outerzone of both kidneys (short arrows). **b** The abdominal T2-weighted axial MRI of Patient 2 at 4 months of age shows hepatomegaly and a non-expansive hypointense lesion of the right liver lobe (long arrows). **c** The lesion is also hypointense on the T1-weighted coronal MRI (long arrows). **d** The axial fat saturation T1-weighted MRI of Patient 2 shows that the lesion (arrowheads) in the liver is less enhanced by the gadolinium contrast agent than the rest of the liver. **e** Normal echogenity of the right liver lobe of Patient 2 at the age of 4 months. **f** A slightly enlarged thymus of Patient 2, which was seen in the coronal T1-weighted MRI (asterisk)

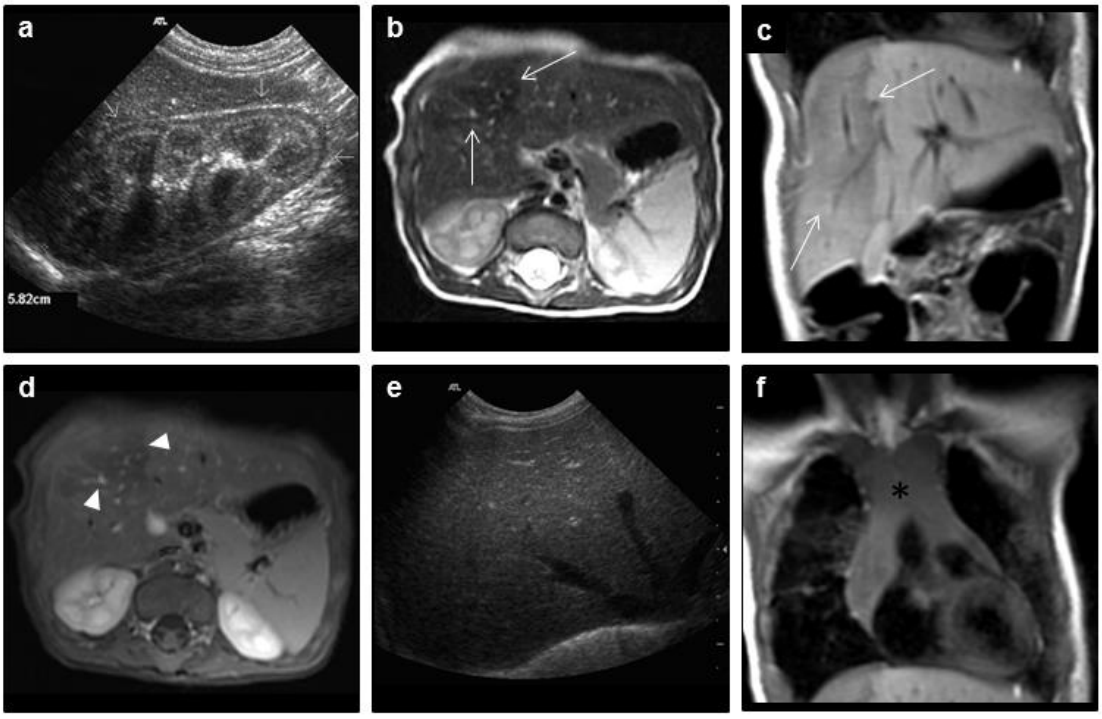


Fig. 10 a Electropherograms showing both heterozygote NM_198514:c.442G>T and c.601_602delAG variants in genomic DNA from the patient. **b** The RNA extracted from the patients' fibroblasts has only c.442G>T. Thus, Sanger sequencing of mRNA confirms that the variants reside in separate haplotypes and have been inherited as compound heterozygotes in all patients. Furthermore, the absence of the c.601_602delAG variant indicates that the transcript with the frameshift deletion is processed through nonsense mediated RNA decay

