



NHLRC2 variants identified in patients with fibrosis, neurodegeneration, and cerebral angiomas (FINCA): characterisation of a novel cerebropulmonary disease

Johanna Uusimaa^{1,2,3} · Riitta Kaarteenaho^{4,5} · Teija Paakkola^{1,3} · Hannu Tuominen^{6,7} · Minna K. Karjalainen¹ · Javad Nadaf^{8,9} · Teppo Varilo¹⁰ · Meri Uusi-Mäkelä¹¹ · Maria Suo-Palosaari¹² · Ilkka Pietilä^{1,3} · Anniina E. Hiltunen^{1,3} · Lloyd Ruddock^{3,13} · Heli Alanen^{3,13} · Ekaterina Biterova^{3,13} · Ilkka Miinalainen³ · Annamari Salminen¹ · Raija Soinen^{3,13} · Aki Manninen^{3,13} · Raija Sormunen^{3,6} · Mika Kaakinen³ · Reetta Vuolteenaho³ · Riitta Herva⁶ · Päivi Vieira^{1,2} · Teija Dunder^{1,2} · Hannaleena Kokkonen^{14,15} · Jukka S. Moilanen^{1,16} · Heikki Rantala^{1,2} · Lawrence M. Noguee¹⁷ · Jacek Majewski⁸ · Mika Rämetsä^{1,2,11} · Mikko Hallman^{1,2} · Reetta Hinttala^{1,3}

Received: 8 September 2017 / Revised: 30 January 2018 / Accepted: 31 January 2018 / Published online: 8 February 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

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 2 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, neuropathology revealed increased angiomas-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 angiomas.

Riitta Kaarteenaho and Teija Paakkola contributed equally to this work.

Hannu Tuominen, Minna K. Karjalainen, and Javad Nadaf contributed equally to this work.

Teppo Varilo and Meri Uusi-Mäkelä contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00401-018-1817-z>) contains supplementary material, which is available to authorized users.

Extended author information available on the last page of the article

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 [2, 21]. 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 [3]. Chronic, progressive pulmonary disorders include interstitial fibrosis, granulomatous changes and diffuse alveolar damage caused by various pathogenetic and genetic mechanisms [24]. 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 uncharacterised, 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 suggest 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 to 2003 (Patients 1 and 2 from Family A) and from 2010 to 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. The Sanger sequencing revealed no pathogenic variants in the coding region of mtDNA, 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 (see 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- μ m sections. Hematoxylin and eosin and luxol fast blue stainings were performed as previously described [1].

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 days post-fertilisation (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 (see Electronic Supplementary Material)

Whole-exome sequencing (WES) [25] 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 [29].

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 weeks, total $n = 306$) sampled prospectively at Oulu University Hospital from 2004 to 2007 ($n = 199$) [10] 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 [13] (<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 (see 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 [9].

Nhlrc2 knockout mice (see 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 [26]. 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 [18]. Embryos that were grown overnight in a microculture drop were transferred onto a gelatinised 24-well plate, each into its own well with 300 μ l of embryonic stem cell medium, and left to grow for 10–12 days, after which DNA was extracted, as described [22].

Generation of zebrafish morphants (see Electronic Supplementary Material)

Two morpholinos against zebrafish *nhlrc2* (ENS-DARG00000089581) 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 entirely within the imaged area were counted and analysed for their integrity using iTEM software (Olympus Software Imaging Solutions GmbH, Münster, Germany). 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's t test was used to analyse the statistical significance. Zebrafish dechoriation 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 hepatitis and toxemia. 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, 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

Table 1 Clinical findings for the three patients with FINCA disease

Findings	Patient 1	Patient 2	Patient 3
Duration of gestation	38 weeks	38 + 6 weeks	36 + 1 weeks
Birth weight/length/head circumference	3690 g/49 cm/35 cm	3280 g/51 cm/34.5 cm	2910 g/46 cm/33 cm
Apgar score	10/10	9/9	10/10
Age at disease onset	2 months	2 months	Birth to 2 months
Neurological symptoms			
Irritability	+	+	+
Axial hypotonia	+	+	+
Dystonia	+	+	+
Developmental delay	+	+	+
Feeding problems	+	+	+
Poor visual contact	+	+	+
Seizures	+	–	+
Respiratory symptoms			
Recurrent respiratory infections	+	+	+
Progressive respiratory insufficiency	+	+	+
Ophthalmological findings			
Decreased vision	+	+	+
Strabismus	+	+	+
Gastrointestinal and cardiac symptoms and signs			
Diarrhoea	+	+	+
Hepatomegaly	+	+	+
Feeding problems, poor weight gain, growth failure	+	+	+
Cardiomegaly	+	–	–
Haematological findings			
Chronic haemolytic anaemia	+	+	+
Metabolic crises			
Transient liver dysfunction during viral infection	+	+	–
Transient kidney dysfunction during viral infection	+	–	–
Progressive course of disease	+	+	+
Age at death	1 year 9 months	1 year 1 month	1 year 2 months

