Aus dem Institut für Zell- und Neurobiologie der
Charité – Universitätsmedizin Berlin

DISSERTATION

Effects of pre- and postnatal deletion of the transcription factor Nkx2-1 on the expression of *NGF*, *trkA*, *trkB* and $p75^{NTR}$ in mice and a clinical update on *NKX2-1* haploinsufficiency in humans

zur Erlangung des akademischen Grades
Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von

Sara Ersözlü

aus Solingen

Datum der Promotion: 04. Juni 2021

Preface

[The inner parts] are for the most part unknown—at least, those of man are, and hence [we] have to refer to those of other animals, the natural structure of whose parts those of man resemble, and examine them.

Aristotle, Historia Animalium, I, XVI

Partial results of the present dissertation have been previously published in:

1) *Magno L, Kretz O, Bert B, Ersözlü S, Vogt J, Fink H, Kimura S, Vogt A, Monyer H, Nitsch R, Naumann T. The integrity of cholinergic basal forebrain neurons depends on expression of Nkx2-1. Eur J Neurosci. 2011 Dec;34(11):1767-82.*

2) *Ersözlü S, Magno L, Naumann T. Effects of pre- and postnatal deletion of the transcription factor Nkx2-1 on the expression of NGF, trkA, trkB, and p75^{NTR} in mice. Abstract Book. 2013 July;6:45-46.*

   Abstract and oral presentation at ISC 2013 – International Student Congress, Medical University of Graz, Austria, July 4–6, 2013

   (Prize:"1st Place of all oral presentations")

3) *Ersözlü S, Magno L, Naumann T. Effects of pre- and postnatal deletion of the transcription factor Nkx2-1 on the expression of NGF, trkA, trkB, and p75^{NTR} in mice. BMC Proceedings. 9. A6-A6. 10.1186/1753-6561-9-S1-A6.*

   Abstract and oral presentation at ICHAMS 2013 – International Conference for Healthcare and Medical Students, Royal College of Surgeons in Ireland, Dublin, October 11–12, 2013

   (Prize: "1st Overall oral presentation winner")

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ABC	Avidine-biotine-peroxidase-complex
ACh	Acetylcholine
AChE	Acetylcholine esterase
ADHD	Attention deficit hyperactivity disorder
ANOVA	Analysis of variance
AOV	All other variants (i.e. not WGD or missense)
BDNF	Brain-derived neurotrophic factor
BF	Basal forebrain
BHC	Benign hereditary chorea
β-gal	β-galactosidase
C57Bl/6	Black 6 Inbred Mouse Strain
CB	Calbindin
cf.	Confer (english: compare)
CGE	Caudal ganglionic eminence
ChAT	Choline acetytransferase
ΔCOOH	Deletion of the transactivation domain 2 (c.753_*881del)
CPu	Caudate putamen (striatum)
CT	Computed tomography
C-TAD	C-terminal transactivation domain
DAB	Diaminobenizidine
Domain I	Inhibitory domain
E	Embryonic day
ECMO	Extracorporeal membrane oxygenation
EDC	Ethyl cysteinate dimer
e.g.	Exempli gratia (english: for example)
ERK	Extracellular regulated kinase
FFS	Fimbria-fornix-system
FFT	Fimbria-fornix-transection
FDG-PET	Fluorodeoxyglucose positron emission tomography
GABA	Gamma-aminobutyric acid
GAD67	Glutamate-decarboxylase isoform 67
GAP-43	Growth associated protein 43

GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
HD	Homeodomain
hDB-SI	Horizontal limb of the diagonal band-substantia innominata
hg18/19	Human Genome Browser Gateway 18/19
HGVS	Human Genome Variation Society
HPRT	Hypoxanthine-guanine phosphoribosyltransferase
i.e.	Id est (english: that is)
IHC	Immunohistochemical
ILD	Interstitial lung disease
IQ	Intelligence quotient
IQR	Interquartile ranges
IRDS	Infant respiratory distress syndrome
Isl1	LIM homeobox 1
Ldb 1	LIM domain-binding protein 1
LGE	Lateral ganglionic eminence
LGP	Lateral globus pallidus
Lhx6/7, LHX6/7	LIM homeobox 6/7
LS	Lateral septum
M	mole per liter
MGE	Medial ganglionic eminence
MRI	Magnetic resonance imaging
MS	Medial septum
MSvDB	Medial septum-vertical limb of the diagonal band of Broca
n	Number
ΔNH2	Deletion of the transactivation domain 1 (c.-96_565del)
NEHI	Neuroendocrine hyperplasia of infancy
NGF	Nerve growth factor
NK2-SD	NK2-specific domain
Non-HD	Non-homeodomain variant (not affecting the homeodomain)
Non-HD+	Non-homeodomain variant affecting the homeodomain
Non-WGD	All variants that are not whole gene deletions
NPY	Neuropeptide Y
N-TAD	N-terminal transactivation domain

p75^{NTR}	*p75* neurotrophin receptor
P	Postnatal day
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PET	Positron emission tomography
PFA	Paraformaldehyde
POA	Preoptic area
pro-SP-A/-B/-C/-D	Pro-surfactant protein A/B/C/D
PV	Parvalbumin
qRT-PCR	Quantitative real-time polymerase chain reaction
RT	Room temperature
SHS	Septohippocampal system
SP-A/-B/-C/-D	Surfactant protein A/B/C/D
SPECT	Single photon emission computed tomography
SST	Somatostatin
Ta	Temperature
TAZ	Tafazzin
T/EBP	Thyroid-specific enhancer-binding-protein (synonym for *NKX2-1*)
TITF-1	Thyroid transcription factor 1 (synonym for *NKX2-1*)
TG	Thyreoglobulin
TN	Tinman-domain
trkA/B	Tyrosine kinase receptor A/B
TTF-1 (=TTF1)	Thyroid transcription factor 1 (synonym for *NKX2-1*)
VAChT	Vesicular acetylcholine transporter
vs.	Versus
WGD	Whole gene deletions
X-gal	5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside
Zic4	Zic family member 4

Abstract

Background: Expression of the transcription factor Nkx2-1 is essential for forebrain, lung, and thyroid development in mice. Humans with *NKX2-1* haploinsufficiency suffer from a rare condition ("brain-lung-thyroid syndrome") with various combinations of organ manifestations and no known genotype-phenotype correlation. There is typically a triad of choreoathetosis, infant respiratory distress syndrome, and congenital hypothyroidism. Especially the neurological phenotype varies strongly between patients, and little is known about the impaired cell types. In recent years, cholinergic and parvalbumin-containing (PV) gamma-aminobutyric acid (GABA)ergic neurons of the basal ganglia and magnocellular nuclei of the basal forebrain (BF) have been found to maintain Nkx2-1 synthesis throughout life[1,2], and their deletion to cause motor deficits and impairments of spatial memory and learning[3]. Cholinergic BF neurons are strongly influenced by the "nerve growth factor" *(NGF)*, which is mainly synthesized by cortical GABAergic neurons[4,5]. They undergo shrinkage and degeneration without access to *NGF*[6]. To elucidate whether Nkx2-1 deficiency is accompanied by effects on the neurotrophin system, we investigate the expression levels of *NGF*, certain neurotrophin receptors, and the axonal sprouting marker growth-associated protein 43 (*GAP-43)* in conditional knockout mice with pre- or postnatal Nkx2-1 deletion. Additionally, a comprehensive literature review and genotype-phenotype correlation analysis of human cases is conducted to compare the results in mice.

Methods: Expression levels of *NGF, $p75^{NTR}$, trkA, trkB*, *GAP-43,* and the housekeeping genes *HPRT* and *GAPDH* were measured using quantitative real-time PCR. Therefore mice with prenatal mutation (*$GAD^{cre/+}$;$Nkx2\text{-}1^{c/c}$*), postnatal mutation (*$ChAT^{cre/+}$;$Nkx2\text{-}1^{c/c}$*) and corresponding controls were used (each n = 5). Moreover, the Cre-recombination in cholinergic neurons of the postnatal line was monitored using ROSA-26 reporter mice. The study was carried out according to institutional guidelines for animal welfare. Additionally, a literature review of human cases (1998 to May 2nd 2020) was performed and a disease severity score was developed and studied.

Results: Significantly decreased expression levels of *trkA* (-82%) and *$p75^{NTR}$* (-60%) in GAD67-mutants, and *trkA* (-56%) in ChAT-mutants were detected. The analysis of 351 human cases revealed a predominant neurological phenotype, an early lung-associated mortality, and certain genotype-phenotype correlations.

Conclusions: In summary the study demonstrates that: (1) the loss of cholinergic neurons after pre- and postnatal Nkx2-1 deletion is accompanied by a corresponding loss of *trkA*; (2) local axonal sprouting and cortical GABAergic *NGF* synthesis are not impaired; (3) similarly to mice,

cognitive impairments play a role in humans; and (4) evidence for certain genotype-phenotype correlations exists for disease manifestations and severity. (400 words)

Abstract (German)

Hintergrund: Die Expression des Transkriptionsfaktors Nkx2-1 ist essentiell für die Entwicklung des Vorderhirns, der Lungen und der Schilddrüse in Mäusen. Menschen mit *NKX2-1*-Haploinsuffizienz leiden an einer seltenen Erkrankung („Gehirn-Lungen-Schilddrüsen-Syndrom") mit verschiedenen Organmanifestationen und ohne bekannte Genotyp-Phänotyp-Korrelation. Typischerweise findet sich eine Symptomtrias aus Choreoathetose, Neugeborenen-Atemnotsyndrom und kongenitaler Hypothyreose. Insbesondere der neurologische Phänotyp unterscheidet sich stark zwischen den Patienten und es ist nur wenig bekannt über die betroffenen Neuronenpopulationen. In den letzten Jahren wurden im Mäusegehirn cholinerge und Parvalbumin(PV)-positive Gamma-Aminobuttersäure (GABA)erge Neurone der Basalganglien sowie des magnozellulären basalen Vorderhirns (BF) als Nkx2-1-positiv identifiziert und es wurde nachgewiesen, dass diese lebenslang Nkx2-1 exprimieren[1,2] sowie, dass ihr Verlust im Gehirn zu motorischen Defiziten und Beeinträchtigungen des Lernens und des Gedächtnisses führt[3]. Cholinerge Neurone des BF werden stark vom „Nervenwachstumsfaktor" (*NGF)* beeinflusst, der vor allem von GABAergen kortikalen Neuronen gebildet wird[4,5]. Sie schrumpfen und degenerieren ohne Zugang zu *NGF*[6]. Um herauszufinden, ob ein Nkx2-1-Defizit mit einer Beeinträchtigung des neurotrophen Systems einhergeht, haben wir die Expressionslevel von *NGF* und bestimmten Neurotrophinrezeptoren sowie dem axonalen Sproutingmarker Wachstums-assoziiertes Protein 43 (*GAP-43*) in konditionalen Mausmutanten mit prä- oder postnataler Nkx2-1-Deletion untersucht. Zusätzlich wurde eine ausführliche Literaturanalyse und Genotyp-Phänotyp-Korrelationsanalyse bei menschlichen Fallberichten durchgeführt, um die Ergebnisse in Mäusen zu vergleichen.

Methoden: Die Expressionslevel von *NGF, $p75^{NTR}$, trkA, trkB*, *GAP-43* und der Housekeeping-Gene *HPRT* und *GAPDH* wurden mit quantitativer real-time-PCR gemessen. Für die Analyse wurden Mutanten mit pränataler Mutation (*$GAD^{cre/+}$;$Nkx2\text{-}1^{c/c}$*) und postnataler Mutation (*$ChAT^{cre/+}$;$Nkx2\text{-}1^{c/c}$*) sowie entsprechende Kontrollen (jeweils n = 5) verwendet. Zusätzlich wurde die Cre-Rekombination in cholinergen Neuronen der postnatalen Mauslinie unter Venwendung von ROSA-26-Reportermäusen überprüft. Die Experimente erfolgten gemäß den Tierschutzbestimmungen der Forschungsinstitution. Außerdem wurde eine Literaturanalyse menschlicher Fallberichte (1998 bis 2. Mai 2020) durchgeführt sowie ein Score für den Schweregrad der Erkrankung entwickelt und untersucht.

Ergebnisse: Eine signifikant reduzierte Expression ließ sich für *trkA* (-82%) und *$p75^{NTR}$* (-60%) in GAD67-Mutanten und für *trkA* (-56%) in ChAT-Mutanten nachweisen. Die Analyse von 351

menschlichen Fallberichten zeigte einen vordergründig neurologischen Phänotyp, eine frühe lungenassoziierte Mortalität und gewisse Genotyp-Phänotyp-Korrelationen.

Schlussfolgerungen: Zusammenfassend zeigt die Studie, dass (1) der Verlust cholinerger Neurone infolge prä- und postnataler Deletion von Nkx2-1 von einer Expressionsminderung von *trkA* begleitet wird, (2) lokales axonales Sprouting und die kortikale GABAerge *NGF*-Synthese nicht beeinträchtigt werden, (3) wie in Mäusen, kognitive Beeinträchtigungen auch bei Menschen eine Rolle spielen und (4) sich Beweise für das Vorhandensein gewisser Genotyp-Phänotyp-Korrelationen für die Krankheitsmanifestationen und den Schweregrad finden. (392 Worte)

1. Introduction

1.1 Nkx2-1 is essential for brain, lung and thyroid gland development

Nkx2-1* is a homeobox transcription factor belonging to the NK-2 gene family[7], first discovered and characterized in the fruit fly, *Drosophilia melanogaster*, by Kim and Nierenberg[8]. Nkx2-1 is highly conserved among mammalian species (98-99%)[7] and located in syntenic regions of the human and murine genome[7], namely, on chromosome 14 in humans[7,9] (region 14q13, *genomic coordinates according to Enseml[10] for hg18 chr14:36.516.392_36.521.149 and for hg19: chr14:36.985.602_36.990.354; last accessed 01.11.2020*) and on chromosome 12C in mice[11]. For the human *NKX2-1* gene, two major transcript isoforms have been described: isoform/transcript 1 (NM_001079668.2: three exons, 401 amino acids long) and isoform/transcript 2 (NM_003317.3: two exons, 371 amino acids long)[12]. Isoform 1, which is consistent with the nomenclature of the Human Genome Variation Society (HGVS), is the one that genetic testing laboratories mainly use as a reference sequence when screening for variants[13,14]; therefore, it is also used in the present study. A schematic drawing of the human isoform 1 illustrated in Figure 1, which contains the most current knowledge on the homeodomain (HD) and other different, rarely reported domains of the *NKX2-1* gene that were identified from different studies and genome databases for the purpose of this study. The HD is responsible for DNA recognition and nuclear localization of the protein[15,16], but there is no consensus on its exact location within the *NKX2-1* gene[15-17]. The two sequences that can be found for isoform 1 in two widely utilized genome databases are of close proximity: c.571-750[18] (used in our study as a reference) and c.565-759[19]. Furthermore, two independent transcriptional transactivation domains – the

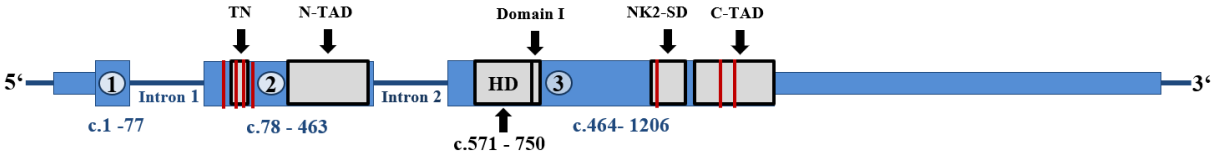

Figure 1. The human NKX2-1 isoform 1. *Three exons, the homeodomain (HD), two introns and NKX2-1-specific domains and sites are illustrated. The coordinates correspond to the Human Genome Browser Gateway 19 (hg19; https://genome.ucsc.edu/cgi-bin/hgGateway?db=hg19) for isoform 1 (NM_001079668.2). The exons and introns are shown in blue[21] and exons are indicated by circles containing the numbers 1–3: exon 1 (c.1-77;p.1-26), exon 2 (c.78-463;p.26-155), and exon 3 (c.464-1206;p.155-402). The domains are marked as gray boxes, and their names are either provided inside the boxes or indicated by labels and black arrows above them: the Tinman domain (TN, c.115-150;p.39_50)[22], the N-terminal activation domain (N-TAD, c.241_459;p.81_153)[20], the HD (c.571-750;p.191-250)[18], the inhibitory domain (Domain I, c.751_777;p.251_259)[23], the NK2-specific domain (NK2-SD, c.844-912;p.282_304)[22] and the C-terminal activation domain (C-TAD, c.973-1203;p.325-304)[20]. The seven phosphorylation sites are presented as thin red bars at the following base pair positions: c.100_102;p.34[28], c.124_126; p.42[28], c.142_144;p.48[28], c.157_159;p.53[18,24,28], c.850_852;p.284[28], c.1069_1071; p.357[28] and c.1096;p.366[28].*

* The transcription factor is written in small letters when referring to animals. When addressing the human gene it is written in capital and italic letters.

the N-terminal transactivation domain (N-TAD) and the C-terminal transactivation domain (C-TAD)[20] – exist. In addition, three other conserved domains within the NK2-gene class have been described: the Tinman domain (TN)[22], which has been discussed as a potential promoter-specific transcriptional activator or inhibitor[23]; the NK2-specific domain (NK2-SD)[22,25-27], which is suggested to act as an accessory DNA-binding domain and to be involved in protein-protein interactions[22,23,25,28] without playing a role in the specificity or affinity of DNA-binding[25,28]; and the inhibitory domain (Domain I) adjacent to the HD with unknown function[23]. While DNA-binding has been reported to be phosphorylation-independent in mice[24], seven phosphorylation sites have been discovered that have been found to play a role in posttranslational modifications, at least in mice[24,29]. Nkx2-1 generally encodes certain proteins that bind DNA and thereby regulate gene expression during development. For instance, the LIM homeobox genes 6 and 7 (Lhx6 and Lhx7)[30] and their LIM domain-binding protein 1 (Ldb-1) have been found to act downstream of Nkx2-1[31].