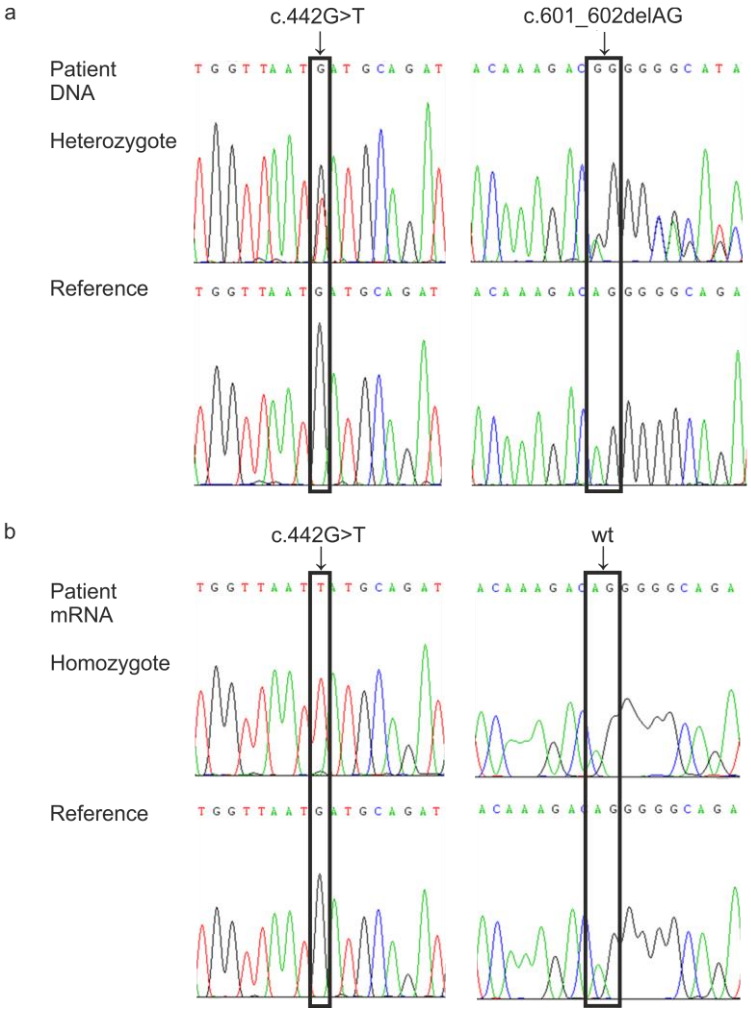


Fig. 11 Immunoblotting to detect NHLRC2 protein expression in human tissue homogenates from control autopsy samples. NHLRC2 was detected in all studied human organs including the heart (He), kidney (Ki), muscle (Mu), liver (Li), lung (Lu) and brain (Br). GAPDH was used as a loading control