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, 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

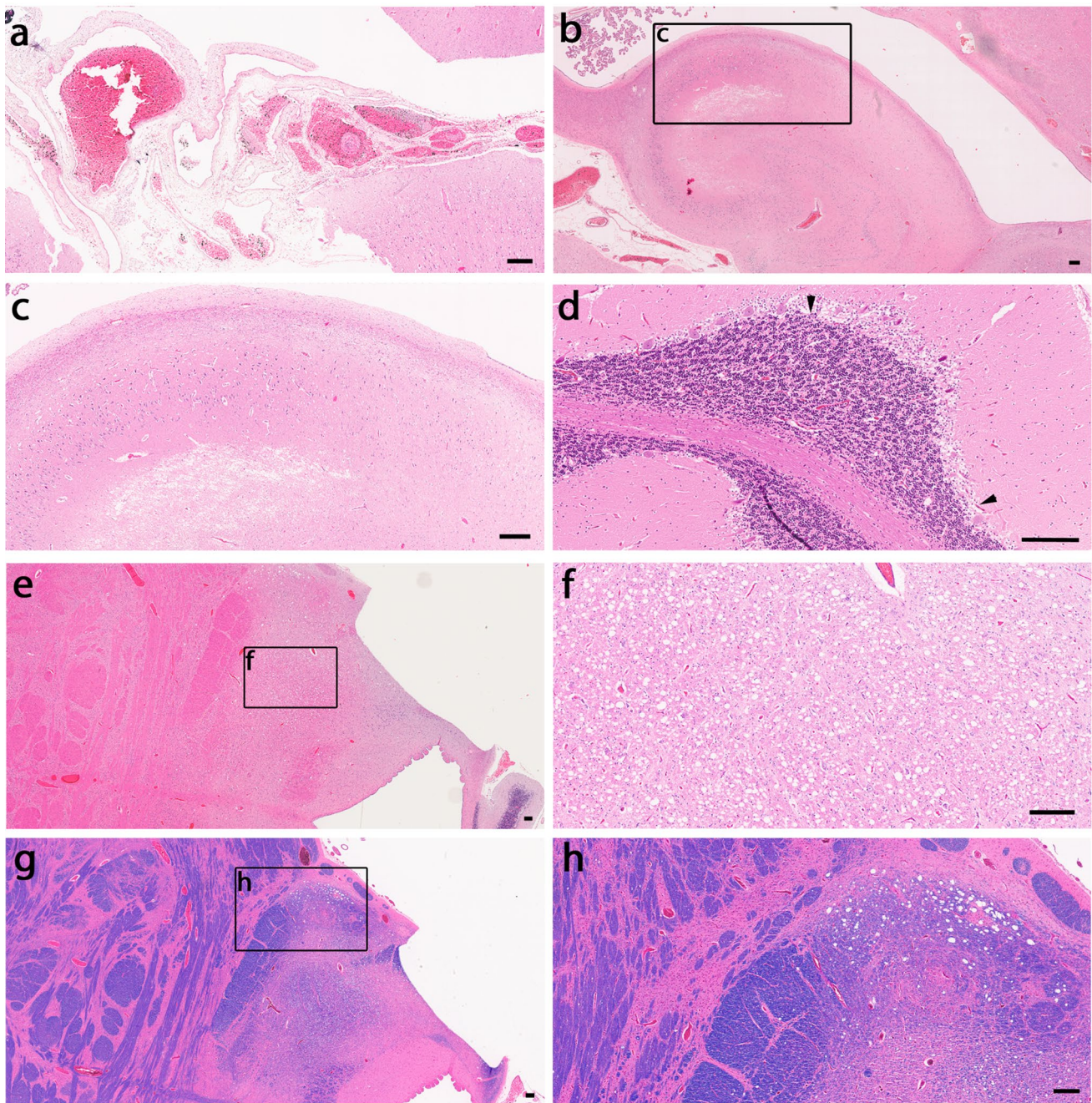


Fig. 1 Histopathological findings in brain autopsy samples from patients with NHLRC2 variants. **a** Autopsy findings from Patient 1 revealed increased angiomas-like leptomenigeal vascularisation. **b** Degeneration and depletion of the pyramidal cells of the hippocampus. **c** Higher magnification of the Ammon's horn showing gradual depletion of pyramidal neurons. **d** Depletion of 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 μ m

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, h). Hyaline-resembling fibrotic areas were observed around the arteries in all cases, but the walls of the vessels were otherwise normal (Fig. 2d, 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 I alveolar cells were detected (Fig. 3a, 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, f).

Variants of *NHLRC2* in the two families

WES revealed compound heterozygous variants in *NHLRC2*, in all three affected children (Fig. 4a). The *NHLRC2* gene is located in chromosome 10:

113,854,661–113,917,194 and encodes for the poorly characterised *NHLRC2* containing 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) [23]. 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, 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, b, Online Resource Fig. 10). Both the variants were segregated from their healthy parents. Genealogies of the patients, traced back 7–9 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 ($n = 8252$), and African ($n = 5202$) 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.