The Nkx2-1 gene has also been found to be essential for the morphogenesis and differentiation of the brain[32], lungs[33], and thyroid gland of vertebrates[34], where it is differentially regulated[24]. It is also expressed in various other parts of the body (e.g. parathyroid, skin and enteric ganglia) and involved in the neuroendocrine and immune system, calcium homeostasis, the regulation of the circadian rhythm, food intake, and osmoregulation[35] (see Malt et al., 2016[35] for an extensive review of cell types with Nkx2-1 expression).

In mice, Nkx2-1 can be detected as early as Embryonic Day 10 (E10; see Figure 2) in the developing telencephalon and from E11 – E13 onwards in the fetal bronchial epithelium and the thyroid anlage[32]. In the developing mouse brain, Nkx2-1 is essential for the differentiation and migration of various neuronal populations of the basal forebrain (BF), including the basal ganglia, pituitary gland, and hypothalamus[36,37]. In the lungs, Nkx2-1 is necessary for pulmonary branching, alveolarization, and surfactant production[9,38] and the developing thyroid gland requires it for the differentiation of thyrocytes and for the establishment and functional maintenance of their follicular architecture[39,40]. Due to its initial isolation from the thyroid gland, Nkx2-1 is also widely known as the "thyroid transcription factor 1" (TTF1 and TTF-1)[41] or the "thyroid-specific enhancer-binding-protein" (T/EBP)[42].

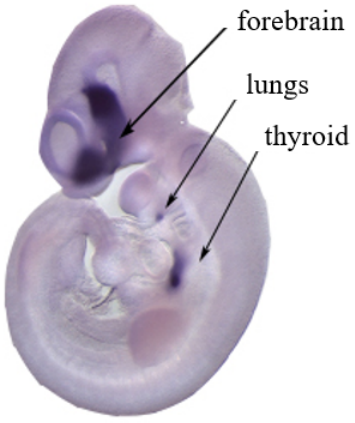

Figure 2. Nkx2-1 expression in a mouse embryo at Embryonic Day 10 (E10)*. Whole mount in-situ hybridization of a mouse embryo at E10 is shown. Nkx2-1 expression (violet staining) is visible in the forebrain, lungs, and thyroid gland* ***(Schrumpf P., Charité Berlin).***

Evidence of the critical role of the Nkx2-1 gene during the development of these three organs was provided in homozygous Nkx2-1$^{-/-}$ null mice that had severely altered basal ganglia, had a rudimentary thyroid gland, and were stillborn due to pulmonary agenesis[32,43]. Two other studies have provided further evidence of the role of the functional domains and phosphorylation sites for proper organogenesis in mice. Using three different mutant mouse lines (see Figure 3 for details), Silberschmidt et al. (2011)[32,291] demonstrated that partial deletion of Nkx2-1 can also result in organ agenesis. They found that phosphorylation is essential for lung development in homozygous mice with simultaneous mutation of the seven phosphorylation sites that had severe lung dysgenesis leading to early postnatal death[44]. They further showed that two different regions that do not involve the HD are equally essential for lung development, as homozygous mice with deletion of either of these, namely the transactivation domain 1 (also referred

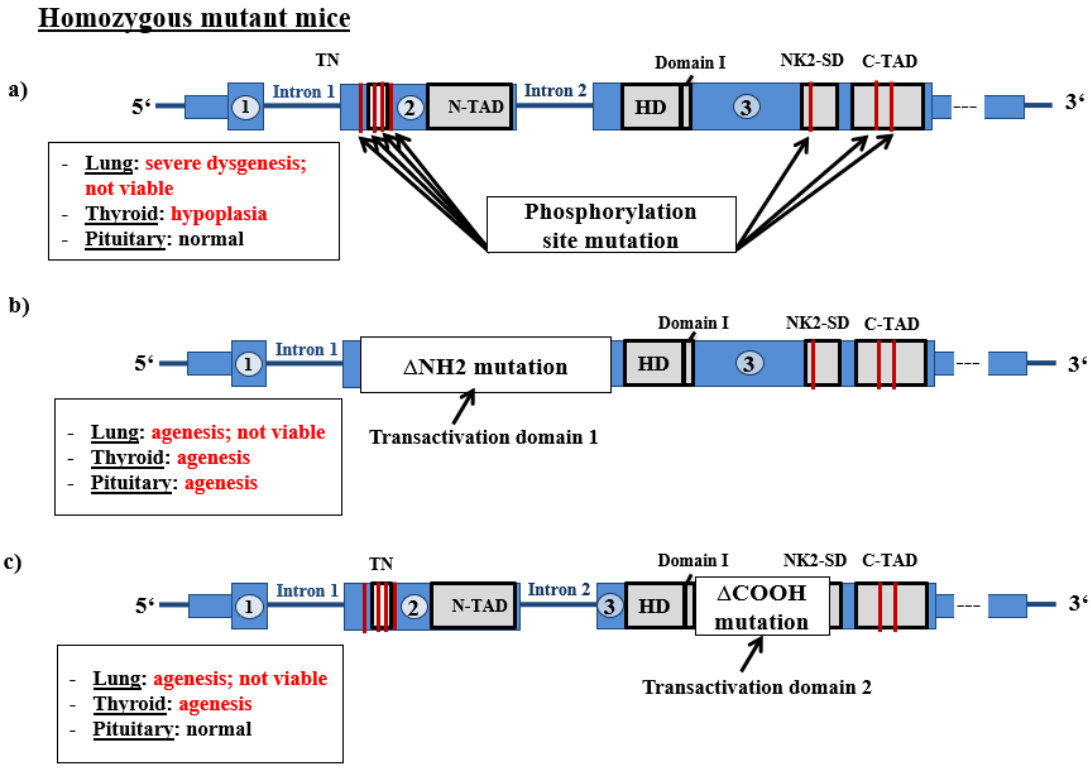

Figure 3. Effects of three different mutations of Nkx2-1 in mice on lung, thyroid and pituitary development. *The schematic shows the three mutant mouse lines with different mutations of Nkx2-1 investigated by Silberschmidt et al. (2011)[32,291]. The Nkx2-1 gene, with its exons (indicated by circles with the numbers 1–3) and introns, different domains (gray boxes), and seven phosphorylation sites (red bars), is illustrated: C-TAD = C-terminal transactivation domain, Domain I = inhibitory domain, HD = homeodomain, NK2-SD = NK2-specific domain, N-TAD = N-terminal transactivation domain, and TN = Tinman domain. The deleted regions of the Nkx2-1 gene are presented for the three mutations by arrows or white bars: a) selective mutation of the phosphorylation sites at serine residues (indicated by arrows); b) the ΔNH2-mutation with a deletion of transactivation domain 1 before the HD (approximate translated location for isoform 1: c.96_565del) encompassing the TN, the phosphorylation sites c.100_102;p.34, c.124_126;p.42, c.142_144;p.48, and c.157_159;p.53 and the N-TAD; and c) the ΔCOOH-mutation with deletion of transactivation domain 2 after the HD (approximate translated location for isoform 1:c.753_*881del) encompassing the NK2-SD and the phosphorylation site c.850_852;p.284. The main results for homozygous mutations are indicated on the left side in the white boxes. The information for a)–c) was obtained from* ***Silberschmidt et al. (2011)***[32,291] *and additional information about the lung manifestations for a) was obtained from from* ***De Felice et al. (2003)***[24].

to as ΔNH2 mutation; approximate translation to isoform 1: c.96_565del; comprising the TN, the phosphorylation sites c.100_102;p.34, c.124_126;p.42, c.142_144;p.48, c.157_159;p.53 and the N-TAD) or the transactivation domain 2 (also referred to as ΔCOOH mutation; approximate translation to isoform 1: c.753_*881del), have lung agenesis[29]. Similarly, the deletion of either the transactivation domain 1 or 2 caused thyroid agenesis in homozygous mice. In the brain, only the pituitary gland was studied, and, in contrast to the lungs and thyroid, only homozygous deletion of the transactivation domain 1 lead to agenesis (Figure 3). Compared to homozygous mice, heterozygous mutants survived in all of the above-mentioned studies (Nkx2-1 deletion[32,45], phosphorylation site mutation, ΔNH2 deletion and ΔCOOH deletion[24,32,291]).

1.2 *NKX2-1* haploinsufficiency in humans – the "brain-lung-thyroid syndrome"

In humans, a rare genetic condition known as *NKX2-1* haploinsufficiency[46] gives rise to the so-called "brain-lung-thyroid syndrome." It was first termed by Willemsen et al. in 2005[47] (OMIM Database[48] accession number #610978; ORPHA Database[49] accession number #209905) and typically causes a triad of symptoms (choreoathetosis, infant respiratory distress syndrome [IRDS], and congenital hypothyroidism) involving the name-giving three organs[46]. Already in the era before the molecular basis of this syndrome was known, a familial phenotype of "benign hereditary chorea" (BHC) was first described in 1966[50]. The prevalence of BHC was estimated to be 1:500,000 in 1978[51,52]; however more recent data are not available[50]. Later in 2000, the genetic region 14q12–22 with the markers D14S49–D14S1064 was suggested to be its genetic locus[53].

BHC is widely regarded as the hallmark of the disease[50]; and therefore, the term is often used to describe patients with *NKX2-1* variants[54-56] independent of the disease manifestations, rather than "*NKX2-1* haploinsufficiency" or the "brain-lung-thyroid syndrome." In an attempt to unify the terminology Patel et al.[50] proposed the use of "*NKX2-1* related disorders" instead. However, in this study, I proceed with the well-established terms "*NKX2-1* haploinsufficiency" and the "brain-lung-thyroid syndrome," as I find them to aptly describe the genetic and clinical condition that I regard as a syndrome, and not a variety of different disorders that can be associated with *NKX2-1* variants.

On a genetic level, "haploinsufficiency" refers to a heterozygous state, where the absence or inactivation of one wild-type allele leads to a deficiency in the gene product that cannot be fully compensated for by the remaining allele, due to semi-dominance of the affected gene[57,58]. In fact, *NKX2-1* has been found to have a particularly high DECIPHER haploinsufficiency score (1.49; 0–10% indicates a high probability of a gene be haploinsufficient)[59].

Overall, knowledge of the brain-lung-thyroid syndrome is scarce, especially regarding the mechanisms underlying this puzzling condition. A rise in the number of case reports can be noted in recent years describing a great variability in manifestations among patients with the same or different variants, extending the typical triad. However, the level of information provided about the patients in the case reports, both on a basic level (i.e. age, sex, family history) and pertaining to specific aspects of the syndrome (i.e. thyroid function parameters, brain imaging results, genetic analysis), is heterogeneous.

So far, only a few reviews on the syndrome exist[50,54,56,60]. Among them, the one that serves as the major reference, especially in terms of statistics, is by Carré et al. (2009)[59,60]. The numbers provided by them also constitute the basis for Patel et al.'s most recent National Center for Biotechnology Information[21] GeneReviews®[†] article[50] (2014). Carré et al. analyzed 46 patients from 19 publications (28 families with 27 different variants) and found a triad of symptoms in 50% of cases. Other organ phenotypes were combinations of the brain and thyroid in 30% and brain and lung in 4%, or isolated brain (13%) or isolated thyroid (2%) manifestations. There were no isolated lung or lung/thyroid phenotypes.

The overwhelming majority had neurological symptoms (93%), including chorea, muscular hypotonia, microcephaly, and developmental delay. Brain magnetic resonance imaging (MRI) was performed in 85% of cases, and pathologies were found in only 18% of these patients, ranging from cerebral atrophy, dysgenesis of the basal ganglia, and cystic masses in the sella turcica, to hyperdensities of the cerebellum, and agenesis of the corpus callosum. Other than developmental delay, no cognitive impairments were reported[60]. The second most frequently affected organ was the thyroid (87%), with findings ranging from congenital hypothyroidism, subclinical hypothyroidism, and thyroid hypoplasia to athyrosis. Pulmonary manifestations were found in 54%, including IRDS, frequent pulmonary infections, interstitial lung disease (ILD) and lung carcinoma. The mortality rate in their study was 7%, with lethal cases only observed among patients with respiratory manifestations. Additional clinical features, such as growth retardation, oligo-/hypodontia, and facial dysmorphism were rarely reported.

In subsequent years, some authors of larger case series[12,34-38] reviewed the literature, mainly to compare their own results, and they provided brief statistical updates on certain aspects of the syndrome[14,26,61-64]. For instance, Gras et al.[61] (2012; 124 patients, 32 families) found a higher

[†] *GeneReviews* is an international point-of-care resource providing clinically relevant and medically actionable information for inherited conditions in a standardized journal-style format covering diagnosis, management, and genetic counseling for patients and their families (https://www.ncbi.nlm.nih.gov/books/NBK1116).

frequency of females (ratio 0.7) and neurological manifestations (97%), with a mean age of 2 years at onset of chorea. The distribution of organ phenotypes in their study was comparable with Carré et al.[60], with the exceptions of a higher incidence of isolated brain phenotypes (22%), the presence of learning difficulties (25%), and severe intellectual disability (13%). Most authors found chorea to progress until adolescence and then either remain stable or improve[50,61,65].

The mode of inheritance in this disorder is autosomal dominant[50], with different rates of familial and *de novo* variants reported in the literature. Several authors have found a higher rate of *de novo* variants[60,62] compared to familial cases. Whilst Gras et al.[61] (2012; analysis of 124 patients from the literature), Nettore et al.[64] (2013; analysis of 37 variants from the literature), and Thorwarth et al.[26] (2014; analysis of 161 patients with 77 variants from the literature) could not find a genotype-phenotype correlation, Inzelberg et al.[54] (2011) and Peall et al.[56,63] (2014; 2015) described whole gene deletions (WGD) to be associated with severe phenotypes. Similarly, Hamvas et al. (2013) identified severe respiratory phenotypes in patients with large deletions in their own case series with 16 patients[13]. In addition, Veneziano et al. observed an association of cystic pituitary masses and abnormalities of the sella turcica with large WGD, as well as missense and nonsense variants[62]. Meanwhile, other authors have observed milder phenotypes in patients with missense variants[50,61,66]. It should be noted that no classification of mild or severe phenotypes exists for this disorder and that the terminology was thus based on the assessment of the clinicians. The occasional finding of hypo- and oligodontia in these patients was linked to concomitant variants of the *PAX9* gene[14,67-69].

In summary, numerous aspects of this syndrome are not well studied and understood yet. In particular, the question of whether there is a genotype-phenotype correlation, which is of great significance for clinicians, remains unclear. To date, all of the available reviews have excluded certain articles from their analyses for declared (e.g. non peer-reviewed articles and/-or non-English articles[56], articles lacking information on lung and/ or thyroid manifestations[54,60]) or undeclared reasons[61,62,64]. This makes it difficult to obtain a comprehensive overview of all the available data on this rare condition. Since Nkx2-1 exerts its main role in the BF, which is a highly heterogeneous region partly lacking a precise definition, knowledge of this brain region is essential for the purpose of this study. Therefore, the following chapter is devoted to the functions of the BF, its anatomy and development, and finally the expression pattern of Nkx2-1 in mice.

1.3 Expression of Nkx2-1 in the basal forebrain (BF) of mice

1.3.1 The BF

The BF, also known as the *pars basalis telencephali*, is a ventro-rostrally located region of the telencephalon, comprising various limbic and cortical structures that are important for cognition, learning and memory formation, emotion, and behavior, as well as for movement control and certain endocrine functions. Disturbances in this region can hence give rise to a set of neuropsychiatric disorders such as schizophrenia and Huntington's disease[70]. Despite its pivotal role in the above-mentioned functions, little is known about the organization of the BF and its neuronal populations and circuits. Even its anatomical boundaries remain controversial. This might be the reason that the historical term *"substantia innominata"*[71] is still used to describe this region. However, it is generally accepted that the BF can be divided into three parts (see Table 1), which are reciprocally connected to one another[70,72].

Table 1: Parts and regions of the basal forebrain (BF)

Part of the BF	Regions
(1) Magnocellular nuclei complex	▪ Septal region ▪ Vertical and horizontal limb of the diagonal band of Broca ▪ Nucleus basalis of Meynert
(2) Ventral striatopallidum	▪ Ventral striatum (= ventromedial caudate-putamen) ▪ Ventral pallidum ▪ Nucleus accumbens ▪ Extensive parts of the olfactory tubercle
(3) Extended amygdala	▪ Neuronal continuum between the following: o the bed nucleus of the stria terminalis and o the centromedial amgydala

Due to its obvious link to neurodegenerative disorders such as Alzheimer's disease[73], the main BF-related research focus has been on the magnocellular nuclei complex[74], which is characterized by a high number of cholinergic neurons responsible for the forebrain's main acetylcholine production. They are therefore also referred to as the "cholinergic centers". The term "magnocellular" is derived from the observation that these neurons are particularly large and heterochromatic[74-76]. Another main characteristic of these neurons is their extensive and seemingly very complex axonal network. In general, the BF is characterized by a huge diversity of cell types and complex projection systems (see Table 1 for the general parts and regions of the BF). The predominant cell types are cholinergic and gamma-aminobutyric acid (GABA)ergic neurons partially containing the calcium-binding protein parvalbumin (PV)[77,78]. Other cell profiles in this region partially colocating and/or interacting with the above-mentioned cell types are calretinin-, calbindin (CB)-[78-80], neuropeptide Y (NPY)-, galanin-, and somatostatin (SST)-positive neurons[81].