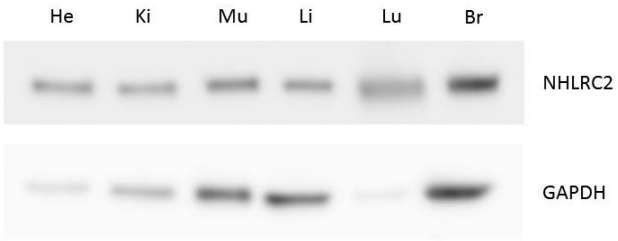
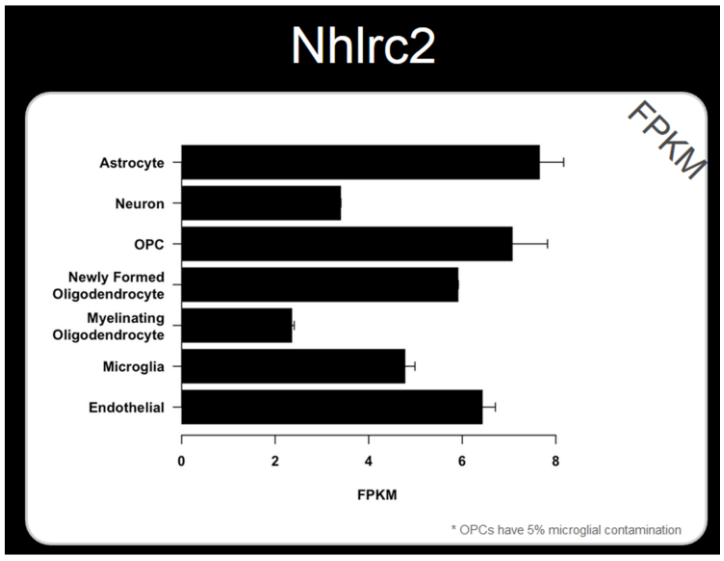
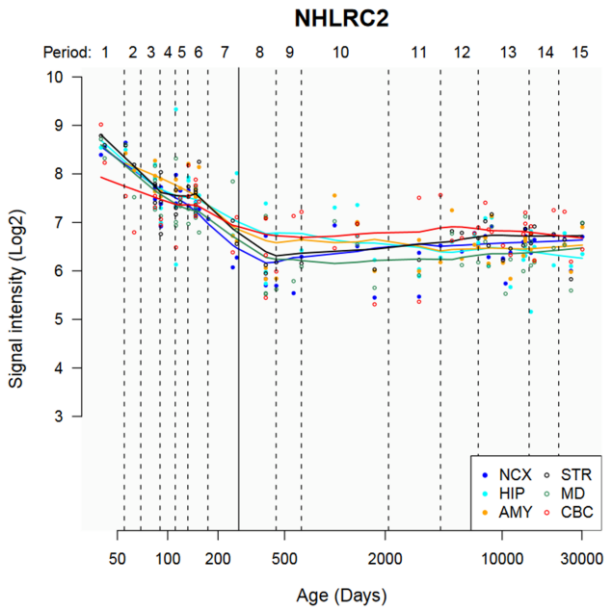


Fig. 12 *Nhlrc2* is expressed in various cell types of the mouse brain and *NHLRC2* expression is highest in the early human life. **a** An RNA-Seq transcriptome and splicing database of the glia, neurons and vascular cells of the cerebral cortex show the highest *Nhlrc2* expression in astrocytes, oligodendrocyte progenitor cells, newly formed oligodendrocytes and endothelial cells. *Nhlrc2* was also expressed in neurons, myelinating oligodendrocytes and microglia. (OPG) oligodendrocyte progenitor cells [18]. Abbreviations: FPKM, fragments per kilobase per million. **b** The average expression levels of *NHLRC2* are similar between the cerebellar cortex (CBC), mediodorsal nucleus of the thalamus (MD), striatum (STR), amygdala (AMY), hippocampus (HIP) and 11 areas of the neocortex (NCX) in the human brain (http://hbatlas.org/). **c** The expression is highest in the early development of the human brain. The number of samples is indicated on the top of the x-axis. Data from www.brainspan.org [5]. Abbreviations: RPKM, reads per kilobase per million; pcw, post conception week; mos, months; yrs, years

a



b


NHLRC2
Period: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Signal intensity (Log2)
10
9
8
7
6
5
4
3
NCX
HIP
AMY
STR
MD
CBC
50 100 200 500 2000 10000 30000
Age (Days)