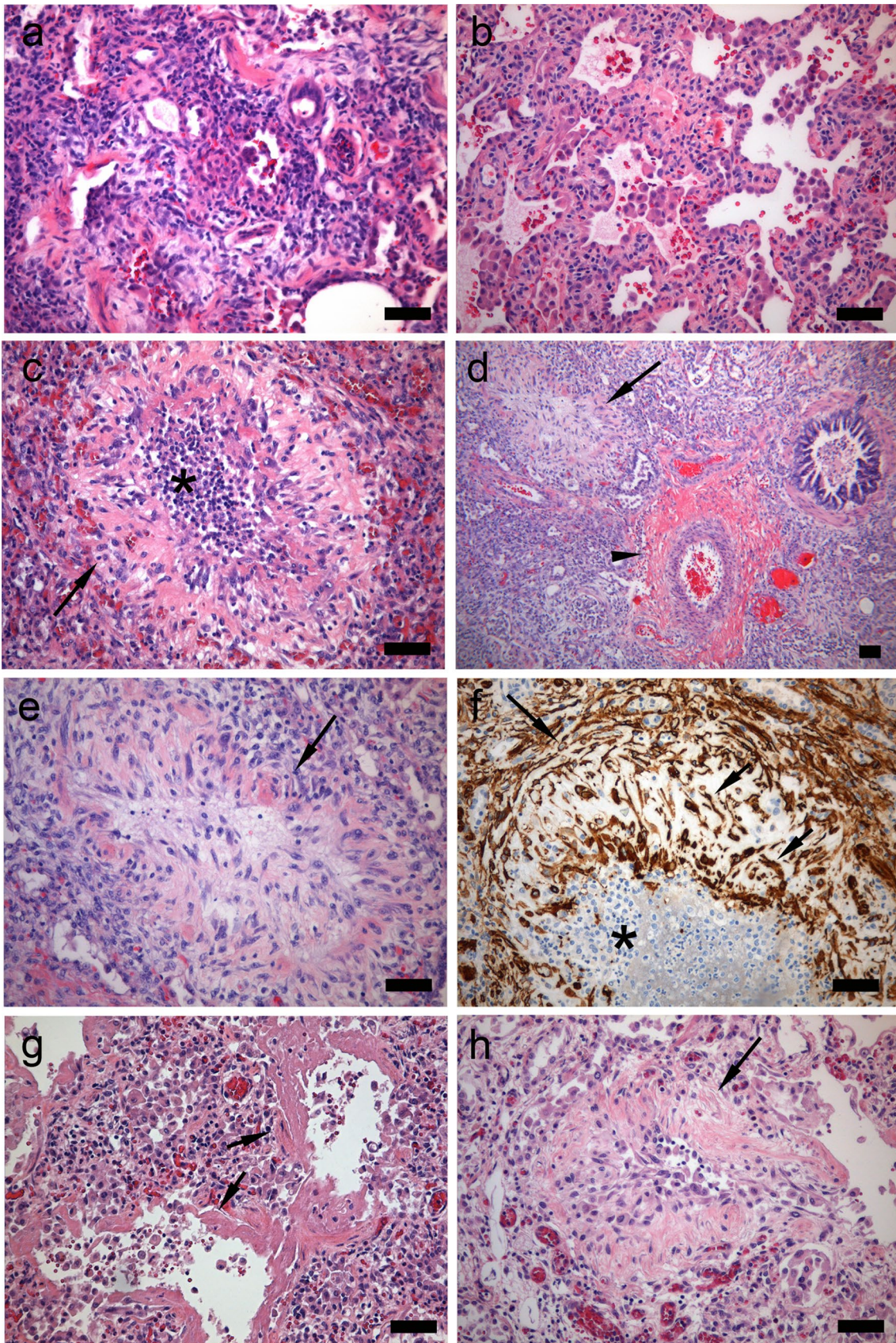


Fig. 2 Histological findings in lung samples from patients with NHLRC2 variants. **a, b** Interstitial fibrosis was present in the lung biopsy samples. **c–f, h** Granuloma-like lesions (arrows) within regions with variable amounts of necrosis (stars) were present in the autopsy samples from each patient. **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. 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, h**). Scale bars: 80 μ m

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 twofold 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 tenfold decrease in the levels of NHLRC2 in whole-cell extracts from patients compared to control fibroblasts (Fig. 4f, 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 (<http://www.brainspan.org>) [12], 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 (<http://www.brainspan.org>) [16], 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 dechoriation

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 h and the larvae display food seeking and avoidance behaviour 72 h after conception [11]. 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 [20]. A slight delay was observed in the timing of the dechoriation (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