1.3.2 Prenatal Nkx2-1 expression and function

During the development of the mouse brain, Nkx2-1 expression can first be detected at E8.75 in the neuroepithelium of the ventromedial parts of the prosencephalic vesicle[32,82]. From E10 onwards, it is expressed in the following subpallial progenitor domains (see Figure 4): the medial ganglionic eminence (MGE), the preoptic area (POA), and the septal primordium[43,83,84]. Among them, the MGE – which is the main source of cholinergic and GABAergic BF populations[2,43,83,85-87] – stands out due to a particularly strong Nkx2-1 expression[43]. This expression is required for the specification of its cellular identity between E9.5 and E12.5[88]. Moreover, Nkx2-1 is suggested to act as a repressor of lateral ganglionic eminence (LGE) genes that otherwise suppress the MGE[88] and promote the transformation of the MGE into an LGE-like structure producing LGE-derivates (e.g. a striatum instead of the pallidum)[43].

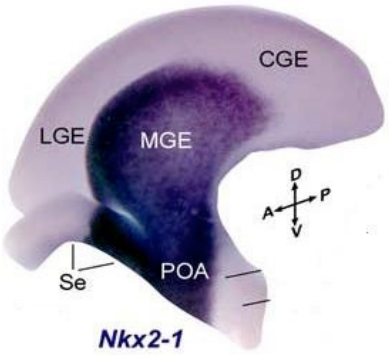

Figure 4. Nkx2-1 expression in its progenitor domains in the mouse brain. *Whole mount Nkx2-1 expression (violet) of a mouse brain is illustrated for the medial ganglionic eminence (MGE), preoptic area (POA), and the septum (Se). Further shown are the caudal and lateral ganglionic eminences (CGE and LGE respectively). Modified figure taken from Flames et al. (2007)[82] and reprinted with permission from the Journal of Neuroscience: The Society for Neuroscience, Delinea-tion of multiple subpallial progenitor domains by the combinatorial expression of transcriptional codes.* ***Flames et al. (2007)*** *Copyright © 2007.*

In general, cholinergic BF neurons arise between E11 and E15[89], resulting in two different neuronal types: (1) projection neurons of the magnocellular BF complex that migrate radially to integrate into their target region and (2) tangentially migrating interneurons to the striatum. Although other transcription factors are also involved in the developing cholinergic system that act independently of Nkx2-1[90], Nkx2-1 can be regarded as the key transcription factor of these neurons, since its homozygous deletion results in the total absence of cholinergic neurons in the magnocellular BF complex at E18.5[91]. Additionally, cholinergic neurons also arise from other Nkx2-1-positive progenitor domains, namely the POA[84] and the septal primordium[83], and certain minor subpopulations also either derive from Nkx2-1-negative domains, such as the neighboring LGE, or downregulate Nkx2-1 during development. Cholinergic neurons of the medial septum-vertical limb of the diagonal band of Broca (MSvDB) that express the neurotrophin receptor *p75^{NTR}* were recently shown to exclusively derive from the Nkx2-1-positive septal progenitors co-expressing the transcription factor Zic Family Protein 4 (Zic4)[3].

In the striatum, nearly all cholinergic and PV-containing GABAergic interneurons were identified as Nkx2-1-positive, and they were found to originate from the MGE and tangentially migrate to the postmitotic LGE[91]. In concordance with these findings, a subtotal loss of these

neurons was observed in Nkx2-1$^{-/-}$ null mice at E18.5[32]. Similarly, 80% of the cells destined for the globus pallidus, which is a structure predominantly composed of PV-containing GABAergic neurons deriving from the ventral MGE and the dorsal POA[86], were found to be Nkx2-1-positive. In line with this, no globus pallidus could be found in Nkx2-1 null mice[43]. Severely reduced cell numbers were also observed for the hippocampal formation that is thought to contain 60–70% MGE-deriving interneurons[92]. Although half of the GABAergic interneurons derive from the MGE[43,93,94], Nkx2-1 expression could not be found in migrating or postmitotic neocortical neurons[27]. Nobrega-Pereira et al. (2008)[27] provided an explanation for this by demonstrating that postmitotic GABAergic interneurons require downregulation of Nkx2-1 as a prerequisite for tangential migration to the cortical mantle[83,85,91]. This migration was reported to begin at E12.5[92,95]. Among these cortical interneurons, 65% were found to be PV-containing, and the remaining ones were identified as SST-positive by fate-mapping studies[88,94-96].

1.3.3 Postnatal Nkx2-1 expression pattern

Our working group has recently demonstrated that Nkx2-1 synthesis is not only continued into early postnatal life[27,85], but also maintained throughout life by the cholinergic and PV-containing GABAergic population of the BF[1,2]. We have further provided evidence that continuous Nkx2-1 expression is needed for the integrity of cholinergic neurons[1‡]. A detailed description of the postnatal distribution pattern of Nkx2-1-positive cells can be found in Magno et al. (2009)[2] and is briefly outlined (see also Figure 5). At birth (P0), a particularly high number of Nkx2-1-positive cells were found in the following regions: (1) the striatum (CPu= Caudate putamen), (2) the lateral globus pallidus (LGP), (3) the lateral septum (LS), (4) the MSvDB complex, and (5) the horizontal limb of the diagonal band-substantia innominata (hDB-SI). Less intense labeling was observed in the hippocampal formation and the medial amygdaloid complex. The latter is thought to derive from the subpallium[84]. Intense Nkx2-1 expression was also described for the diencephalon, particularly in the mammillary bodies,

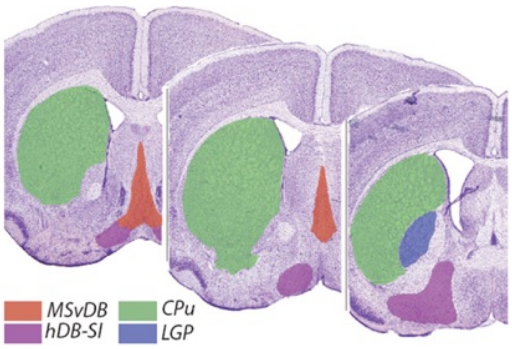

Figure 5. Regions with a high expression level of Nkx2-1 in the mouse brain. *The four regions with high Nkx2-1 expression are illustrated at three different levels along the rostro-caudal axis of the mouse brain: the Caudate putamen (CPu), the horizontal limb of the diagonal band-substantia inominata (hDB-SI), the lateral globus pallidus (LGP) and the medial septum-vertical limb of the diagonal band of Broca (MSvDB). Figure published in Magno et al. (2011)[1]. Reprinted with permission from the European Journal of Neuroscience: John Wiley and Sons, The integrity of cholinergic basal forebrain neurons depends on expression of Nkx2-1.* ***Magno et al. (2011)*** *Copyright © 2011.*

‡ These data are not subject of this dissertation as they were mainly obtained by Dr. Lorenza Magno as part of her PhD thesis[97].

the nucleus arcuatus, the ependyma of the third ventricle, and several periependymal hypothalamic nuclei[2]. Moreover, persisting Nkx2-1 synthesis was shown for the ventral tips of the subventricular zone of the lateral ventricles[2], which are thought to be remnants of the MGE[98]. Almost the same Nkx2-1 expression pattern was found in adult and aged mice[2]. However, a significant downregulation of Nkx2-1 was detected during the first postnatal week[2].

Immunohistochemical (IHC) analysis for choline acetyltransferase (ChAT) and PV revealed the following neuronal profiles. (1) In the striatum, Nkx2-1 expressing neurons were mainly PV-containing GABAergic or cholinergic neurons. (2) In the LGP, nearly all Nkx2-1-positive cells were found to contain PV, which was in turn expressed by two thirds of the LGP's GABAergic neurons[2]. (3) The LS, known to be 100% GABAergic[99], displayed no double-labeling for Nkx2-1 and PV[2]. Instead, partial co-expression with the glutamate-decarboxylase isoform 67 (GAD67) was detected, albeit less pronounced in adult mice due to the high turnover of GAD67. (4) In the MSvDB and hDB-SI, most of the Nkx2-1 expressing neurons were revealed as either PV-containing or cholinergic. (5) In the hippocampal formation nearly all of the neurons positive for Nkx2-1 were identified as GABAergic. They were mostly found in the gyrus dentatus' stratum moleculare and oriens, as well as in the hippocampus proper. (6) In the medial amygdala, Nkx2-1 was found in either PV-containing or cholinergic neurons, and in the lateral and medial mammillary bodies, where nearly half of the neurons express Nkx2-1, it was mostly identified in PV- or CB-containing neurons. (7) The periependymal hypothalamic nuclei were either Nkx2-1- or GAD67-positive, and double-labeling was only occasionally observed. To a limited extent, Nkx2-1 expression was also described for non-neuronal cells, (e.g. glial cells of the subfornical organ and the third ventricle's ependymal).

1.3.4 Effects of pre- and postnatal Nkx2-1 ablation

1.3.4.1 Pre- and postnatal ablation targeting parvalbumin (PV)-positive and/or cholinergic neurons

Since the neurological impairments observed in *NKX2-1* haploinsufficient patients could be a consequence of both pre- and postnatal deficiency in the *NKX2-1* gene, our working group generated two different conditional mutant lines and studied the effects of pre- and postnatal lack of Nkx2-1 in GABAergic and cholinergic BF neurons[1]. As these mouse lines were subsequently also used in the present study, they are described in detail in the methods section of this thesis (see Section 2.1.1). Briefly, cell-type-specific deletion was carried out using Cre-recombinase activities under the control of two promoters: (1) the GAD67-promoter to target GABAergic neurons ("prenatal mutation") and (2) the ChAT promoter to target cholinergic neurons ("postnatal mutation"). Interestingly, both the pre- and postnatal mutation resulted in a

loss of cholinergic neurons[1,2]. This is most likely due to a common Lhx6-positve progenitor of GABAergic and cholinergic neurons. The prenatal mutation supposedly occurs around E10 as GAD67 and Nkx2-1 become detectable around this time, namely, GAD67 at E10[100] and Nkx2-1 at E9.5[43,101]. The neurons thus mutate at an immature stage prior to upregulation of their neurotransmitters, GABA[100] and ChAT[99,102,103]. In line with this, postnatal monitoring of the prenatal mutation using X-gal staining in ROSA26-floxed reporter mice (results of our working group, not part of this dissertation) exhibited strong 5-Bromo-4-chloro-3-indolyl-β-D-galacto-pyranoside (X-gal) labeling in GABAergic regions early postnatally (e.g. P2), as well as slightly later, around P5/P6, in the cholinergic centers[1] (see Magno et al., 2011). The onset of the post-natal mutation was investigated in this study and is therefore presented in the results (see Section 3.1). The theory of a common progenitor for the striatum was already proven by Fragkouli et al.[104] who demonstrated that approximately 80% of the cholinergic neurons populating the striatum derive from the same Nkx2-1-positive precursor as GABAergic striatal neurons. Current research suggests the following underlying mechanism (see also Figure 6): Nkx2-1-ex-

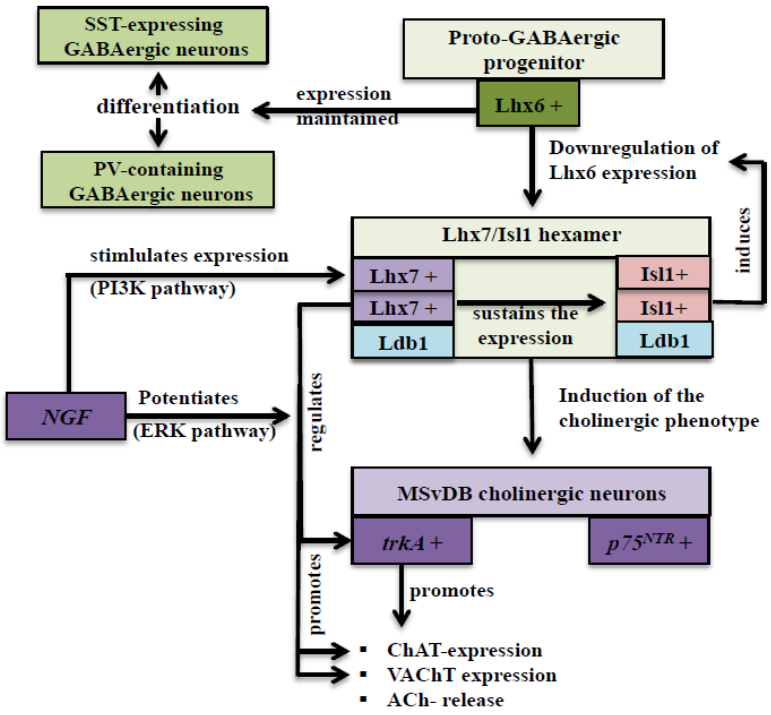

Figure 6. Differentiation of the proto-GABAergic progenitor in mice. *The proto-GABAergic progenitor initially expresses Lhx6[104], which, if maintained, leads to differentiation into GABAergic (SST-expressing or PV-containing) neurons[105,106]. Lhx6-downregulation, induced by Isl1 and sustained by Lhx7[31], as well as Lhx7/Isl1 co-expression[104] (Lhx7-Isl1/Ldb1 hexamers[107]) causes differentiation into magnocellular trkA-, and partially p75^{NTR}-positive cholinergic neurons[104]. Lhx7 expression, stimulated by NGF via the PI3K-pathway, directly regulates trkA expression[108], which is potentiated by NGF[108] via the ERK-pathway, that promotes expression of ChAT, VAChT, and ACh-release[109]. Abbreviations: Ach = acetylcholine, ChAT = choline acetyltransferase, ERK = extracellular regulated kinase, Isl1 = LIM homeobox 1, Ldb 1 = LIM domain-binding protein 1, MSvDB = medial septum-vertical limb of the diagonal band of Broca, NGF = nerve growth factor, PV = parvalbumin, p75^{NTR} = p75^{NTR} neurotrophic receptor, SST = somatostatin, trkA = tyrosine kinase A receptor, VAChT = vesicular acetylcholine transporter.*

pressing progenitors of the MGE enter the subventricular zone and express Lhx6 and GABA after leaving the cell cycle[104]. Then, the proto-GABAergic precursors either continue or stop expressing Lhx6. Those that maintain Lhx6 expression differentiate into PV-containing or SST-expressing GABAergic neurons[105,106]. Meanwhile those that downregulate Lhx6 due to induction by Isl LIM homeobox 1 (Isl1) expression[104], and then start to express Lhx7 (also known as Lhx8 or L3) as well as Isl1 instead, differentiate into magnocellular tyrosine kinase A (*trkA*)-positive cholinergic neurons[104]. Moreover, the presence of Lhx7 in these neurons sustains the downregulation of Lhx6[31]. Lhx7/Isl1 co-expression is known to be essential for developing cholinergic BF neurons[30,110-112], and Lhx7 is known to be a postnatal marker of cholinergic neurons. These factors were recently shown to play a role as inductors of the cholinergic phenotype, including the promotion of ChAT-expression and acetylcholine (ACh) release[109] in the form of Lhx7-Isl1 hexamers[107] consisting of two Lhx7, two Isl1, and two Ldb 1. In addition, Lhx7 alone enhances the expression of ChAT and the vesicular acetylcholine transporter (VAChT)[109].

The significance of Lhx7 for the cholinergic system was also demonstrated in $Lhx7^{-/-}$ null mice, where a severe loss of cholinergic neurons of the BF and the striatum was observed[111,113]. Lopes et al. (2012)[114] recently proposed that Lhx6/Isl1 antagonism is not only responsible for the generation of the cholinergic population, but also required for the maintenance of the populations' identity. Using Lhx7-deficient mice, they demonstrated that the lack of Lhx7 causes a loss of Isl1, resulting in the upregulation of Lhx6 and the subsequent transformation of immature "cholinergic" neurons into GABAergic ones[114].

1.3.4.2 Neuronal loss, motor impairments and spatial memory deficits

In the prenatal GAD-cre line, the following changes were observed at 3 months of age (results of the PhD thesis by Dr. Lorenza Magno[97]): (1) cholinergic and PV-containing neurons were significantly reduced in the CPu, LGP, MSvDB, and hDB-SI; (2) nearly no cholinergic fibers were found in their target regions; and (3) almost no Nkx2-1 expression was detectable in the BF from E14 onwards[1]. Postnatal Nkx2-1 ablation in the ChAT-Cre line reduced the number of cholinergic neurons by half in the above-mentioned regions, was accompanied by a substantial reduction in the number of cholinergic axonal terminals in the neocortex, and resulted in a less pronounced reduction in the basal ganglia and the hippocampal formation[1]. The reduction in cholinergic and PV-containing GABAergic neurons in the prenatal line[§] led to (1) severe motoric impairments due to disturbances in basal ganglia circuits revealed by a significantly

§ see PhD thesis of Dr. Lorenza Magno[97]

reduced performance on the Rotarod test, and (2) spatial memory deficits detected in the Morris water maze test, where mutants failed to learn the location of the platform. In line with the cell counts, the impairments were attenuated in postnatal mutants. In Chat-Cre mutants, no motor impairments were detected in the Rotarod test, as GABAergic neurons were not affected, and spatial memory deficits were only detected in female mutants. The reason for the latter remains unclear, as no sex-specific effects of the mutations could be identified[1].