c

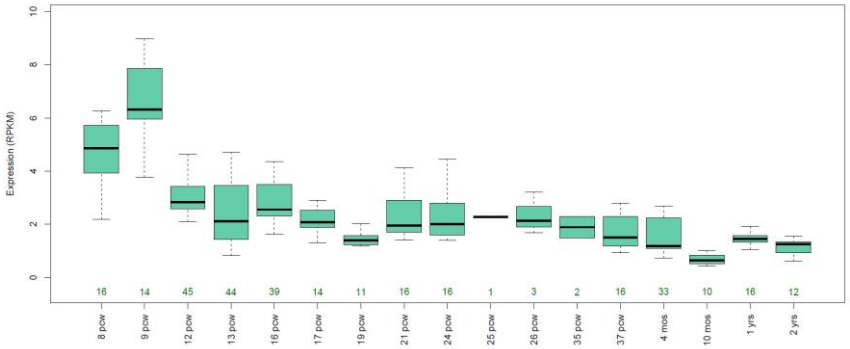

Expression (RPKM)
0
2
4
6
8
10
16 14 45 44 39 14 11 16 16 1 3 2 16 33 10 16 12
8 pcw
9 pcw
12 pcw
13 pcw
16 pcw
17 pcw
19 pcw
21 pcw
24 pcw
25 pcw
26 pcw
35 pcw
37 pcw
4 mos
10 mos
1 yrs
2 yrs

Fig. 13 Wild-type (WT) and heterozygous (het) *Nhlrc2* KO morulae and their growth *in vitro*. The morulae were isolated at embryonic day (E) 2.5 and co-cultured in a micro drop culture overnight. On the next day, the E3.5 compacted morulae and early blastocysts were transferred to separate gelatinised wells. At E4.5, the embryos had matured to the blastocyst stage and at E5.5 they hatched and attached to the bottom of the plate. Cells were grown for several days to detect possible phenotypes and to obtain enough material for genotyping (E7.5, E10.5 and E13.5). No homozygous morulae were detected

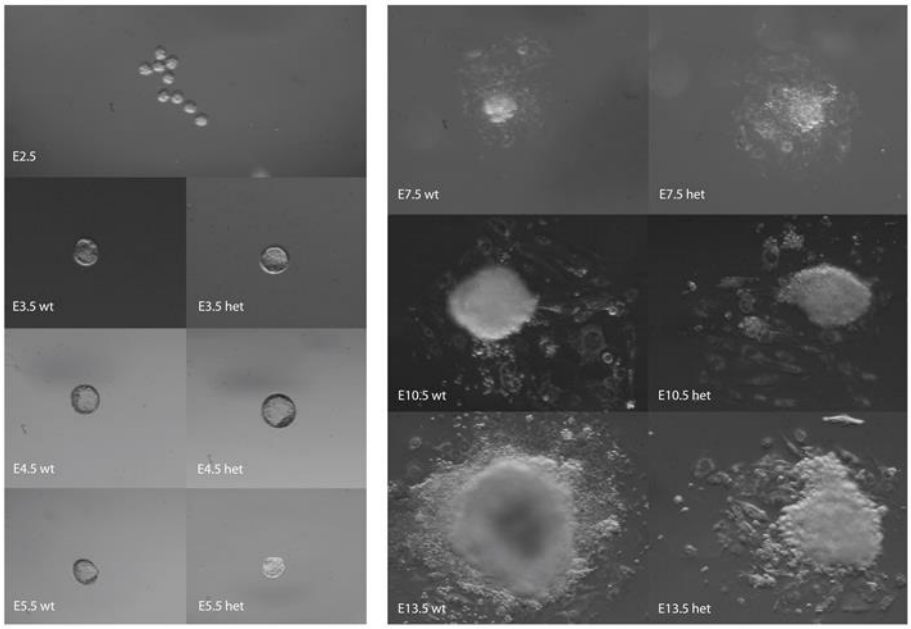
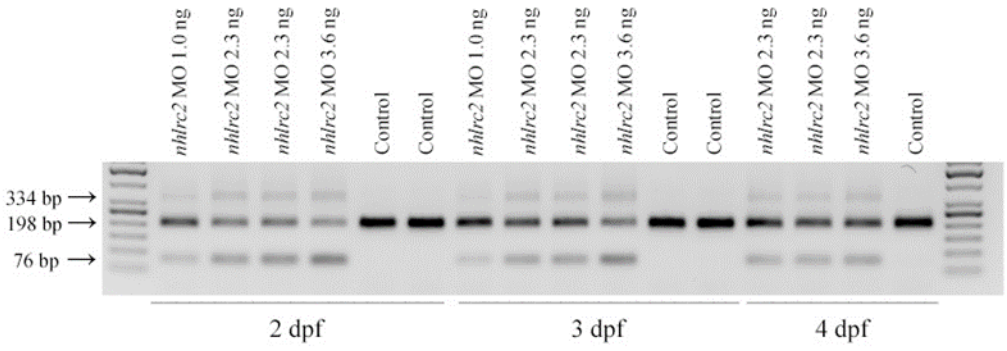