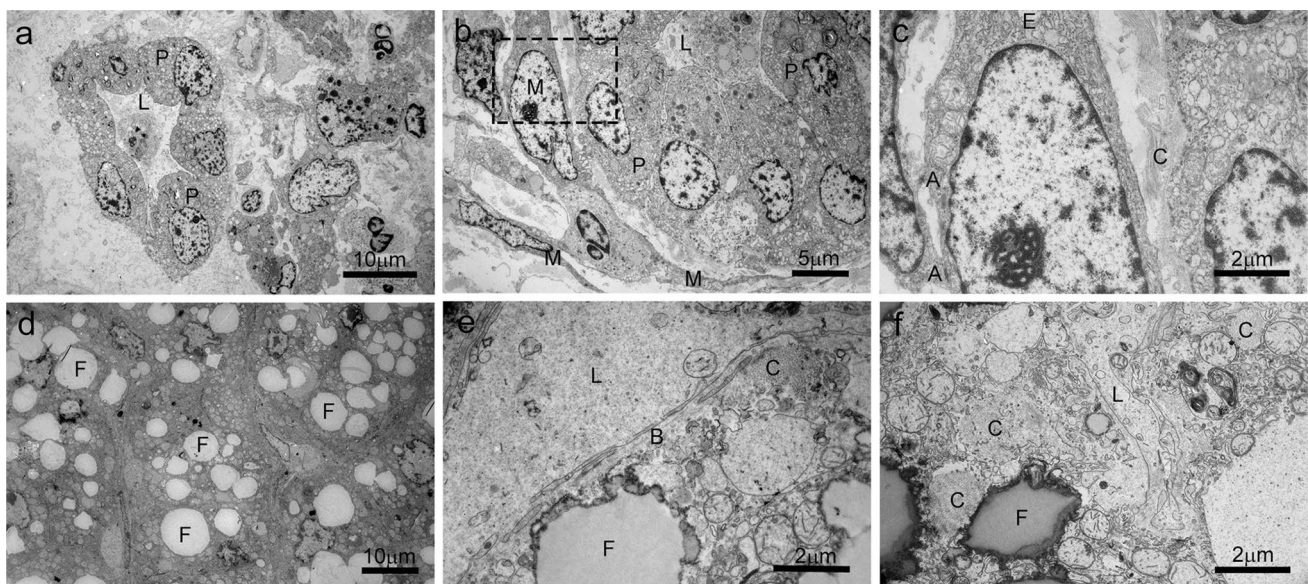


Fig. 3 Electron microscopy analysis of lung (a–c) and liver (d–f) autopsy samples. **a** Lung alveoli contained type II pneumocytes (P) and the alveolar space was very limited (L, the lumen of the alveolus). **b** In addition to type II pneumocytes (P), myofibroblasts (M) were present with diverging ‘limb-like’ projections. The square is delineating the area magnified in **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)

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 angiomas. 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 [17]. 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 [15] and non-redox functions. An example of a non-redox function occurs in SasA, a circadian clock-associated histidine kinase [28]. 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.