In a recent fate-mapping study, Magno et al. (2017; separate study not related to our working group) provided further evidence of the role of cholinergic BF neurons and the septohippocampal system (SHS) in spatial learning and memory, using Zic4-Cre mice with focal septal Nkx2-1 deletion[3]. Their mutants displayed short- and long-term memory impairments in the novel object recognition test, as well as severe spatial learning and memory impairments in the T-Maze test, while social recognition, memory, and motor behavior were normal[3]. The authors further investigated the hippocampal theta rhythm – an oscillatory brain activity associated with cognitive functions such as memory and spatial learning (cf. [3,115]) – and found a shift in the oscillation frequency at running speed that might cause impairments of hippocampal processing and memory formation[3]. As the focus of this study is on the cholinergic BF population and the role of neurotrophins in Nkx2-1-deficient mice, the link between cholinergic neurons and neurotrophins is briefly outlined in the following section.

1.4 Cholinergic BF neurons and the septohippocampal system (SHS)

1.4.1 Cholinergic BF neurons and their projections

Cholinergic neurons constitute an important neuronal population in the BF although they only account for about 20% of the total number of BF neurons[70]. They are well known to widely project into neocortical fields, including the allocortical hippocampal formation. The projection system between the septal region and the hippocampal formation constitutes a distinct functional entity known as the SHS (see Figure 7[116]), which also comprises a retrograde projection by hippocamposeptal neurons via the fimbria-fornix system (FFS)[117].

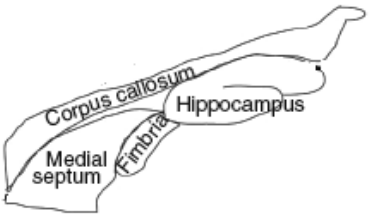

Figure 7. Location of the septohippocampal system in the mouse brain*. Modified figure, taken from Park et al. (2010)[116], reprinted with permission from Nature Neuroscience: Springer Nature. P75NTR-dependent, myelin-mediated axonal degeneration regulates neural connectivity in the adult brain.* ***Park et al. (2010)****. Copyright © 2010.*

The septal region can be subdivided into four nuclei: the dorsal, lateral, medial, and posterior septum[118,119]. Septohippocampal projection neurons mainly arise from cholinergic and PV-containing GABAergic neurons of the MSvDB-complex[120-124] that equally contribute to the projections[125]. To be specific, more than 95% of the MSvDB's cholinergic neurons project into

the hippocampal formation[126]. While cholinergic neurons display a diffuse innervation pattern, PV-containing GABAergic neurons specifically form synapses on hippocampal GABAergic interneurons[121,122]. Similarly, corresponding neurons located in more caudal regions (e.g. the hDB) give rise to rather diffuse projections to the entire neocortical mantle[77,117,127]. In rodents, the development of the septal region is accomplished around E17, and the FFS is subsequently fully established at the end of the first postnatal week[123].

1.4.2 Cholinergic BF neurons depend on the "nerve growth factor" (*NGF*) – The "neurotrophin hypothesis" for septohippocampal cholinergic neurons

Cholinergic BF neurons depend on a certain amount of the nerve growth factor (*NGF*) during brain development[4,5]. It leads to proper differentiation, including the synthesis of the marker proteins ChAT and VAChT[128-130]. *NGF*, initially discovered by Hamburger and Levi-Montalcini[131], is the first neurotrophin (synonym: neurotrophic factor) that was identified, and nowadays, it is the most important representative of the neurotrophin class, which is a group of closely related proteins playing an essential role in development and differentiation, maintenance of function and morphology, and neuronal survival and plasticity[132,133]. *NGF* is mainly synthesized by neuronal targets that are connected to *NGF*-sensitive neurons through synapses[134].

In the SHS, hippocampal GABAergic interneurons synthesize *NGF*[4,5], which is taken up by the synaptic endings of cholinergic BF neurons and subsequently transported backwards to their somata[134,135]. *NGF*-binding is mediated by the high-affinity *trkA* receptor[136-138] and the low-affinity $p75^{NTR}$ receptor[139-143]. The co-expression of ChAT, *trkA,* and $p75^{NTR}$ is characteristic of cholinergic BF neurons[144]; particularly for those of the MSvDB, where more than 90% express all three markers[144-146], high-affinity *NGF*-binding was only observed in cholinergic neurons with this co-expression pattern[146]. The role of $p75^{NTR}$ appears to involve the simultaneous modulation of *trkA* binding (i.e. to promote *NGF* binding[147]) and inhibition of binding of other neurotrophins[148,149], as well as potentiation of *NGF* binding at suboptimal *NGF* concentrations[150,151]. *TrkA* and $p75^{NTR}$ further promote axonal growth and innervation of the cholinergic target structures[152,153]. During prenatal development, the synthesis level of *NGF* is very low[154]. However, from P1 onwards, it steadily increases, reaching its maximum at P8[154] (see Figure 8). During this

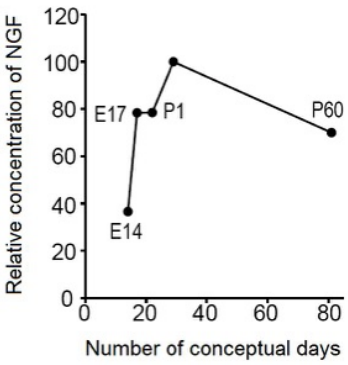

Figure 8. NGF mRNA expression in the septal region of the developing mouse brain from E14 onwards*. Nkx2-1 expression normalized to the actin beta expression is shown. Modified figure taken from Schnitzler et al. (2010)[154] and reprinted with permission of the Journal of Neuroscience: Society of Neuroscience. BMP9 (bone morphogenetic protein 9) induces NGF as an autocrine/paracrine cholinergic trophic factor in developing basal forebrain neurons.* ***Schnitzler et al. (2010).*** *Copyright © 2010.*

period, the phenotypic development of cholinergic BF neurons takes place[155-157], and they first become detectable at P4/5[99,102,103]. At the end of the first postnatal week, a fully established cholinergic cellular pattern can be observed[103] and at P12–15, a mature pattern of cholinergic fibers is established[99,103]. In the meantime, increasing expression of the *trkA* and *p75^{NTR}* receptors results in an enhanced uptake of *NGF,* thus further triggering differentiation and axonal pathfinding, so that the vast majority of cholinergic projections to the hippocampal formation and cortex is detectable within two weeks after birth[123,124,157].

The critical role of developmental *trkA*- and *p75^{NTR}*-mediated *NGF* signaling is demonstrated by the fact that neurons without access to *NGF* degenerate and probably die[102,135,158-161]. Similar observations have been made for caudo-ventral projections of the hDB-SI to the neocortical fields[5]. This selection process has led to the formulation of the "neurotrophin hypothesis," which attempts to explain how the primarily large number of neurons is reduced by apoptosis during brain development[102,135,159-162]. Unlike cholinergic BF neurons, GABAergic neurons of this region have been less intensively investigated, and in particular, knowledge about the role of neurotrophins in these neurons is scarce. To date, no dependency on *NGF* or expression of *trkA* has been described.

The available data rather suggest an important role of developmental *trkB*-mediated "brain-derived neurotrophic factor"[163.164] (*BDNF*), which has not been described for cholinergic MSvDB neurons[165]. *BDNF* and *trkB* expression can be found in the septal complex and the hippocampal formation, and *trkB/BDNF* signaling seems to be of particular importance for synaptic plasticity[163]. In line with this, both *trkB*- and *BDNF*-knock-out mice are unviable. While *trkB*-knockouts die from starvation because of their inability to suckle due to abnormalities in the trigeminal ganglion and the facial motor nucleus[166], *BDNF*-knockouts die within the first two postnatal weeks displaying severe impairments of motor functions and coordination in addition to significant neuronal loss in all sensory ganglia[167].

1.5 Aim of the study

In the past, numerous studies have provided evidence of the essential role of *NGF* and its receptors (*trkA* and *p75^{NTR}*) for the development, survival, and maturation of cholinergic neurons during early postnatal development, and for their axonal pathfinding to the various target structures (see above). For instance, mice deficient in *NGF* or *trkA* do not form the septohippocampal projections[168,169], and these neurons therefore never gain access to *NGF*-synthesizing target neurons and hence degenerate. In addition, the deletion of the low-affinity receptor *p75^{NTR}* further impairs the fate of these neurons[103,170-173]. As demonstrated by our working group (see Magno et al., 2009 and 2011)[1,2], deletion of the transcription factor Nkx2-1 leads to substantial loss of both cholinergic and PV-containing BF neurons.

In the present thesis, I aim to determine whether these two systems are linked and how my data could be discussed in the context of the current knowledge of *NKX2-1* haploinsufficiency in humans. Since, to date, no review of this rare syndrome is available that includes the majority of cases and provides analysis for different relevant quantitative and qualitative criteria, it is difficult to obtain a comprehensive overview. Therefore, I perform a systematic review of all case reports in the literature between 1998 (when the first patient was reported) and May 2nd 2020, with an emphasis on the neurological phenotype and including the questions 1) what role cognitive impairments play in this disorder, 2) if there are data suggesting an involvement of the BF and the same neuronal groups in humans, and 3) whether a genotype-phenotype correlation exists. To this end, I examine the different variants for their genotypes, phenotypes, and effects on gene function. Based on these results a disease severity score is developed that allows for a comparison of different patient groups and that could additionally serve as a tool for clinicians. Finally, the first statistical analysis for genotype-phenotype correlation including disease severity and Kaplan-Meier analysis for survival in *NKX2-1* haploinsufficient patients is conducted.

2 Materials and Methods

2.1 Mouse lines, conditional mutants and genotyping

2.1.1 Mice

All animals were housed according to institutional guidelines for animal welfare on a 12/12 hour day/night cycle with access to food and water *ad libitum*. The experiments were performed under the permit #G0004/09 of the state authority of Berlin, Landesamt für Gesundheit und Soziales (LAGeSo). In the present study, the following mouse lines were used to generate conditional mice with a pre- or postnatal mutation of Nkx2-1. For this purpose wild-type mice were also used.

2.1.1.1 Glutamate-decarboxylase isoform 67 (GAD67)-Cre line ("prenatal mutation")

The GAD67-Cre mouse line is a knock-in line that was generated and characterized by Dr. Angelika Voigt (Dissertation, University Heidelberg, Germany, 2007)[174] and kindly provided by Professor Hannah Monyer (Heidelberg, Germany). In contrast to knockout mice, where the target gene is deleted, in knock-in mice, it is specifically integrated into a certain DNA sequence (the Cre-recombinase) and then deactivated via homologous recombination. Here, a target vector was used that acts under the control of the GAD67 promoter causing a specific Cre-recombinase expression in all GABAergic neurons. Therefore, a targeting vector containing a fragment of the murine GAD67 gene and the cDNA of the Cre-recombinase, which was placed 30 base pairs upstream of the start codon of GAD67, was used. The vector was integrated into E14 embryonic stem cells by means of electroporation. Then, the stem cells with the desired mutation were injected into the blastocysts of Black 6 Inbred Mouse Strain (C57Bl/6) mice and subsequently implanted into pseudopregnant foster mice. Afterwards, male knock-in mice were backcrossed to females from the C57Bl/6 line that gave birth to viable and fertile heterozygous GAD67-Cre mice.

2.1.1.2 Choline acetyltransferase (ChAT)-Cre line ("postnatal mutation")

The ChAT-Cre mouse line, also a knock-in line, was purchased from the Jackson Laboratories (Charles River, Sulzfeld, Germany; stock number #6410). The vector used in this line acts under the control of the endogenous ChAT-promoter, causing a specific expression of the "internal ribosome entry-site-linked Cre-recombinase" (IRES-Cre) in all cholinergic neurons. This vector was electroporated into 129S6/SvEvTac-derived W4 embryonic stem cells. The stem cells with the desired mutation were subsequently injected into the blastocysts of C57Bl/6 mice, which were then implanted into pseudopregnant foster mice. Finally, male knock-ins were crossed to wild-type female siblings that had viable and fertile heterozygous offspring.

2.1.1.3 Nkx2-1-floxed line

The Nkx2-1-floxed mouse line[175] is a knockout line that was kindly provided by Dr. Shioko Kimura (National Cancer Institute, National Institutes of Health, Bethesda, USA). In this line, exon 2 of Nkx2-1 was modified by homologous recombination using a vector with two loxP sites between exon 1 and 2. It should be noted here that the description of the Nkx2-1 gene used by Kimura et al. (1999) is the one based on the initial characterization of Oguchi et al. (1995)[11], where there are only two exons, with the HD located on exon 2. Stem cells with the desired mutation were injected into the blastocysts of C57Bl/6 mice that were implanted into pseudo-pregnant foster mice. Male knockouts were bred with C57Bl/6 females that had viable and fertile heterozygous offspring.

2.1.1.4 ROSA26-floxed line

The ROSA26-floxed mouse line is a reporter mouse line that was purchased from the Jackson Laboratories (Charles River, Sulzfeld, Deutschland; stock number #2073). It was developed and characterized by Dr. Philippe Soriano (Fred Hutchinson Cancer Center, Seattle, WA)[176] and is used to check the Cre expression pattern, allowing for an assessment of when and where the Cre-enzyme is activated in mutant lines. ROSA26-floxed mice contain a floxed gene (the reporter gene), which is expressed under the control of a ubiquitous promoter, and also the LacZ gene, which encodes the enzyme β-galactosidase (β-gal). Transcription of the LacZ gene is blocked in these mice by a floxed neo sequence. However, in cells with Cre expression, the floxed neo sequence is recognized and cut out, resulting in the transcription of the LacZ gene (i.e. β-gal). This transcription product can then be visualized using β-gal immunohistochemistry or X-gal staining – the latter resulting in blue reaction products.

2.1.2 Generation of conditional mutants

The conditional mice were generated in two steps:

(1) $GAD^{cre/+}$ and $ChAT^{cre/+}$ were bred with $Nkx2\text{-}1^{fl/fl}$ mice.

(2) Mice of the F1-generation with the Cre-sequence were bred with homozygous floxed mice resulting in an F2 generation with Nkx2-1 deletion in the following:

a) GAD67-positive cells: $GAD^{cre/+}$;$Nkx2\text{-}1^{c/c}$ or

b) ChAT-positive cells: $ChAT^{cre/+}$;$Nkx2\text{-}1^{c/c}$.

Mice without the "Cre" were used as controls: $GAD^{+/+}$;$Nkx2\text{-}1^{fl/fl}$ or $ChAT^{+/+}$;$Nkx2\text{-}1^{fl/fl}$. We also analyzed heterozygous floxed mice (both lines) as additional "controls," but no differences regarding cell numbers, among other things were found when compared to the homozygous floxed mice.

2.1.3 Genotyping

The genetic tissue required for genotyping was obtained from mouse-tail biopsies and extracted using the InvisorbSpin Tissue Mini Kit (Invitek GmbH, Berlin, Germany). Polymerase chain reactions (PCRs) were carried out to amplify the obtained DNA with two reagents, namely, GoTaq Polymerase (Promega, Madison, WI) and desoxynucleotides (Sigma-Aldrich, Inc., St. Louis, MO) as well as mouse line-specific primers (see Table 2) and the cycling parameters listed in Table 3. The PCR products were subsequently run through a 1% agarose gel (in TRIS-Acetat-EDTA buffer) containing ethidium bromide (5 µl/100 ml agarose solution). The DNA bands were finally visualized with ultraviolet light and their size was determined using a DNA ladder (Fermentas).

Table 2: Mouse line-specific primers and primer hybridization temperatures

Mouse line	Origin	Primer sequences	Ta
GAD67-Cre line: "prenatal mutation"	Developed by Dr. A.Voigt (Heidelberg)[174] and kindly provided by Prof. Dr. H. Monyer (Heidelberg).	CCGGCGCCCCTTAGCTGTGAG GGTTCTTGCGAACCTCATCACTCG	63°C
ChAT-Cre line: "postnatal mutation"	Purchased from the Jackson Laboratories (Charles River, Sulzfeld, Germany; stock number #6410).	The first pair detects the wild-type allele: GTTTGCAGAAGCGGTGGG AGATAGATAATGAGAGGCTC The second pair detects the Cre-allele: AGATAGATAATGAGAGGCTC CCTTCTATCGCCTTCTTGACG	55°C
Nkx2-1-floxed line	The "neo primers"[175] were kindly provided by Dr. S. Kimura (Bethesda, USA). The "internal primer" was developed by Dr. Katharina Paulick (Dissertation, Free University Berlin, 2014).[177]	The first and second primers detect the floxed sequence ("neo primer"): TGCCGTGTAAACACGAGGAC GACTCTCAAGCAAGTCCATCC The "internal primer" detects the intermediate sequence between the floxed ones: GAAGTGGCGAAAGCTACAGG	55°C
ROSA26-floxed line	The primers were developed by Dr. K. Paulick.[177]	LacZ a: TCCCAACAGTTGCGCAGCCTGAAT LacZ b: ATATCCTGATCTTCCAGATAACTGCCG	63°C

Abbreviations: Ta = primer hybridization temperature

Table 3: Genotyping cycling parameters

Step	Duration	Temperature
1. Initial activation	5 min	95°C
2. 35 cycles of	1.) 30 sec of denaturation 2.) 30 sec of annealing 3.) 30 sec of extension	1.) 94°C 2.) specific temperatures 3.) 72°C
3. Final extension	5 min	72°C

2.2 Immunohistochemistry

2.2.1 Animal groups and tissue preparation

Histochemical and IHC investigations were carried out on perfused and vibratome-cut brains of $ChAT^{cre/+}$;Nkx2-1R26 mice at different developmental ages (P0–P15). All animals were

anaesthetized with a weight-adjusted intramuscular injection of a mixture of ketamine, acepromazine, and xylazine (0.3 ml per 20 g of body weight). Newborn and early postnatal mice (P0–P4) underwent decapitation, and their brains were rapidly stored in the fixative. For all other mice transcardial perfusion was first performed using physiologic saline solution (0.9% NaCl) and subsequently either 4% paraformaldehyde (PFA) or 0.1 M glutaraldehyde in 4% PFA. Glutaraldehyde was only used for brains that were chosen for X-gal staining, as it requires a stronger fixation that cannot be achieved with PFA. After fixation and careful removal of the skull and meninges, the brains were post-fixed overnight in 4% PFA at 4°C. Thereafter, they were embedded in agarose gel and coronally vibratome-cut as agarose blocks in phosphate buffer saline (PBS). The slice thickness was 100 µm for P0–P4 mice and 50 µm for older mice. If not directly used, the slices were stored in PBS at 4°C. Before the experiments were performed, they were freed from agarose remains and trimmed with a razor blade.