Fig. 14 *nhlrc2* splice site-blocking morpholino alters the mRNA spicing in zebrafish embryos but not the gross morphology of the embryos. Indicated amount of splice site targeting morpholino (together with an equal amount of p53 morpholino) was injected into the yolk sacs of fertilised eggs at the 1-2 cell stage. **a** Arrows on the left show the expected product sizes after the splice site-blocking events (334 bp intron inclusion, 198 bp WT product, 76 bp exon exclusion), and below the image are indicated the times of RNA extraction as days post fertilisation (dpf). **b** Un-injected controls on the left, *nhlrc2* splice site-blocking morpholino (3.6 ng) treated morphants in the middle and *nhlrc2* translation-blocking morpholino (2.3 ng) treated morphants on the right. The embryos appear to have WT morphology; only minor swelling was observed in the abdominal area in some individuals in both morphant groups at 3 dpf. However, this swelling diminished by 4 dpf. The images were taken with Axiovision software version 4.8 using a Zeiss SteREO Lumar V12 fluorescence microscope, equipped with 1.5X camera (Carl Zeiss MicroImaging GmbH, Göttingen, Germany)

a



b

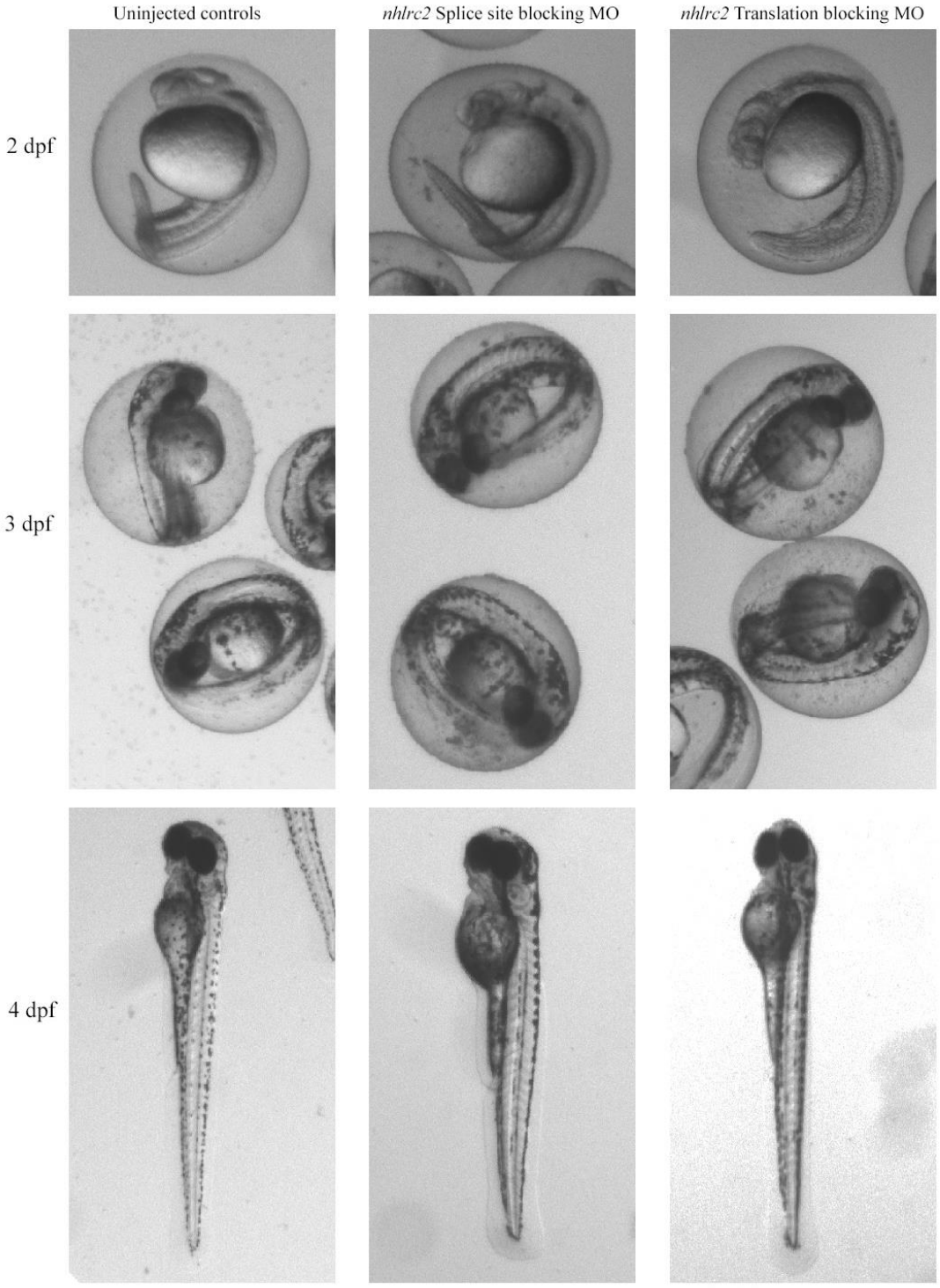

Uninjected controls
nhlrc2 Splice site blocking MO
nhlrc2 Translation blocking MO
2 dpf
3 dpf
4 dpf

Fig. 15 The timing of dechorionation of the *nhlrc2* morphants. The fraction of dechorionated embryos was recorded for selected hours post fertilisation (hpf). The significances were calculated with a log rank Mantel-Cox test. PBS control N=91, untreated control N=47; p53 MO control N=23; untreated control N=36; random control MO 3.6 ng N=63; untreated control N=34 and splice site-blocking MO 3.6 ng N= 76, untreated control N=34. Graphs were generated with GraphPadPrism 5