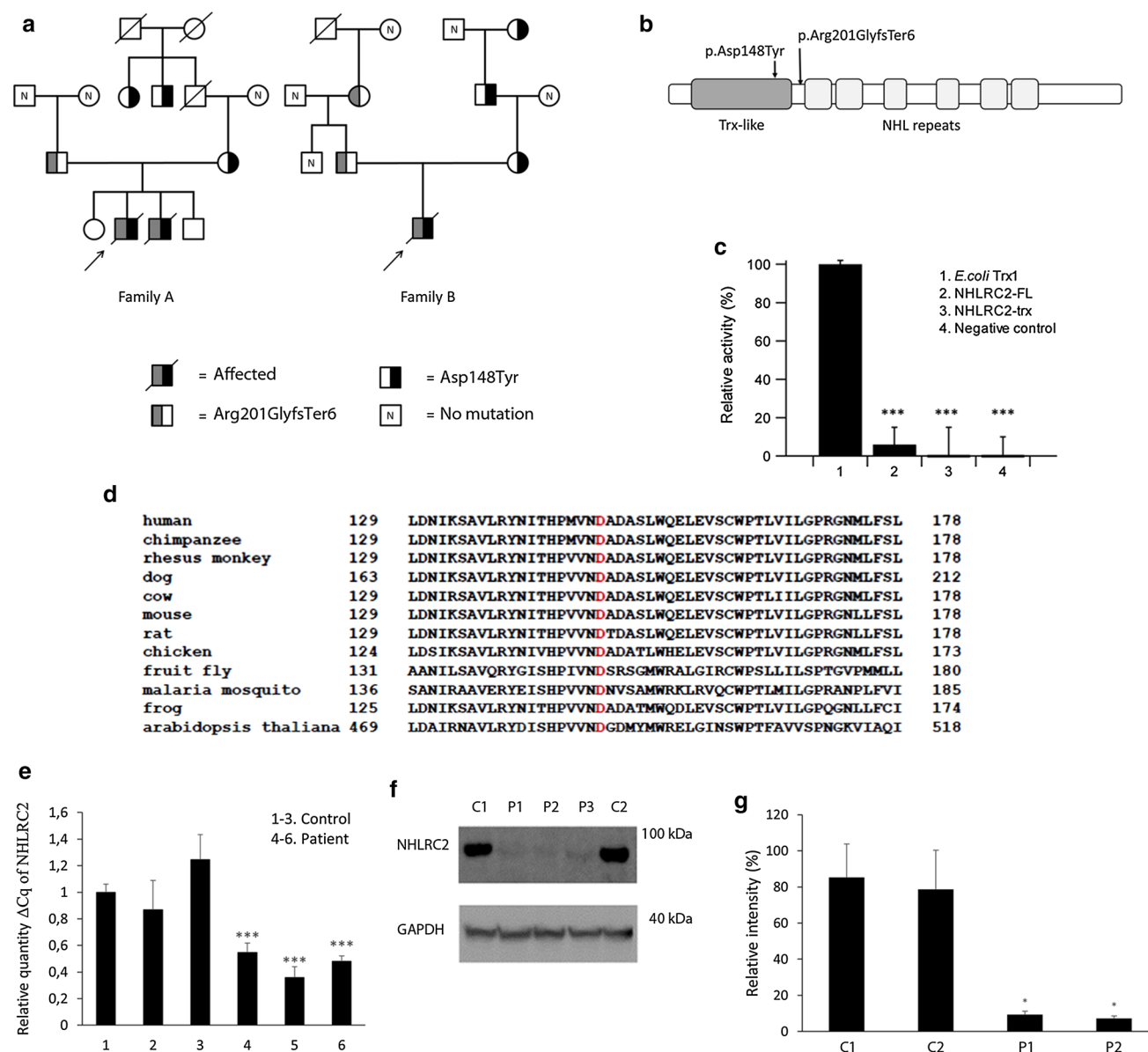


Fig. 4 Variants p.Asp148Tyr and p.Arg201GlyfsTer6 in NHLRC2 in patients with FINCA disease from two non-consanguineous families. **a** Pedigrees and the segregation of heterozygous variants p.Asp148Tyr and p.Arg201GlyfsTer6 in Families A and B. The probands of the families are indicated with an arrow 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's *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's *t*-test. Abbreviations: 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

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

Table 2 Genotyping summary of *Nhlrc2* knockout mouse offspring obtained through heterozygous breeding

Genotyping	WT (%)	Het (%)	Hom (%)	Total
P14 mice	32.0	68.0	0.0	154
E10.5 embryos	24.8	75.2	0.0	53
E2.5 embryos	31.1	68.9	0.0	45

χ^2 test showed significant deviation from Mendelian segregation ($p < 0.001$)

P14 2 weeks, E10.5 embryonic day 10.5, E2.5 embryonic day 2.5, WT wild-type, Het heterozygous, Hom homozygous

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 leptomenigeal, 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, which were enriched