2.2.2 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) staining

X-gal staining is used to visualize cell-type-specific expression of the LacZ gene in reporter mice. The presence of the LacZ gene is revealed by the conversion of a colorless substrate of the enzyme β-gal named X-gal, into an irresolvable blue reaction product (5-Brom-4-chloro-Indigo). The staining protocol was as follows. Free-floating sections were initially preincubated for 10 min with 0.2% glutaraldehyde solution before being washed twice, first with PBS and then with the washing solution (0.02% Igepal, 0.01% sodiumdeoxycholat, and 2 mM MgCl2 in PBS). Staining was subsequently carried out overnight at 37°C in the staining solution [5 mM $K_4Fe(CN)_6$, 5 mM $K_3Fe(CN)_6$, 2 mM $MgCl_2$, 1 mg⁄mL X-gal in dimethylformamide, in 1x PBS]. On the next day, the staining result was fixed with 4% PFA solution for 10 min. After several PBS washes, the slices were transferred to object slides or double-labeled with antibodies (Nkx2-1, ChAT or PV).

2.2.3 Immunohistochemistry for β-galactosidase (β-gal), ChAT, PV, and Nkx2-1

Immunohistochemistry is a combination of histochemical and IHC techniques that are used to detect and visualize proteins in tissues using specific antibodies. The detection takes place in two steps: first, a specific, unconjugated antibody (primary antibody) binds the protein; then, the secondary antibody, which is linked to an enzyme, binds the primary antibody. In this study, secondary antibodies that were biotinylated and linked to the enzyme peroxidase were used. Due to the high affinity of biotin for the glycoprotein avidine that is added to the incubation solution, large avidine-biotine-peroxidase-complexes (ABCs) are formed (ABC-method), allowing for maximal signal amplification. The staining result is visualized by diaminobenizidine

(DAB) staining; that is, the sections are incubated in a staining solution containing the catalyst hydrogen peroxide and the substrate DAB, which is oxidized by the peroxidase into a black/brown precipitate. The following protocol was used: the sections were first preincubated with 0.6% hydrogen peroxide for 20 min to block the endogenous peroxide activity. After washing with PBS, the sections were preincubated at room temperature (RT) for 1h. The composition of the preincubation solution (in 0.1 M PBS) for the various stainings is presented in Table 4. The sections were subsequently incubated in the incubation solution containing 0.2% Triton-X (TX), 10% of the specific normal serum, and the respective primary antibodies (see Table 5) for 48 hours at 4°C.

Table 4: Composition of the preincubation solutions

Serum	β-gal	ChAT	PV	Nkx2-1
Normal serum*	10%	10%	10%	5%
BSA	5%	5%	5%	1%
TX	0.2%	1%	0.2%	0.3%

*Abbreviations and symbols:*from species in which the secondary antibody was raised, β-gal = β-galactosidase, BSA = bovine serum albumin, ChAT = choline acetytransferase, PV = parvalbumin, TX = Triton-X.*

Table 5: Primary antibodies

Antibody	Host	Type	Manufacturer	Concentration
anti-β-galactosidase	rabbit	polyclonal	Cappel	1:3,000
anti-ChAT	goat	polyclonal	Chemicon Int. Inc., Temecula, CA	1:100
anti-PV	rabbit	polyclonal	Swant, Bellinzona, Switzerland	1:5,000
anti-Nkx2-1	rabbit	polyclonal	Santa Cruz Biotechnology, Santa Cruz, CA	1:2,000

Abbreviations: ChAT = choline acetytransferase, PV = parvalbumin.

Afterwards the sections were washed with PBS and incubated with the respective secondary antibody (Vector Laboratories, Burlingame, CA, USA) for 2 h at RT, at a concentration of 1:250 (Nkx2-1 and β-gal with goat anti-rabbit IgG and ChAT with rat anti-goat IgG). After another PBS washing step, the sections were incubated for 1h with the ABC reagent (Elite Vectastain ABC Kit; Vector Laboratories, Burlingame, CA) at a concentration of 1:250. Just before the incubation time was over, the DAB solution was freshly prepared with the following reagents: 0.07% DAB, 0.001% H_2O_2, 1% cobalt chloride, and 1% nickel ammonium sulfate in 0.1M PBS. Before adding the DAB solution, the sections were carefully washed several times with PBS. The staining result was checked under the light microscope and once the desired staining intensity was achieved, the reaction was stopped by washing in PBS. Thereafter, the sections were either directly mounted onto gelatinized slides or double-labeled. Whenever double-labeling was planned, cobalt chloride and nickel ammonium sulfate – both used to intensify the staining result – were omitted. The sections were dehydrated using graded alcohols in an ascending order and mounted in Shandon Xylene Substitute Mountant (Thermo Fisher Scientific, Pittsburgh, PA).

2.2.4 Image processing

Pictures were taken with an Olympus digital camera (Olympus BX-51, Hamburg, Germany). The contrast, brightness, and color levels were adjusted in Adobe Photoshop CC (Adobe Systems Inc., San Jose, CA, USA) and Adobe Illustrator CC (Adobe Systems Inc.).

2.3 Quantitative real-time-polymerase chain reaction (qRT-PCR)

2.3.1 Animal groups

The quantitative real-time PCR (qRT-PCR) study was carried out with the brain tissue of 25 mice at P15. To compare the mutants and controls of the pre- and postnatal line, five animals of each genotype were used: the mutants ($GAD^{cre/+}$;$Nkx2\text{-}1^{c/c}$ and $ChAT^{cre/+}$;$Nkx2\text{-}1^{c/c}$) and their controls ($GAD^{+/+}$;$Nkx2\text{-}1^{fl/fl}$ and $ChAT^{+/+}$;$Nkx2\text{-}1^{fl/fl}$). In addition, five wild-type mice were used for the analysis of growth-associated protein 43 (*GAP-43)*.

2.3.2 Tissue preparation, RNA extraction, and reverse transcription

For the qRT-PCR study, cervical dislocation was used to euthanize the mice. All animals were previously anaesthetized with a weight-adjusted intramuscular injection of a mixture of ketamine, acepromazine, and xylazine (0.3 ml per 20 g of body weight). The skulls were immediately removed on ice, and the brains, after snap freezing in liquid nitrogen, were stored at -80°C until use. RNA was extracted and reverse transcription were carried out as follows: brain tissues were homogenized and the RNA extracted using 1 ml TRIzol reagent (Invitrogen, Carlsbad, CA) per 100 mg tissue. After adding 0.2 ml chloroform, the samples were incubated at RT for 10 min and then centrifuged (12,000 x g at 4°C for 15 min). This resulted in the separation of the organic phase from the aqueous RNA-containing phase. The latter was subsequently carefully removed and transferred into a new tube, which was again incubated at RT for 10 min after the addition of 0.5 ml of isopropyl alcohol. Thereafter, another centrifugation followed (12,000 x g at 4°C for 10 min). The visible pellets (i.e. RNA precipitates) at the bottom of the tubes were washed twice with 75% ethanol and then left to dry at RT. Once dried, the samples were resuspended with diethyl dicarbonate-treated water, and their RNA concentrations were measured with the BioMATE™ 3UV-Visible spectrometer (Thermo Fisher Scientific Inc. Pittsburgh, PA). The RNA (5 µl in a total volume of 50 µl) was then retrotranscribed using a cDNA Reverse Transcription Kit (Applied Biosystems Inc., Foster City, CA) according to the manufacturer's instructions. Afterwards, the obtained cDNA was diluted to a ratio of 1:3, and before usage for qRT-PCR, the samples were checked with β-Actin PCR.

2.3.3 qRT-PCR and statistical analysis

qRT-PCR was performed using an ABI PRISM 7500 Fast Real-Time-PCR Detection System (Applied Biosystems Inc., Foster City, CA), with reagents purchased from Applied Biosystems (Universal TaqMan Master Mix, primers, and probes). Measurements of *NGF*, *trkA*, $p75^{NTR}$, *trkB,* and *GAP-43,* as well as measurements of the two housekeeping genes, namely, glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) and hypoxanthine-guanine phosphoribosyltransferase (*HPRT*) – which are expressed by all mammalian cells and therefore used as reference genes – were always performed in duplicate at a final volume of 20 µl, according to the manufacturer's recommendation. Therefore, the following thermal profile was used: 3 sec denaturation at 95°C and 30 sec annealing/extension at 60°C.

For validation of the amplification efficiency, standard curves normalized to *GAPDH* expression were generated. The results of the relative gene expression were calculated with GraphPad Prism version 4.00 for Windows (GraphPad Software, San Diego, CA, USA) using the $2^{-\Delta CT}$ method. Statistically relevant differences between the mutants and the controls were assessed using the two-tailed unpaired Student's t-test with a confidence interval of 95% and the significance level $p < 0.05$. The results in the graphs are presented with standard error of the mean.

2.4 Literature review

2.4.1 Case finding, inclusion, and exclusion criteria

A search for the keywords "*NKX2.1,*" "*NKX2-1,*" "*TTF1,*" "*TTF-1,*" "*TITF-1,*" "*T/EBP,*" and "brain-lung-thyroid syndrome" was carried out on PubMed using the search filter "humans" (PubMed; last accessed May 2nd 2020). Additionally, HGMD Professional (time-limited access; last accessed August 21st 2017 for case finding) and the reference lists of all of the included case reports were screened for cases not found by a PubMed search.

The inclusion criteria were case reports of patients with *NKX2-1* variants with at least one organ manifestation of the brain-lung-thyroid syndrome. Patients with brain-lung-thyroid syndromes without *NKX2-1* variants were excluded. Cases published in more than one paper were only counted once for statistics, but the information in the different articles was merged.

2.4.2 Extraction, standardization, and classification of data

2.4.2.1 Publications

The included publications were subgrouped according to the year of publication and the number of cases per publication extracted. Then, cases that were published in multiple reports were identified.

2.4.2.2 Patient characteristics

The cases were screened for demographic information (i.e. sex, age at last visit, and country of origin). In lethal cases, data pertaining to age at death; cause of death; and whether an autopsy was performed, along with the respective results, if applicable, were extracted; and in all cases with a brain phenotype, the age at onset of neurological symptoms was extracted. All cases were checked for genetic and phenotypic information on the parents and affected relatives, as well as the mode of inheritance. Multigenerational families with affected family members were identified and grouped accordingly. Furthermore, a check was performed in all cases to determine whether information on the thyroid was provided.

2.4.2.3 Phenotypes

All of the available phenotypic information was extracted from the case reports and categorized as belonging to either the brain, lung, or thyroid phenotype. Any other findings were also collected and categorized as additional clinical features and disorders. Then, subcategories for the various manifestations of the three organ phenotypes and additional clinical features and disorders were formed, and the findings were assigned to the respective categories.

2.4.2.4 Genotypes

2.4.2.4.1 Nomenclature

All variants were first checked for the information provided on DNA (c.) and protein level (p.) and their nomenclature. For the purpose of comparability, the nomenclature was standardized according to the current recommendations of the HGVS[178] to the REFSEQ isoform 1 (NM_001079668.2; see Figure 1 for details). The initial standardization was performed using time-limited access to HGMD Professional 2017.2 (full demo version; last accessed August 21st 2017 by me)[179]. Later, all variants were checked using different online genome databases (for this, I gratefully received the help of Janic Teutsch MSc, biologist, Division of Human Genetics, University of Bern): HGMD Professional 2018.3[179] (paid online license; accessed by the Institute for Human Genetics Bern; last accessed November 15th 2018 by Janic Teutsch). WGD were further checked using Genoglyphix Build Version 3.1-2.ion (PerkinElmer; paid online license; accessed by the Division for Human Genetics Bern; last accessed November 15th 2018 by Janic Teutsch) and LiftOver (UCSC Genome Browser; free online tool; accessed via usegalaxy.org v.1.0.6; last accessed November 15th 2018 by Janic Teutsch and May 30th 2020 by me).

For the conversions of the genomic positions to isoform 1 Alamut Visual (v.2.11 Jan 2018, by Interactive Biosoftware, SOPHiA Genetics; Desktop, Pay; accessed via the paid license of the Institute for Human Genetics Bern; last accessed November 15th 2018 by Janic Teutsch) and

LUMC Mutalyzer[180] (mutalyzer.nl; free online tool; last accessed November 15th 2018 by Janic Teutsch and May 30th 2020 by me) were used. All cases where the nomenclature in this study differs from the original article are indicated, and the original names are provided in the respective chapters.

2.4.2.4.2 Subgroups

After standardization of the nomenclature the variants were analyzed and subgrouped for

a) their type: WGD, missense, other deletions, nonsense, insertion, insertion/deletion (synonym indel), duplication, splice site variant, and stop loss;
b) their position within the *NKX2-1* gene: WGD, in the HD, outside the HD but affecting the HD (non-HD+), and variants outside of the HD with no effect on the HD (non-HD);
c) the deletion size of WGD (in Mb);
d) location inside or outside of the gene regions c.96–565 and c.753–881;
e) their involvement of functional domains (i.e. TN, N-TAD, HD, Domain I, NK2-SD, and C-TAD) and phosphorylation sit1es (i.e. c.100_102;p.34, c.124_126; p.42, c.142_144;p.48, c.157_159;p.53, c.850_852;p.284, c.1069_1071; p.357 and c.1096;p.366);
f) their mode of inheritance (*de novo* or familial variants);
g) concomitant deletion of other genes in all patients and the subgroup of WGD, and
h) subanalysis of WGD with additional deletion of other genes for the number of additionally deleted genes and their location: proximal (chromosomal coordinates hg19: < chr14:36.985.602) and/or distal location (chromosomal coordinates hg19: > chr14:36.990.354).

For the statistical analysis (see below), the following genotypic groups were defined (see Tables 6a and 6b), and each patient assigned to the respective groups depending on his or her variant.

Table 6a: Genotypic groups based on the impaired gene region (whole gene deletions [WGD], homeodomain [HD], non-HD+, non-HD)

Genotypic group	Definition
Whole gene deletions (WGD)	Deletions encompassing the whole Nkx2-1 gene
Homeodomain variants (HD)	Variants encompassing the HD
Non-homeodomain variants affecting the HD (non-HD+)	Variants located outside of the HD but affecting it
Non-homeodomain variants (non-HD)	Variants located outside of the HD that do not affect it

Table 6b: Genotypic groups based on the type of variant (WGD, missense, all other variants [AOV])

Genotypic group	Definition
Whole gene deletions (WGD)	Deletions encompassing the whole Nkx2-1 gene
Missense	All missense variants
All other non-WGD variants (AOV)	All variants that are not WGD or missense variants

2.4.2.4.3 Impairments of gene function

Variants with information about their effects on gene function were identified, and functional test-related information and the results were extracted for subanalysis. All the variants with the same functional impairments were grouped together and then subanalyzed for the involvement of functional domains, namely, a) the number and combination of impaired functional domains and phosphorylation sites and b) the overall impairment of each functional domain and/or phosphorylation site.

2.4.3 Development of a disease severity score

After analysis of the phenotypes (see Section 2.4.2.2) all brain, lung, and thyroid manifestations with – *on the basis of current knowledge about the syndrome* – plausible correlation with *NKX2-1* variants were chosen, and each manifestation was assigned points between one and three (one for mild, two for moderate and three for severe impairments). In this way, an overall disease severity score and subscores for the three organs were developed. For each of the four categories (overall and three organ subscores), severity score ranges were defined (mild, moderate, severe, and very severe).

2.4.4 Statistical analysis

The following different approaches were used to analyze the patient population and to investigate whether a genotype-phenotype correlation exists in this heterogeneous patient population. All statistical analyses were performed using the SPSS/PC software package (versions 25.0–27.0; SPSS Inc., Chicago, IL, USA).

2.4.4.1 Publications

To analyze the development of publications on the syndrome, the number of first-time publications per year and the number of published cases per year were calculated. Follow-up studies were excluded from this calculation.

2.4.4.2 Age distribution

The median age was calculated for all patients with available information on the last visit and/or death. The age at death was further subanalyzed for pulmonary causes and age at death due to

the lung phenotype and non-pulmonary causes. For patients with neurological manifestations and known age at onset of these symptoms, the median age at onset was calculated. The above-mentioned calculations were performed for the whole patient population, the genotypic groups based on the impaired gene region, and the genotypic groups based on the type of variant. All median ages were indicated with interquartile ranges (IQRs).