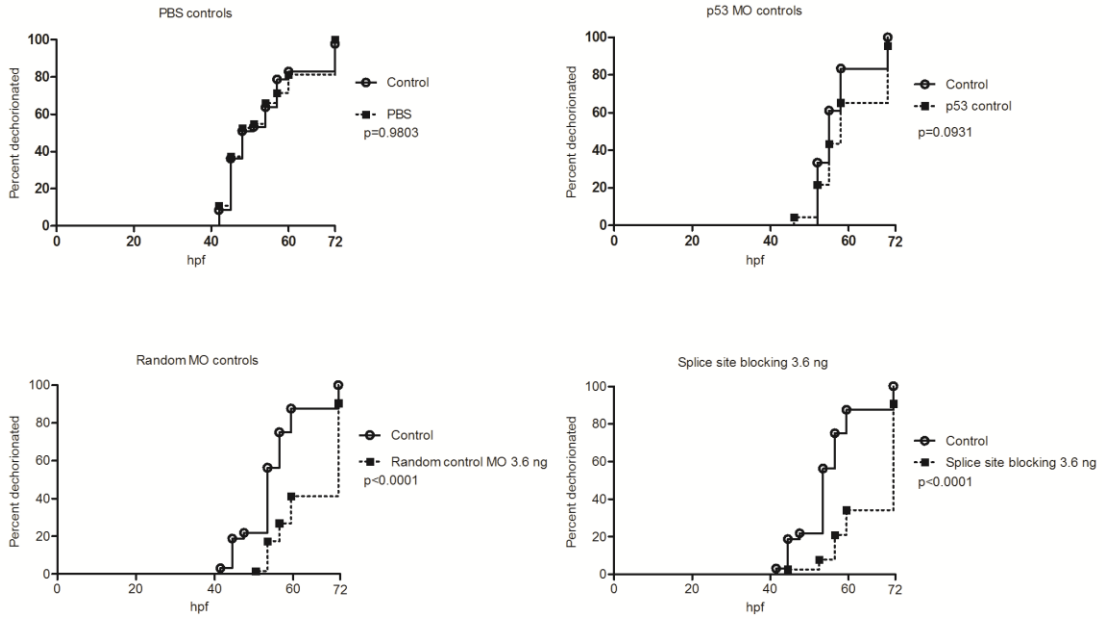
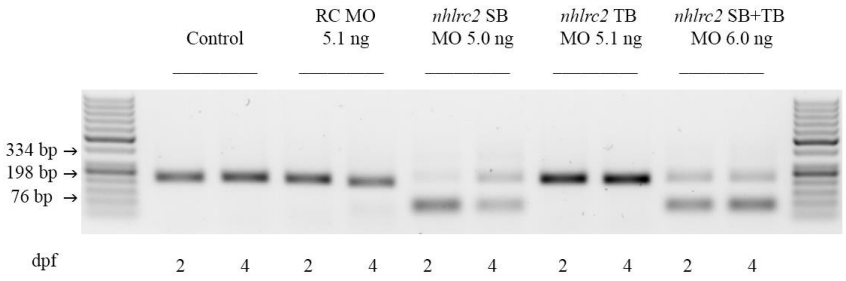
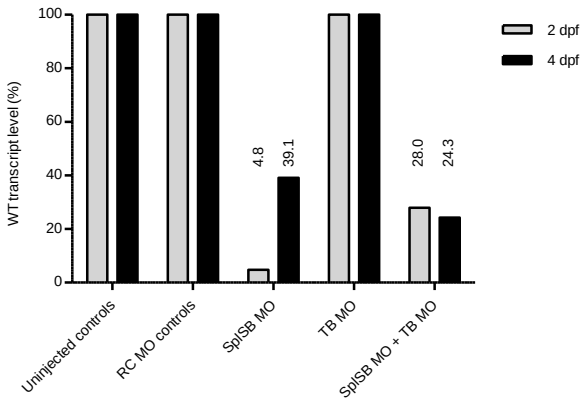


Fig. 16 Knockdown of *nhlrc2* using splice site-blocking morpholino (nhlrc2 SB MO), translation-blocking morpholino (nhlrc2 TB MO), a combination (nhlrc2 SB+TB MO) of the two morpholinos compared to un-injected controls (Control) and random control morpholino (RC MO). **a** *nhlrc2* splice site blocking morpholino alters the mRNA splicing in zebrafish embryos. The indicated amount of splice site-targeting morpholino (together with an equal amount of p53 morpholino) was injected into the yolk sacs of fertilised eggs at the 1-2 cell stage. After the indicated time, RNA was isolated, PCR was performed and PCR product was run on 2% agarose gel. Arrows on the left annotate the expected product sizes after the splice site-blocking events (334 bp intron inclusion, 198 bp WT product, 76 bp exon exclusion), and below the image are the times of RNA extraction as days post fertilisation (dpf). **b** Knockdown efficiencies were calculated using ImageJ 1.49v. The levels of WT transcript are displayed above the bars for samples containing splice site-blocking morpholinos

a



b


100
80
60
40
20
0
WT transcript level (%)
4.8
39.1
28.0
24.3
2 dpf
4 dpf
Uninjected controls
RC MO controls
SplSB MO
TB MO
SplSB MO + TB MO

References

1. Drozdetskiy A, Cole C, Procter J, Barton GJ (2015) JPred4: a protein secondary structure prediction server. Nucleic Acids Res 43:W389-94. Doi:10.1093/nar/gkv332

2. Gaciarz A, Veijola J, Uchida Y, Saaranen MJ, Wang C, Horkko S, et al. (2016) Systematic screening of soluble expression of antibody fragments in the cytoplasm of E. coli. Microb Cell Fact 15:22-016-0419-5. Doi:10.1186/s12934-016-0419-5