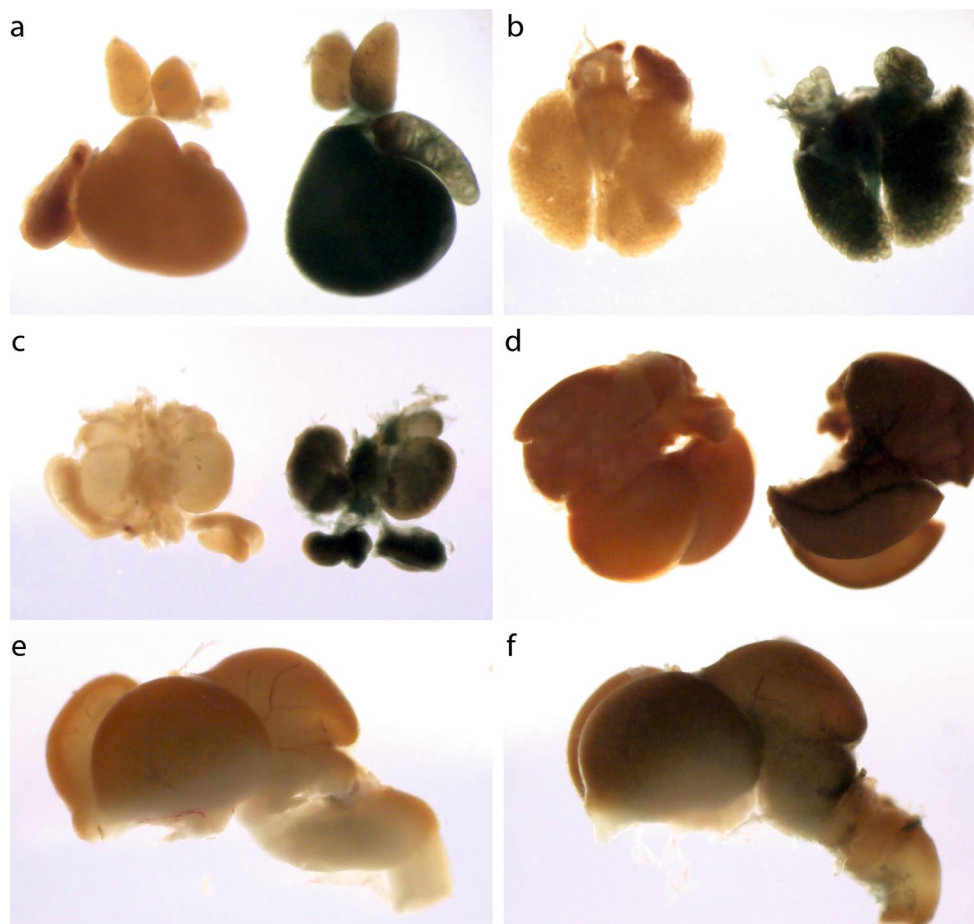


Fig. 5 LacZ mouse tissue staining. Heterozygote *Nhlrc2^{tm1a}* mice had a strong β -galactosidase signal throughout the body. Heart and thymus (a); lungs (b); kidneys, adrenal glands and gonads (c); liver (d)

and brain (e, f) (e WT; f heterozygote *Nhlrc2^{tm1a}* mice). WT littermates were negative for β -galactosidase signal

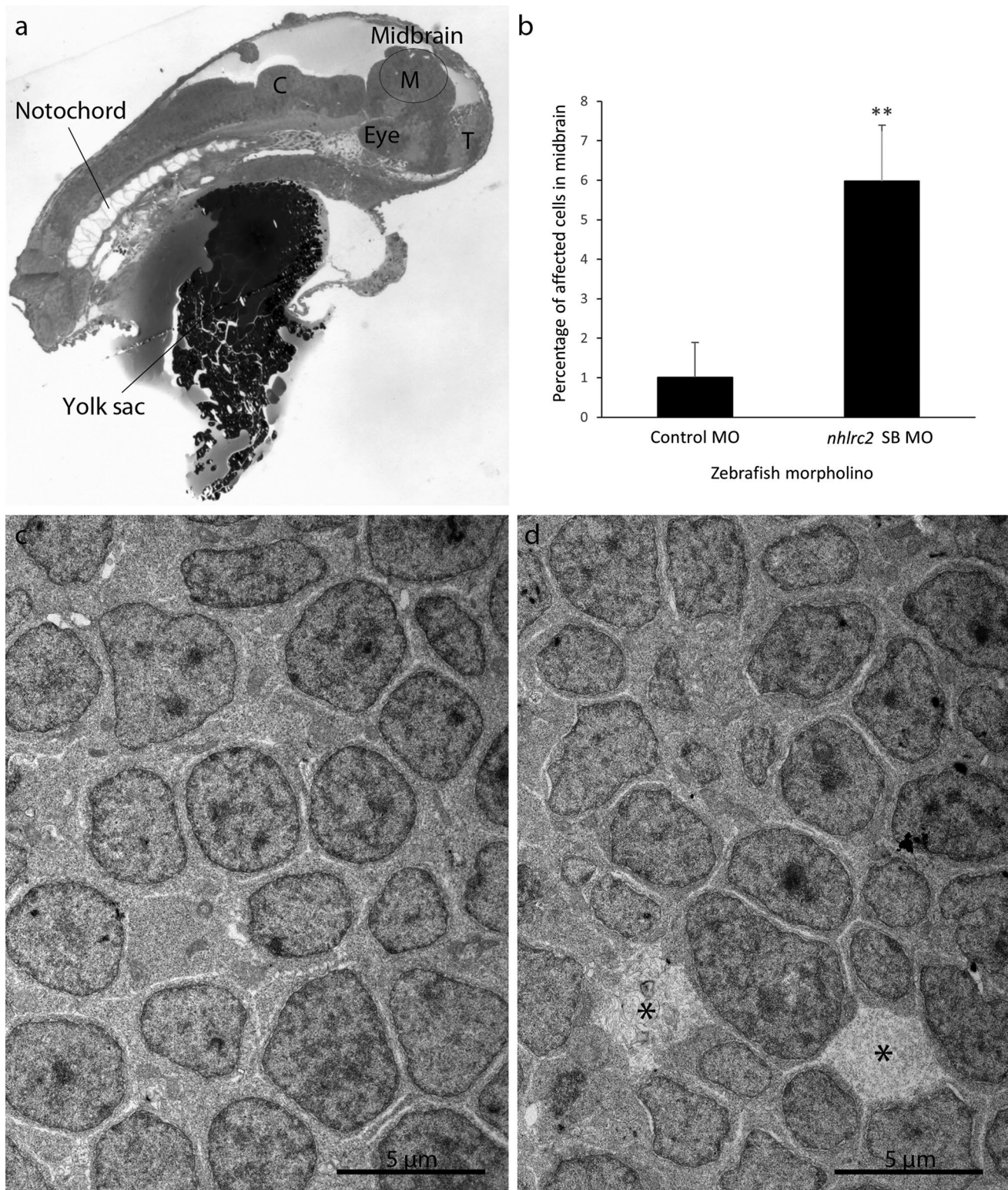


Fig. 6 Transmission electron microscopy from the midbrain region of the 2-dpf-zebrafish morphant. **a** A cross section of a zebrafish embryo at 2 dpf with a circled area specifying the exact location from which the electron micrographs were taken. **b** Cells with excess vacuolisation or disintegration 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: *C* cerebellum, *M* mesencephalon, *T* telencephalon, *SB* splice site-blocking, *MO* morpholino. Differences were considered as statistically very significant (**) with *p* value 0.001–0.01