2.4.4.3 Frequencies

Absolute and relative numbers were calculated for the following categories:

1) Lacking information on: gender, age at last visit or death, age at onset of neurological symptoms (in patients with a brain phenotype), genetic testing of parents, and information on thyroid function. These calculations were performed for the whole patient population, the genotypic groups based on the impaired gene region, and the genotypic groups based on the type of variant.
2) In patients with available information: gender, mode of inheritance, overall mortality, mortality due to the lung phenotype or non-pulmonary causes, and whether autopsies were conducted. These analyses were performed for the whole patient population, the genotypic groups based on the impaired gene region, and the genotypic groups based on the type of variant.
3) The number of affected families, family members, multigenerational families, and patients without affected family members. These calculations were performed for the whole patient population.
4) Location of the variants on the exons (1–3), introns (1–2) , and the gene regions c.96–565 or c.753–881.
5) Types of variants (WGD, missense, other deletions, nonsense, insertion, indel, duplication, splice site variant, and stop loss). These analyses were conducted for the whole patient population, patients with known gender, patients with known mode of inheritance, and the genotypic groups based on the impaired gene region.
6) Non-WGD (i.e. all variants except WGD), all other variants (AOV, i.e. not WGD or missense) and the genotypic groups based on the impaired gene region. In addition, the frequencies of non-HD+ and non-HD for the gene regions c.96–565 and c.753–881 were calculated.
7) Subanalysis of overall involvement of functional domains for each functional domain (TN, N-TAD, HD, Domain I, NK2-SD, and C-TAD) and phosphorylation site, as well as subanalyses of different combinations of functional domains (0–6) and phosphorylation sites

(0–7). These analyses were performed for the whole patient population and the subgroups of missense variants, AOV, and the gene regions c.96–565 and c.753–881.

8) Overall deletion of other genes, number of additionally deleted genes per patient, and frequencies of the additionally deleted genes. Optionally, genes with known overlapping phenotypes with *NKX2-1* deficiency were subanalyzed for the frequencies of typical manifestations. Additionally, WGD were subanalyzed for the additionally deleted gene regions: proximal (chromosomal coordinates hg19: < chr14:36.985.602) and/or distal location (chromosomal coordinates hg19: > chr14:36.990.354).
9) Subanalysis of the respective functional impairments in variants with available information regarding their effects on gene function. Additionally, the variants with functional impairments were analyzed for involvement of functional domains and phosphorylation sites and for their different combinations of domain (0–6) and phosphorylation site (0–7) involvement.
10) The overall frequencies of the different organ manifestations, as well as additional clinical features – calculated for the whole patient population, the genotypic groups based on the impaired gene region, the genotypic groups based on the type of variant, and the gene regions c.96–565 and c.753–881.

2.4.4.4 Disease severity

2.4.4.4.1 Median scores

Each patient was scored according to the disease severity score for overall, brain, lung, and thyroid disease severity. Then, median scores for the overall and three organ disease severity scores were calculated for the whole patient population. In multigenerational familial cases, the scores of the family members were compared to investigate the question of multigenerational differences. Furthermore, each patient was assigned to various subgroups according to the following criteria, if the respective information was available or if the criteria were applicable and median scores were calculated: a) gender; b) mode of inheritance; c) genotypic groups based on the impaired gene region; d) genotypic groups based on the type of variant; e) the gene regions c.96–565 and c.753–881; f) affected functional domains (various combinations of TN, N-TAD, HD, Domain I, NK2-SD, C-TAD, and NK2-SD, or no domain); f) subanalysis of WGD: WGD with proximal and distal deletion and WGD with only additional distal deletion of other genes; and g) subanalysis of AOV: nonsense, other deletions, splice site variants, duplications, insertion, stop loss, and indels. All median scores were indicated with IQRs.

2.4.4.4.2 Unpaired Student's t-test

The differences between the groups in Categories a) to g) in Section 2.4.4.4.1 were further statistically assessed using the two-tailed, unpaired Student's t-test with confidence intervals of 95% and the significance level $p < 0.05$. A complete correction for multiple testing of all the p-values was not possible, and their significance thus remains exploratory.

2.4.4.5 Other median numbers

Median numbers were further calculated for two categories and indicated with IQRs. The first category is the deletion size of WGD with available information. In cases with additional deletion of other genes, the median number of the number of additionally deleted genes was calculated. The second category is the duration of oxygen treatment in patients with available data.

2.4.4.6 Genotype-phenotype correlation

For the genotype-phenotype correlation analysis, the most common organ manifestations that – on the basis of current knowledge about the syndrome – can be related to *NKX2-1* were chosen. These organ manifestations were identified through the previous analysis outlined in Section 2.4.2.3.

2.4.4.6.1 Unpaired Student's t-test

Differences in the frequencies of organ manifestations between the gene regions c.96–565 and c.753–881 were investigated using the unpaired Student's t-test with confidence intervals of 95% and the significance level $p < 0.05$. A complete correction for multiple testing of all the p-values was not possible, and their significance thus remains exploratory.

2.4.4.6.2 One-way analysis of variance (ANOVA)

One-way analysis of variance (ANOVA) was performed for comparison of a) the genotypic groups based on the impaired gene region and b) the genotypic groups based on the type of variant. ANOVA is a method that compares the means between different groups and determines whether these differences are statistically significant. It tests the null hypothesis (i.e. that there is no significant difference between the groups); in cases with a statistically significant result, the null hypothesis is rejected, and the alternative hypothesis of two different group means is accepted. In this study, the null hypothesis was that no correlation exists between the genotypic groups (the predictors, synonym independent variables) and the phenotypes (i.e. the different organ manifestations; the dependent variables). If statistically significant results are found, then the alternative hypothesis (i.e. that the phenotype depends on the genotype) is accepted. All p-values obtained from this study were corrected for multiple testing using the post hoc Tukey

HSD test. However, it should be noted that the significance of theses *p*-values remains exploratory.

2.4.4.7 Kaplan-Meier analysis for survival

For the Kaplan-Meier analysis for survival, patients with known age at last visit and information on survival and/or death were identified, and overall survival was calculated. Furthermore, since the lung phenotype has previously been identified as the main cause of mortality in other case series[13,60], the patient population with available information on survival was divided into two groups, namely, those with a lung phenotype and those without known lung manifestations at the timepoint of reporting, and the survival for the two subgroups was calculated.

3 Results

3.1 Cre-recombinase activity in the ChAT-Cre postnatal line

To study the mutation in cholinergic neurons of the postnatal line, Cre-recombinase activity was monitored using X-gal staining and IHC double-labeling techniques at different stages. Therefore, Cre-positive ChAT mice were bred with the Cre-reporter-line ROSA26 LacZ (R26R), and their LacZ-expressing offspring (ChAT$^{cre/+}$;Nkx2-1R26) were investigated. X-gal precipitates were not detected before P4 (see P2, Figure 9A) in ChAT$^{cre/+}$;Nkx2-1R26 mice.

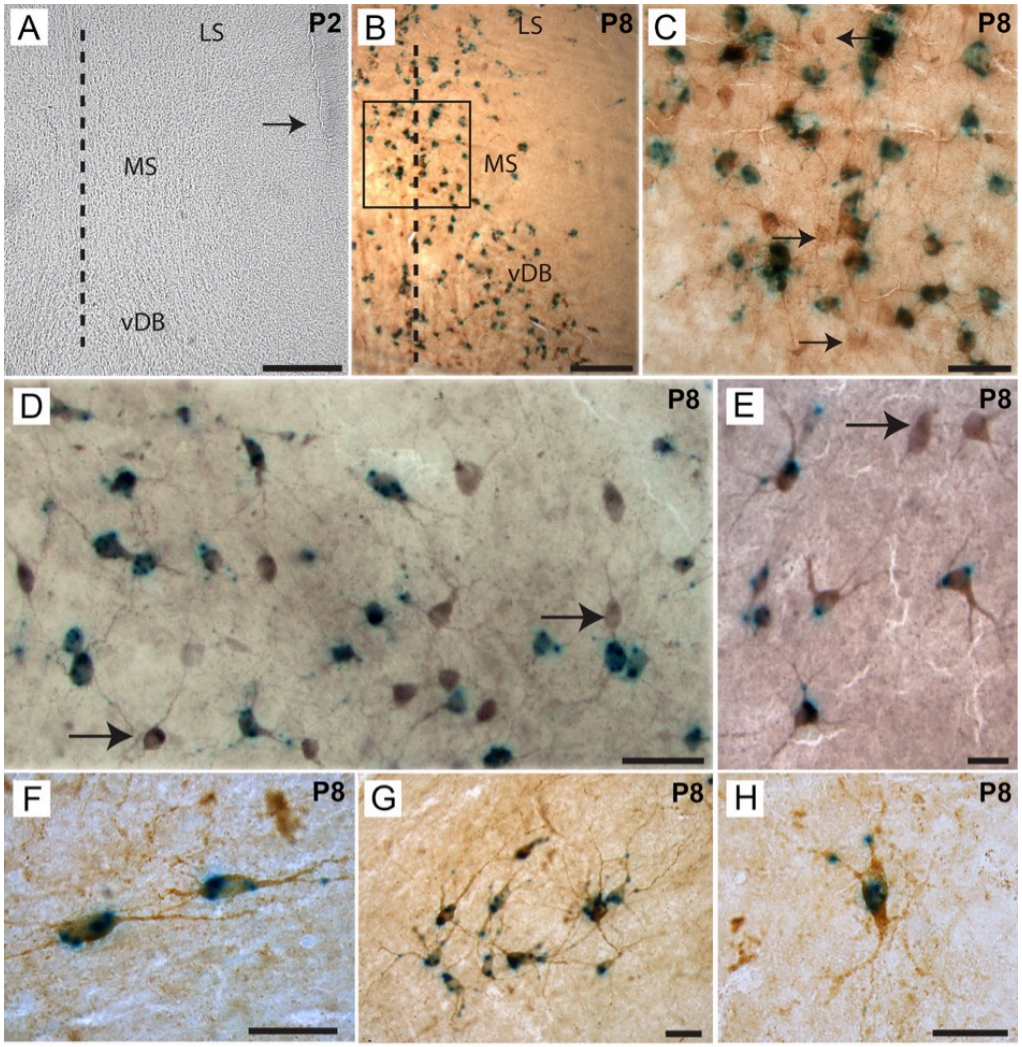

Figure 9. Onset of the Cre-recombinase activity in the choline acetyltransferase (ChAT)-Cre postnatal mutant line*. 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) staining is depicted for representative coronal sections of ChAT$^{cre/+}$;Nkx2-1R26 mice at Postnatal Days P2 (A) and P8 (B–H). The dotted lines indicate the median axis of the brain (A, B). As shown for P2 (A), no X-gal staining was seen before P3/4, not even in the regions surrounding the lateral ventricle. From the first postnatal week onwards, strong X-gal labeling was observed in the cholinergic centers: the MSvDB complex (B; boxed area shown at higher magnification in C), the magnocellular nuclei of the hDB-SI (D), and the CPu (E). Following ChAT immunohistochemistry, these X-gal precipitates were nearly always found in cholinergic neurons (F–H). At higher magnification, the precipitates could be well localized in the cell bodies and axons of cholinergic neurons (F–H). Only occasionally, solely ChAT-immunoreactive cells were observed (see arrows in C–E). Scale bars: 100 µm (A), 200 µm (B), 50 µm (C,D), 10 µm (E–H). Publication of Figures A–E in Magno et al. (2011)[1] and reprinted with permission from the European Journal of Neuroscience: John Wiley and Sons. The integrity of cholinergic basal forebrain neurons depends on expression of Nkx2-1.* ***Magno et al. (2011).*** *Copyright © 2011.*

Then, from P7/P8 onwards, strong labeling of X-gal-positive cells was observed in the cholinergic BF centers (see Figure 9): the MSvDB complex (B and C), the magnocellular nuclei of the hDB-SI (D), and the CPu (E).

The localization of these X-gal precipitates corresponds to the distribution of BF cholinergic neurons in adult mice (see also Figures 11 and 12 for comparison), and the great majority of these neurons could subsequently be confirmed to be cholinergic by additional ChAT staining (Figure 9, B–H). However, a small group of cholinergic BF neurons was X-gal-negative (see Figure 9C–E, black arrows point at some examples), and PV/X-gal double-labeling was not observed in any of the analyzed regions (not shown). Additionally, some of the X-gal stained sections of ChAT$^{cre/+}$;Nkx2-1R26 mice at P8, which were also used for Figure 9, were double-labeled for Nkx2-1 to monitor its expression in these animals (see Figure 10).

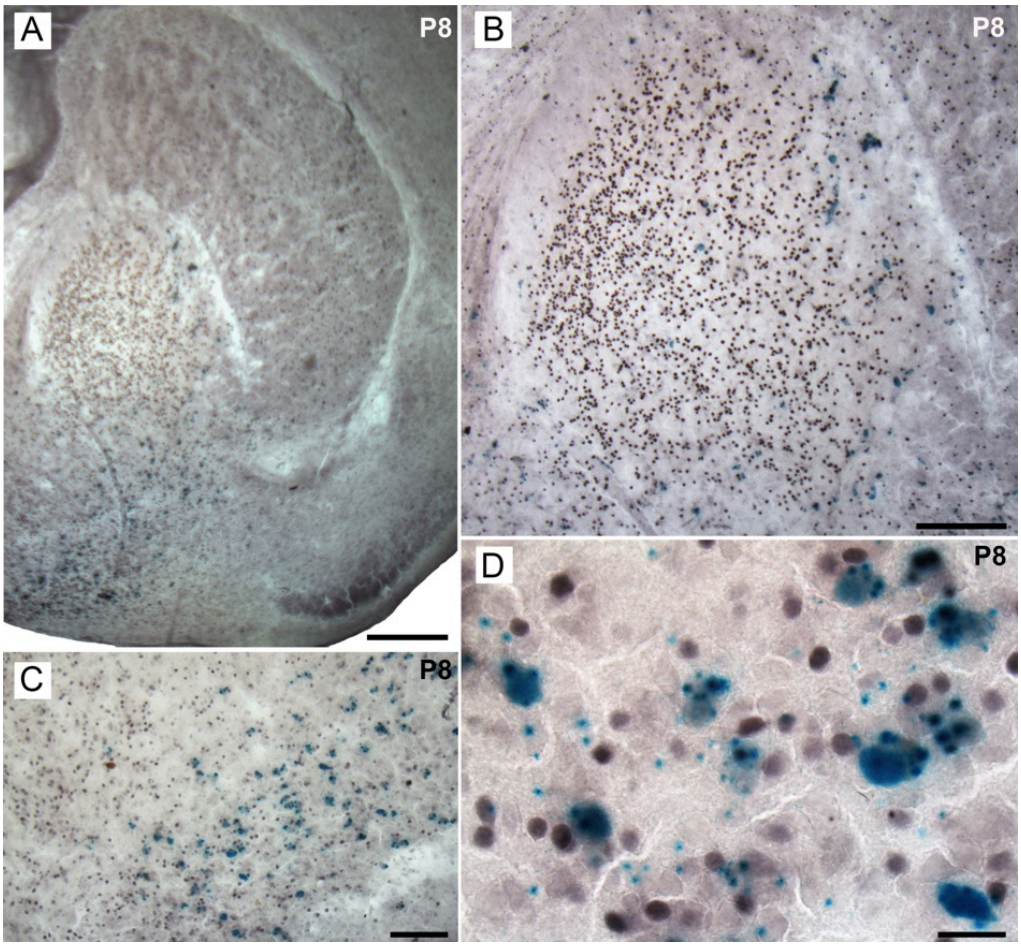

Figure 10. Nkx2-1 expression in ChAT$^{cre/+}$;Nkx2-1R26 reporter mice at P8. *Double-labeling for X-gal and Nkx2-1 is depicted for representative coronal sections at the level of the globus pallidus. Nkx2-1 labeling appears as dark violet to brown staining of the cell nuclei (D). Here, intense Nkx2-1 labeling was observed throughout the CPu (A, B), LGP (A, B), and the HDB (C, D). In line with the lower number of cholinergic neurons in the LGP (B), only a low rate of co-localization with X-gal, and a corresponding high density of co-localizing cells in the HDB (C, D), which contains a high number of cholinergic neurons, was observed. Scale bars: 500 µm (A), 100 µm (B), 50 µm (C), and 10 µm (D).*

This staining resulted in the typical distribution pattern of Nkx2-1-positive cells in the BF, as described in detail in Magno et al. (2009)[2], and confirmed the co-localization of X-gal and Nkx2-1 in ChAT-Cre targeted regions. In the cholinergic centers, such as in the HDB (Figure 10C and D), a high density of X-gal labeling and co-localization with Nkx2-1 was observed. In contrast, the LGP, which is a region known for its high density of predominantly Nkx2-1-positive PV-containing GABAergic neurons, only had low co-localization of both markers (see Figure 10A and B).

Finally, ChAT$^{cre/+}$;Nkx2-1R26 mice at P15 were investigated to monitor the Cre-recombination at the timepoint when the number of cholinergic neurons has reached the level of adult mice. To this end, β-gal immunohistochemistry was used. At P15, intense labeling was observed in the cholinergic centers and in the regions known to contain Nkx2-1-positive cholinergic neurons (see Figures 11 and 12), which is in line with other findings and confirms that only cholinergic neurons are targeted by the postnatal mutation.