3. Karala AR, Ruddock LW (2010) Bacitracin is not a specific inhibitor of protein disulfide isomerase. FEBS J 277:2454-2462. Doi:10.1111/j.1742-4658.2010.07660.x

4. Karjalainen MK, Huusko JM, Ulvila J, Sotkasiira J, Luukkonen A, Teramo K, et al. (2012) A potential novel spontaneous preterm birth gene, AR, identified by linkage and association analysis of X chromosomal markers. PLoS One 7:e51378. Doi:10.1371/journal.pone.0051378

5. Miller JA, Ding SL, Sunkin SM, Smith KA, Ng L, Szafer A, et al. (2014) Transcriptional landscape of the prenatal human brain. Nature 508:199-206. Doi:10.1038/nature13185

6. Mitchell A, Chang HY, Daugherty L, Fraser M, Hunter S, Lopez R, et al. (2015) The InterPro protein families database: the classification resource after 15 years. Nucleic Acids Res 43:D213-21. Doi:10.1093/nar/gku1243

7. Nagy A, Gertsenstein M, Vintersten K, Behringer R (2003) Manipulating the Mouse Embryo, A Laboratory Manual, 3rd ed. p.198-200. Cold Spring Harbor Laboratory Press, New york

8. Saaranen MJ, Karala AR, Lappi AK, Ruddock LW (2010) The role of dehydroascorbate in disulfide bond formation. Antioxid Redox Signal 12:15-25. Doi:10.1089/ars.2009.2674

9. Scavizzi F, Ryder E, Newman S, Raspa M, Gleeson D, Wardle-Jones H, et al. (2015) Blastocyst genotyping for quality control of mouse mutant archives: an ethical and economical approach. Transgenic Res 24:921-927. Doi:10.1007/s11248-015-9897-1

10. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9:671-675

11. Spandidos A, Wang X, Wang H, Seed B (2010) PrimerBank: a resource of human and mouse PCR primer pairs for gene expression detection and quantification. Nucleic Acids Res 38:D792-9. Doi:10.1093/nar/gkp1005

12. Spandidos A, Wang X, Wang H, Dragnev S, Thurber T, Seed B (2008) A comprehensive collection of experimentally validated primers for Polymerase Chain

Reaction quantitation of murine transcript abundance. BMC Genomics 9:633-2164-9-633. Doi:10.1186/1471-2164-9-633

13. Sulonen AM, Ellonen P, Almusa H, Lepisto M, Eldfors S, Hannula S, et al. (2011) Comparison of solution-based exome capture methods for next generation sequencing. Genome Biol 12:R94-2011-12-9-r94. Doi:10.1186/gb-2011-12-9-r94

14. Varilo T, Savukoski M, Norio R, Santavuori P, Peltonen L, Jarvela I (1996) The age of human mutation: genealogical and linkage disequilibrium analysis of the CLN5 mutation in the Finnish population. Am J Hum Genet 58:506-512

15. Wang X, Seed B (2003) A PCR primer bank for quantitative gene expression analysis. Nucleic Acids Res 31:e154

16. Witkowski L, Carrot-Zhang J, Albrecht S, Fahiminiya S, Hamel N, Tomiak E, et al. (2014) Germline and somatic SMARCA4 mutations characterize small cell carcinoma of the ovary, hypercalcemic type. Nat Genet 46:438-443. Doi:10.1038/ng.2931

17. Yao J, Shoubridge EA (1999) Expression and functional analysis of SURF1 in Leigh syndrome patients with cytochrome c oxidase deficiency. Hum Mol Genet 8:2541-2549. Doi:ddc275

18. Zhang Y, Chen K, Sloan SA, Bennett ML, Scholze AR, O'Keeffe S, et al. (2014) An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. J Neurosci 34:11929-11947. Doi:10.1523/JNEUROSCI.1860-14.2014