with myofibroblasts that accumulate in areas surrounding granulomas in several granulomatous lung diseases such as sarcoidosis, atypical mycobacteriosis and tuberculosis [8]. 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) [30, 27]. 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 [27]. Neuroimaging in some NKX2-1 related cases has shown structural brain abnormalities including agenesis of the corpus callosum [19], and autopsies have revealed reduced numbers of striatal and neocortical interneurons consistent with a defect in neuronal migration [12]. In addition to RDS in neonates, pulmonary disease may present with interstitial lung disease in young children, and pulmonary fibrosis in older persons [19]. 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 (<http://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 angiomas. 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 [14]. Our expression analysis with LacZ-reporter mice suggests 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 (JU, decision number 138566; RH, decision numbers 266498, 273790 and 303996; MH, decision number 1126662; LR, decision numbers 266457 and 272573); the Sigrid Juselius Foundation (JU, RH and MH); the Foundation for Paediatric Research, Finland (JU and MKK); the Alma and KA Snellman Foundation (JU and MKK); a Marie Curie International Outgoing Fellowship of the European Union’s Seventh Framework Programme (Grant agreement number 273669 [BioMit]) (RH); Foundation of the Finnish Anti-Tuberculosis Association (RK); the Jane and Aatos Erkko Foundation (MR); 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 (JU); the National Heart, Lung and Blood Institute of the US National Institutes of Health under award number HL-54703 (LMN) and the Eudowood Foundation (LMN).

Compliance with ethical standards

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

(EETMK 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

- Bancroft JD, Stevens A (1991) Theory and practice of histological techniques. Wiley, Churchill Livingstone, Edinburgh
- Barabasi AL, Gulbahce N, Loscalzo J (2011) Network medicine: a network-based approach to human disease. *Nat Rev Genet* 12:56–68. <https://doi.org/10.1038/nrg2918>
- Bossy-Wetzell E, Schwarzenbacher R, Lipton SA (2004) Molecular pathways to neurodegeneration. *Nat Med* 10(Suppl):S2–S9. <https://doi.org/10.1038/nm1067>
- 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. <https://doi.org/10.1371/journal.pbio.1000582>
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>
- 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
- Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Finland (2016) Sequencing Initiative Suomi project (SISu). <http://sisuproject.fi>. Accessed 1/10/2017
- 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
- Karala AR, Ruddock LW (2010) Bacitracin is not a specific inhibitor of protein disulfide isomerase. *FEBS J* 277:2454–2462. <https://doi.org/10.1111/j.1742-4658.2010.07660.x>
- 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. <https://doi.org/10.1371/journal.pone.0051378>
- Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF (1995) Stages of embryonic development of the zebrafish. *Dev Dyn* 203:253–310. <https://doi.org/10.1002/aja.1002030302>
- 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. <https://doi.org/10.1002/ana.10637>
- 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. <https://doi.org/10.1038/nature19057>
- Long J, Pan G, Ifeachor E, Belshaw R, Li X (2016) Discovery of novel biomarkers for Alzheimer's disease from blood. *Dis Mark* 2016:4250480. <https://doi.org/10.1155/2016/4250480>
- Martin JL (1995) Thioredoxin—a fold for all reasons. *Structure* 3:245–250
- 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. <https://doi.org/10.1038/nature13185>
- 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–D221. <https://doi.org/10.1093/nar/gku1243>
- Nagy A, Gertsenstein M, Vintersten K, Behringer R (2003) Manipulating the mouse embryo, a laboratory manual, 3rd edn. Cold Spring Harbor Laboratory Press, New York, pp 198–200
- 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)
- Robu ME, Larson JD, Nasevicius A, Beiraghi S, Brenner C, Farber SA et al (2007) P53 activation by knockdown technologies. *PLoS Genet* 3:e78
- Santiago JA, Potashkin JA (2013) Integrative network analysis unveils convergent molecular pathways in Parkinson's disease and diabetes. *PLoS ONE* 8:e83940. <https://doi.org/10.1371/journal.pone.0083940>
- 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. <https://doi.org/10.1007/s11248-015-9897-1>
- Slack F, Ruvkun G (1998) Heterochronic genes in development and evolution. *Biol Bull* 195:375–376. <https://doi.org/10.2307/1543152>
- Spagnolo P, Grunewald J, du Bois RM (2014) Genetic determinants of pulmonary fibrosis: evolving concepts. *Lancet Respir Med* 2:416–428. [https://doi.org/10.1016/S2213-2600\(14\)70047-5](https://doi.org/10.1016/S2213-2600(14)70047-5)
- Sulonen AM, Ellonen P, Almusa H, Lepisto M, Eldfors S, Hanula S et al (2011) Comparison of solution-based exome capture methods for next generation sequencing. *Genome Biol* 12:R94. <https://doi.org/10.1186/gb-2011-12-9-r94>
- 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. <https://doi.org/10.1002/gene.20012>
- Thorwarth A, Schnittert-Hubener S, Schrupf 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. <https://doi.org/10.1136/jmedgenet-2013-102248>
- 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. <https://doi.org/10.1016/j.jmb.2004.07.010>
- 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
- de Vries BB, Arts WF, Breedveld GJ, Hoogbeem JJ, Niermeijer MF, Heutink P (2000) Benign hereditary chorea of early onset maps to chromosome 14q. *Am J Hum Genet* 66:136–142
- 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. <https://doi.org/10.1038/ng.2931>