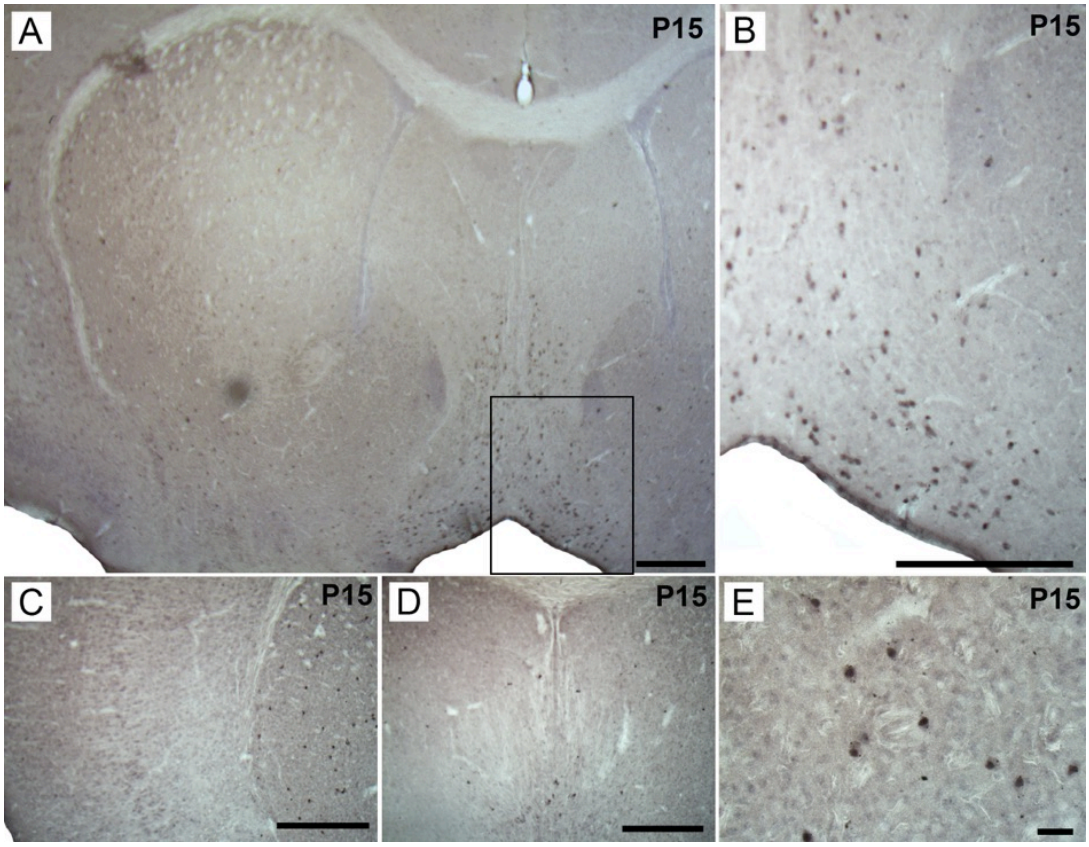

Figure 11. Cre-recombinase activity in ChAT$^{cre/+}$;Nkx2-1R26 mice at P15 (Part 1). *Here, β-galactosidase (β-gal) immunohistochemistry is shown for representative coronal sections at the level of the medial septum (MS). Strong β-gal immunoreactivity corresponding to the distribution of cholinergic neurons was observed in the MSvDB (A; higher magnification of the boxed area in B) and the CPu (C; at higher magnification in E). Only few β-gal-positive cells were observed in the dorsal MS (D). The cortex was entirely free of Cre-recombination activity (C). Scale bars: 500 µm (A), 200 µm (C, D), 50 µm (B), and 20 µm (E).*

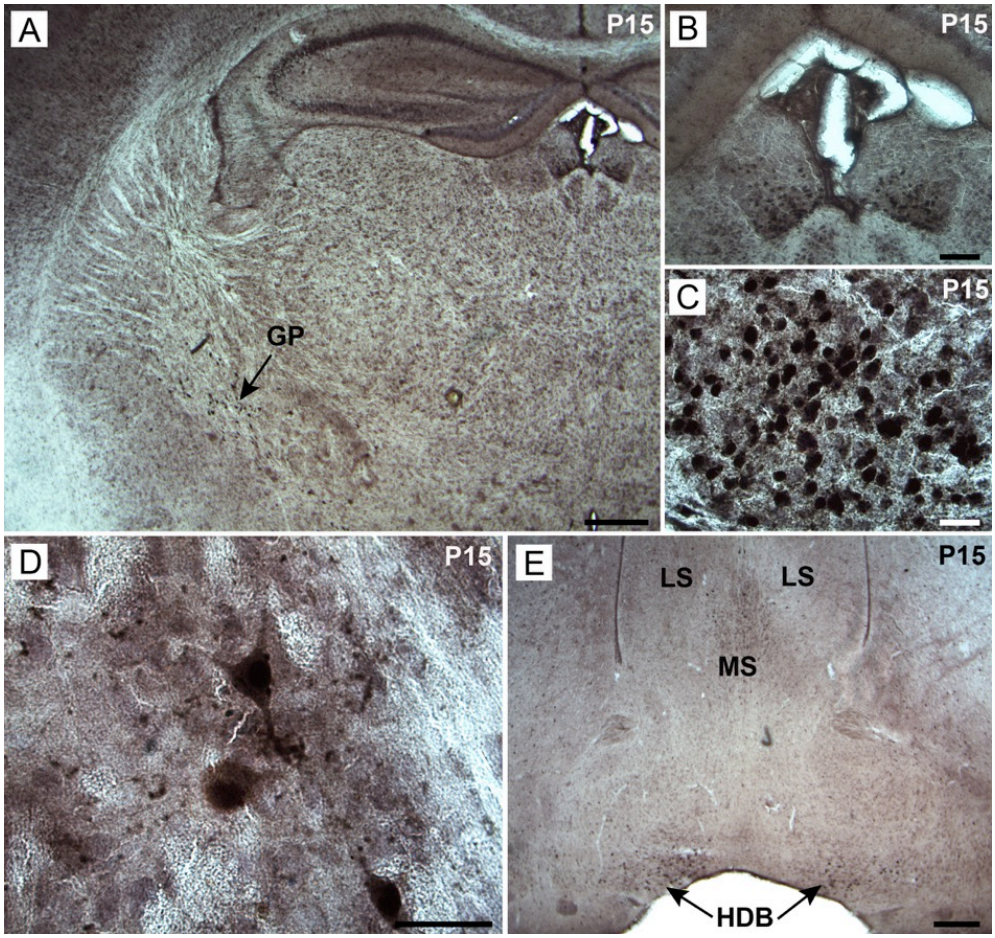

Figure 12. Cre-recombinase activity in ChAT$^{cre/+}$;Nkx2-1R26 mice at P15 (Part 2). *Representative coronal sections at the level of the hippocampal formation and the dorsal HDB (E) are shown. Few β-gal-positive cells were observed in the medial and lateral parts of the globus pallidus (LGP; marked by a black arrow in A and shown at higher magnification in D), and intense labeling was observed in the MS and the HDB (E). Interestingly, strong labeling was also detected in the subfornical organ (B and higher magnification in C), which contains non-cholinergic pallial-derived cells known to express Nkx2-1 and to be involved in body fluid homeostasis and blood pressure regulation. No immunoreactivity was observed in the LS (E), which is entirely GABAergic, thus confirming that GABAergic neurons are not targeted by the postnatal mutation. Scale bars: 200 µm (A, E), 50 µm (B), and 10 µm (C, D).*

3.2 Expression of *NGF* and its receptors tyrosine kinase receptor A (*trkA*) and *p75* neurotrophin receptor (*p75^{NTR}*)

Figure 13 compares the expression levels of *NGF* and its receptors *trkA* and *p75^{NTR}* in the pre- and postnatal mutant line at P15 using qRT-PCR. The expression levels of *NGF*, *trkA*, and *p75^{NTR}* were normalized to the housekeeping gene *GAPDH.* As illustrated in Figure 13a, in GAD$^{cre/+}$;Nkx2-1$^{c/c}$ mice with prenatal mutation of cholinergic and PV-containing GABAergic neurons, the expression levels of both *trkA* (-82%, $p = 0.01$) and *p75^{NTR}* (-60%, $p = 0.0006$) were significantly reduced. The degree of reduction in receptor expression was found to be attenuated in postnatal mutants with deletion of only cholinergic neurons (Figure 13b). Moreover, while in ChAT$^{cre/+}$;Nkx2-1$^{c/c}$ mice, a significant decrease was observed for *trkA* (-56%, $p < 0.0001$), the results for *p75^{NTR}* could not be considered significant ($p = 0.08$). In addition, no

significant differences could be detected for *NGF* between the mutants and their controls in either of the mutant lines (Figure 13 a and b).

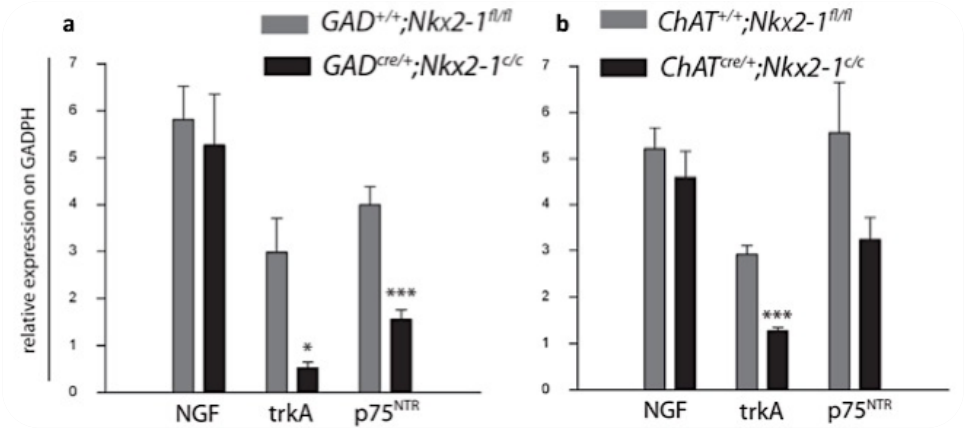

Figure 13. Quantification of nerve growth factor (NGF), tyrosine kinase receptor A (trkA), and p75 neurotrophin receptor ($p75^{NTR}$) expression levels at P15 in pre- and postnatal mutant mice. *The results of quantitative real-time polymerase chain reaction (qRT-PCR) for 5 mice of each group are shown. 14a: $GAD^{cre/+};Nkx2\text{-}1^{c/c}$ (prenatal mutants) and $GAD^{cre/+};Nkx2\text{-}1^{fl/fl}$ (prenatal controls); 14b: $ChAT^{cre/+};Nkx2\text{-}1^{c/c}$ (postnatal mutants) and $ChAT^{cre/+};Nkx2\text{-}1^{fl/fl}$ (postnatal controls). The expression of NGF, trkA, and $p75^{NTR}$ was normalized to GAPDH. The histograms show average values + standard error of the mean (two-tailed, unpaired Student's t-test), and significance is indicated by asterisks (*p < 0.01,***p < 0.001).The figure indicates significantly reduced expression levels for a) trkA and $p75^{NTR}$ in the prenatal line and b) trkA in the postnatal line. No significant differences were found for NGF expression in either of the lines. The figure was already published in Magno et al. (2011)[1] and reprinted with permission from the European Journal of Neuroscience: John Wiley and Sons. The integrity of cholinergic basal forebrain neurons depends on expression of Nkx2-1.* ***Magno et al. (2011)****. Copyright © 2011.*

3.3 Expression of the axonal sprouting marker growth-associated protein 43 (*GAP-43*)

Figure 14 illustrates the results of the quantification of the expression levels of the axonal sprouting marker *GAP-43* in the pre- and postnatal mutant line at P15 using qRT-PCR. The expression levels of *GAP-43* were normalized to *GAPDH* expression. Interestingly, the mutants as well as the controls of both transgenic lines were found to have lower expression levels compared to wild-type mice. However, the analysis did not reveal any significant differences between the mutants and controls of either mutant line.

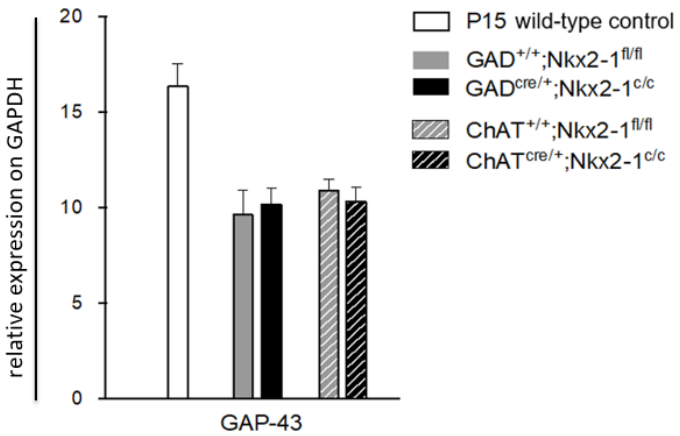

Figure 14. Quantification of growth-associated protein 43 (GAP-43) expression at P15 in pre- and postnatal mutant mice and wild-type controls*. The results of qRT-PCR for five mice of each group are shown. The expression of GAP-43 was normalized to GAPDH expression. The histograms show average values + standard error of the mean (two-tailed, unpaired Student's t-test). The differences between the mutants and their controls were not found to be significant. However, the expression level was higher in wild-type controls.*

3.4 Expression of the brain-derived neurotrophic factor *(BDNF)* receptor *trkB*

Figure 15 illustrates the results of the measurements of the expression levels of the *BDNF* receptor *trkB* in the pre- and postnatal mutant line at P15 using qRT-PCR. The expression of *trkB* was normalized to the expression of the housekeeping gene *GAPDH*. No significant differences in *trkB* expression were detectable between the mutants and their controls ($GAD^{cre/+}$;$Nkx2\text{-}1^{c/c}$: -5.5%, $p = 0.6705$; $ChAT^{cre/+}$;$Nkx2\text{-}1^{c/c}$: -5.15%, $p = 0.4391$).

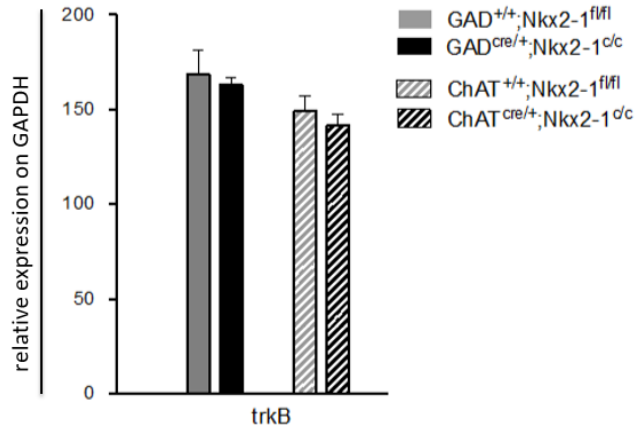

Figure 15. Quantification of tyrosine kinase receptor B (trkB) expression at P15 in pre- and postnatal mutant mice. *The results of qRT-PCR for five mice of each group are shown. The expression of trkB was normalized to the expression of GAPDH. Histograms show average values+ standard error of the mean (two-tailed, unpaired Student's t-test). The differences between the mutants and their controls were not found to be significant.*

3.5 Literature review: clinical update on *NKX2-1* haploinsufficiency

3.5.1 Identification of cases

In total, 91 case reports (88 full articles in peer-reviewed journals and three congress abstracts) were found, describing the medical histories of 351 patients. The cases were identified as follows: first, PubMed was searched (last accessed May 2nd 2020) for certain key words (see Table 7) using the filter "humans," yielding 80 papers. Five additional publications and three congress

Table 7: PubMed search results: key words, inclusion, and exclusion of articles

Key word with the filter "humans"	Search result (n)	Included (n)	Excluded (n)
NKX2.1/NKX2-1	51	40	12
TTF1	23	6	17
TTF-1	77	36	41
TITF-1	14	13	1
T/EBP	0	0	0
Brain-lung-thyroid syndrome	10	8	2

Abbreviations: n = number of articles, T/EBP = thyroid-specific enhancer-binding-protein , TITF-1/TTF1/TTF-1 = thyroid transcription factor 1.

abstracts were found in the online variants database HGMD Professional (time-limited access, last accessed for this purpose: July 31st 2017). Two articles[69,181] were identified from the reference lists of other case reports[182,183]: one from a review article[53] and the other from a follow-up study[184]. Cases without at least one organ manifestation and those with a phenotype but no variant of the *NKX2-1* gene were excluded (cf.[185,186]). Furthermore, four variants affecting 25 patients were excluded, as their location could not be verified on the *NKX2-1* gene. They were

described as *NKX2-1* variants but found on chromosome 9[187]. Overall, 27 patients[13,26,36,46,61,184,188-196] were described twice – first in a primary case report and then in follow-up studies[26,36,61,69,184,191,193,195,197]. The information from these multiple reports was merged, and patients were counted once. Figure 16 depicts the increase in the number of publications and cases published annually in the last 15 years.

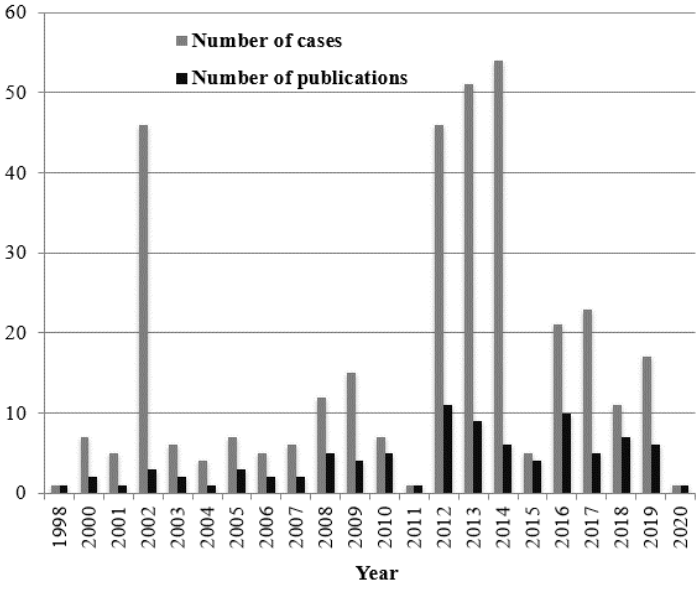

Figure 16. Published cases per year between 1998 and May 2nd 2020. *The gray bars indicate the number of published cases per year, and the black bars denote the number of publications per year. Here, only the first-time publications of patients and their follow-ups were counted.*

3.5.2 Comparability of cases

An analysis of full-text case reports revealed great differences in the level of detail of case presentations. Several papers did not provide basic demographic information, such as gender and age (see Table 8 for cases lacking data). Due to the rare nature of the syndrome, in this study all cases – those with and without certain data – were included to obtain as much information as possible about this patient population. Whenever specific statistical analyses were performed, the cases lacking the respective information were excluded and indicated.