32. Zhang Y, Chen K, Sloan SA, Bennett ML, Scholze AR, O'Keefe 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. <https://doi.org/10.1523/JNEUROSCI.1860-14.2014>

Affiliations

Johanna Uusimaa^{1,2,3} · Riitta Kaarteenaho^{4,5} · Teija Paakkola^{1,3} · Hannu Tuominen^{6,7} · Minna K. Karjalainen¹ · Javad Nadaf^{8,9} · Teppo Varilo¹⁰ · Meri Uusi-Mäkelä¹¹ · Maria Suo-Palosaari¹² · Ilkka Pietilä^{1,3} · Anniina E. Hiltunen^{1,3} · Lloyd Ruddock^{3,13} · Heli Alanen^{3,13} · Ekaterina Biterova^{3,13} · Ilkka Miinalainen³ · Annamari Salminen¹ · Raija Soininen^{3,13} · Aki Manninen^{3,13} · Raija Sormunen^{3,6} · Mika Kaakinen³ · Reetta Vuolteenaho³ · Riitta Herva⁶ · Päivi Vieira^{1,2} · Teija Dunder^{1,2} · Hannaleena Kokkonen^{14,15} · Jukka S. Moilanen^{1,16} · Heikki Rantala^{1,2} · Lawrence M. Noguee¹⁷ · Jacek Majewski⁸ · Mika Rämetsä^{1,2,11} · Mikko Hallman^{1,2} · Reetta Hinttala^{1,3}

✉ Johanna Uusimaa
johanna.uusimaa@oulu.fi

¹ PEDEGO Research Unit and Medical Research Center Oulu, University of Oulu and Oulu University Hospital, PO Box 5000, 90014 Oulu, Finland

² Department of Children and Adolescents, Oulu University Hospital, PO Box 23, 90029 Oulu, Finland

³ Biocenter Oulu, University of Oulu, PO Box 5000, 90014 Oulu, Finland

⁴ Research Unit of Internal Medicine, Respiratory Research, University of Oulu, PO Box 5000, 90014 Oulu, Finland

⁵ Medical Research Center Oulu and Unit of Internal Medicine and Respiratory Medicine, Oulu University Hospital, PO Box 20, 90029 Oulu, Finland

⁶ Department of Pathology, Cancer and Translational Medicine Research Unit, University of Oulu, PO Box 5000, 90014 Oulu, Finland

⁷ Department of Pathology, Oulu University Hospital, PO Box 50, 90029 Oulu, Finland

⁸ McGill University and Génome Québec Innovation Centre, Montreal, QC H3A 0G1, Canada

⁹ St. Jude Children's Research Hospital (SJCRH), 262 Danny Thomas Place, Memphis, TN 38105, USA

¹⁰ Department of Medical Genetics, University of Helsinki, Haartmaninkatu 8, 00251 Helsinki, Finland

¹¹ BioMediTech Institute and Faculty of Medicine and Life Sciences, University of Tampere, Tampere, Finland

¹² Department of Diagnostic Radiology and Medical Research Center Oulu, Oulu University Hospital and University of Oulu, PO Box 50, 90029 Oulu, Finland

¹³ Faculty of Biochemistry and Molecular Medicine, University of Oulu, PO Box 5000, 90014 Oulu, Finland

¹⁴ Northern Finland Laboratory Centre NordLab, Oulu University Hospital, PO Box 500, 90029 Oulu, Finland

¹⁵ Department of Clinical Chemistry and Medical Research Center Oulu, University Oulu and Oulu University Hospital, PO Box 5000, 90014 Oulu, Finland

¹⁶ Department of Clinical Genetics, Oulu University Hospital, PO Box 23, 90029 Oulu, Finland

¹⁷ Division of Neonatology, Johns Hopkins University School of Medicine, CMSC 6-104A, 600 N. Wolfe St., Baltimore, MD 21287, USA