Table 8: Cases with lacking data on basic patient characteristics

Characteristic with unavailable data	n	%
Gender[26,60,66,198-206]	19/351	5
Age at last visit or death[26,64,65,68,188,192,198-200,203-215]	105/351	30
Age at onset of neurological symptoms[14,26,45-47,53,60,61,63-65,68,182,188,193,194,196,198,199,205-232]	107/309*	35
Genetic testing of parents[13,14,26,46,60,198,203,227,232-237]	34/351	10
Thyroid function[13,53,181,184,188,198,215,231]	53/351	15

*Abbreviations and symbols: *This category is only applicable to those with a brain phenotype (309/351 patients), n = number of patients.*

3.5.3 Patient demographics: gender, age, mortality rate, and country of origin

The majority of patients (63%, 221/351) were members of the 58 affected families with familial variants. The case reports were from 23 countries (see Table 9), mainly reporting Caucasians.

An analysis of patients with known gender (see Table 8 for cases lacking information) revealed slightly more females (56%, 185/332 patients)[13,14,17,26,53,55,60-65,67,181,182,184,188-194,196,201, 206-214,216-222,224-226,228,230,231,233,234,236,238-249] than males (44%, 147/332 patients)[13,14,20,26,45-47,53,55,60,63,64, 68,181,183,184,188,192-194,201,206,208-210,212,215-217,219,221-223,226-229,231-233,235-238,242,243,248-254].

Table 9: Country of origin and numbers of affected families, family members and patients without affected families

Country	Overall patients (n)	Families (n)	Affected family members (n)	Patients without affected families (n)	References
USA	69	8	46	23	13,188,191,192,200,212,213,217,220,234,236,250
France	52	12	32	20	14,60,61,68,196,206,226
Germany	47	6	18	29	26,45,46,209,235,251,253
Netherlands	31	2	27	4	47,53,184,188,208
Italy	34	8	24	10	62,64,66,67,188,194,201,202,205,207,208,214,233,237,252
Japan	21	6	17	4	55,184,189,204,208,219,228,229,238,247
UK	19	3	8	11	63,211,218,242
Spain	15	3	10	5	182,223,225,227,230,233,239
Canada*	11	1	5	6	17,65,181,184,241
Israel	11	1	10	1	198,210,215,231
China	11	1	7	4	203,248,249,254
Norway	8	1	8	0	193
Australia	7	2	5	2	199,222,232
Finland	3	0	0	3	216
Austria	1	0	0	1	183
Portugal	3	1	2	1	221,246
Hungary	2	1	2	0	243
Belgium	1			1	184,190
Brazil	1			1	244
Ireland	1			1	224
Saudi Arabia	1			1	245
Switzerland	1			1	240
Vietnam	1			1	181
Total	**351**	**58**	**221**	**130**	**91**

Abbreviations and symbols: * One patient was reported to be hispanic (not further specified); n = number.

The median age of all patients at last visit or death with available data was 11.5 years (IQR 4–28 years; data from 246 patients). The age distribution and the age at onset of neurological symptoms are presented in Figure 17. The overall mortality rate was 6% (21/351 patients) with a median age at death of 1.5 years (IQR 0.5–16 years) for all patients. The main cause of death was pulmonary (86%, 18/21 patients), predominantly due to respiratory failure in infancy (71%, 15/21 patients): death at 28 days[234], 1 month[13], 40 days[17], 3 months[254], 4.5 months[250], 6 months[13], 10 months[251], 14 months[181], 17 months[13], 18 months[14,196], 3 years[13,189], 5 years[13], and 13 years[189]. Other pulmonary causes were lung cancer (10%, 2/21 patients: death at 23 years[47] and 62 years[210] respectively) and rejection of a lung transplant (5%, 1/21 patients: death at 19 years[13]). Non-pulmonary causes included leukemia (10%, 2/21 patients: death at 8 years[209] and

59 years[65] respectively) and a motor vehicle accident (5%, 1/21 patients: death at 23 years[13]). The median age at death due to respiratory causes was 1.5 years (IQR 0.4–7 years), compared to 28 years (8, 28, and 58 years) for patients with non-pulmonary causes. Autopsies were performed in a third of these patients (7/21 patients; see pathological postmortem results in Table 13 for the brain, Table 17 for the lung, and Table 19 for the thyroid).

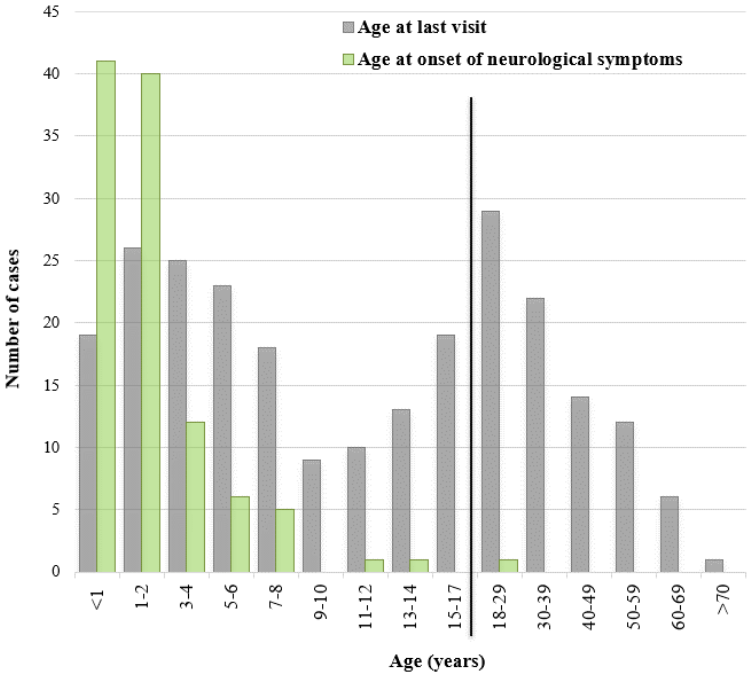

Figure 17. Overall age distribution and age at onset of neurological symptoms. *The gray bars indicate the number of patients with a certain age at last visit or death, and the green bars show the number of patients with known age at onset of neurological symptoms. The black line separates children/teenagers from adults at the age of 18 years.*

3.5.4 Genotypes

Overall, after standardization for HGVS nomenclature (REFSEQ isoform 1; see Figure 1), 167 different variants of the *NKX2-1* gene, including its complete deletion were identified (see Tables 25–28 for a complete list of all variants). In two patients, two concomitant non-WGD variants[13,26,66] were found and therefore counted in multiple variant categories (see Table 26, Part 2 for Patient 132; Table 27, Part 1–2 for Patient 253; and Table 28, Part 1–2 for Patient 334).

3.5.4.1 Mode of inheritance

Among patients with information on the mode of inheritance (90%, 317/351 patients)[13,17,20,26,45-47,53,55,61-68,181-184,188-194,196,199-202,204-226,228-232,236,238-243,245-254], the majority had familial variants (73%, 230/317 patients)[13,14,26,46,53,61-65,68,182,184,188-194,201,205-223,226,228,230-232,236,238,240,242,243,248-250,252,254] compared to *de novo* variants (27%, 86/317 patients)[13,14,17,20,26,45-47,55,60,61, 63,66,67,181,183,184,196,199-202,204,206,208,216,224,225,229,236,239,241,244-247,251,253]. Moreover, one patient had two concomitant variants: one with familial inheritance and one de novo variant (Patient 132; Table 26, Part 2)[26]. There were also two cases of parental mosaicisim (1%, 2/317 patients; Patients 50 and 51; Table 25, Part 2)[208].

3.5.4.2 Types of variants

Two categories of variants can generally be differentiated: 1) non-WGD variants, constituting the large majority of cases (70%, 116/166 variants; 82%, 289/351 patients) and 2) WGD[15,16,28,54,61,62,64,68,69,167,169,170,175,176,184,192-194,202,218,224,225,228,231], accounting for the minority (30%, 50/166 variants; 18%, 62/351 patients). Figure 18 details the locations of non-WGD variants on the *NKX2-1* gene and Table 10 lists the distribution of non-WGD variants. The largest sub-group among all patients comprises missense variants (33%, 115/351 patients; 21%, 35/167 variants), followed by WGD, nonsense variants (16%, 57/351 patients; 13%, 21/167 variants) and other deletions (16%, 56/351 patients; 17%, 28/167 variants).

Table 10: Non-WGD variants: variant types, involved functional domains and phosphorylation sites

	Patients n (%)	Variants n (%)	References
All non-WGD	289 (100%)	116 (100%)	15,16,19,22,28,46-48,56,61-66,168,174,177-180,182,186-192,195-223,225,226,228,230,232-240
Types of non-WGD variants			
Missense	115 (40%)	35 (30%)	15,16,19,28,47,61,62,64,174,177-179,182,186,187,189,190,202,203,205,206,219,220,228,230,232-234,240
AOV	174* (60%)	81 (70%)	13,14,20,26,45-47,55,60-66,182,188,194,196,201,205,206,209-216,218,221-232,235-237,243-247,251-255
Nonsense*	57 (20%)	21 (18%)	15,16,28,47,56,62-64,67,180,187,192,195,196,204,207,209,210,221,222,227,237,238
Other deletions*	56 (20%)	28 (24%)	15,16,28,62,64,65,168,174,182,187,188,191,192,200,202,208,216-218,222,225,229
Splice site variants*	35 (12%)	11 (9%)	14,26,60,61,65,66,211-213,225-228
Duplications*	17 (6%)	12 (10%)	15,16,22,28,46,48,62,222,223,226,241
Insertions*	5 (2%)	5 (4%)	26,46,236,245,246,254
Stop loss*	3 (1%)	2 (3%)	63,215
Indels*	2 (1%)	2 (2%)	26,229
Overall functional domain involvement			
All non-WGD	281 (97%)	111 (96%)	13,14,17,20,26,45-47,55,60,61,63-66,182,188,191-194,196,200-202,204,209-214,216-238,240-244,246-249,251-255
HD (overall)	225 (78%)	79 (68%)	13,14,17,20,26,45,46,55,60-66,188,191-194,196,200-202,206,209-212,214,216-229,233-238,240-245,251,252,255
C-TAD (overall)	174 (60%)	79 (68%)	16,22,28,46-48,56,62-67,168,174,180,182,188,192,195-198,200,204,207-212,215-218,221-223,226,227,233-235,237-239,241228
NK2-SD (overall)	147 (51%)	68 (59%)	14,20,26,45,46,55,60-66,188,194,201,202,209,210,212,214,216,218,221-226,228,229,235,237,243-246,251,252,255
N-TAD (overall)	53 (18%)	27 (23%)	13,14,20,26,45,46,55,61,63,66,201,202,204,206,216,222-224,235-237,243,251,252,255
No domain	8 (2%)	5 (4%)	63,203,205,215
Number of functional domains involved			
One domain (all)	137 (47%)	46 (40%)	13,14,17,46,47,60,61,63,182,188,191-193,200,201,204,216,217,219,220,230-234,236,238,240-242,246-249,253,254
HD-only**	99 (34%)	26 (22%)	13,14,26,60,61,188,191-193,196,200,201,217,219,220,233,234,240-242
C-TAD-only**	30 (10%)	14 (12%)	13,26,47,63,182,230-232,247-249,253,254
N-TAD-only**	5 (2%)	3 (3%)	187,190,202,222
NK-2 SD-only**	3 (1%)	3 (3%)	64,187,236
2: NK2-SD, C-TAD	18 (6%)	12 (10%)	13,14,26,61,188,196,206,216,236,245
3: NK2-SD, C-TAD, HD	78 (27%)	29 (25%)	13,14,26,46,60-65,194,201,209-212,214,218,221,225-229,244
4: NK2-SD, C-TAD, HD, N-TAD	48 (17%)	24 (21%)	13,14,20,26,45,46,55,61,63,66,202,206,222-224,235-237,243,251,252,255
Overall phosphorylation site involvement			
Overall (all)	164 (57%)	74 (64%)	13,14,20,26,45-47,55,60-66,182,188,194,196,201,202,206,209-214,216,221-232,235-237,243-245,247,248,251-253,255
c.850_852;p.284 (located in NK2-SD)	137 (47%)	60 (52%)	13,14,20,26,45,46,55,60-66,194,201,202,206,209-214,216,218,221-229,235-237,243-245,252,255
c.1069_1071;p.357 + c.1096_1098;p.366 (both in the C-TAD)	164 (57%)	74 (64%)	13,14,20,26,45-47,55,61-66,182,188,194,196,201,202,206,209-214,218,221-232,235-237,243-245,247,248,251-253,255

*Abbreviations and symbols: *These variants are part of the "AOV" group. **These variants are part of the "One domain (all)" group. AOV = all other variants (i.e. not WGD or missense), C-TAD = C-terminal transactivation domain, HD = homeodomain, n = number, NK2-SD = NK2-specific domain, N-TAD = N-terminal transactivation domain.*

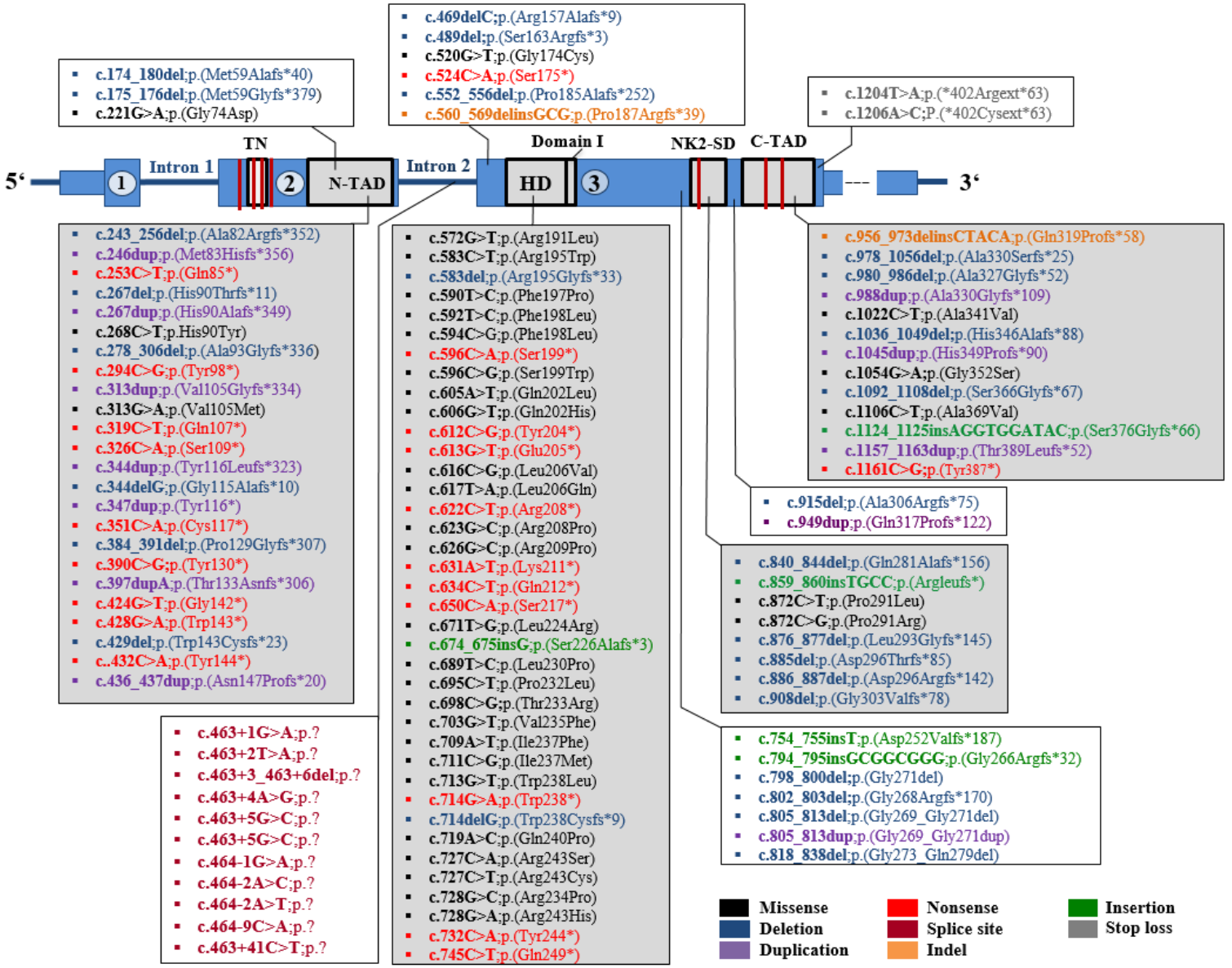

Figure 18. Non-whole-gene-deletion variants (Non-WGD). *The figure portrays the NKX2-1 gene with its three exons (indicated by numbers in the circles), two introns, the functional domains (gray boxes) and the seven phosphorylation sites (red lines). Variants located within a domain are shown in gray boxes, and variants located outside of those domains are shown in white boxes. The respective variants are listed together in the boxes, and the types of variants are denoted by different font colors (see legend in the right lower corner). Abbreviations: C-TAD = C-terminal transactivation domain, HD = homeodomain, NK2-SD = NK2-specific domain, N-TAD = N-terminal transactivation domain, TN = Tinman-domain.*

Splice site variants (10%, 35/351 patients; 7%, 11/167 variants) and duplications (5%, 17/351 patients; 7%, 12/167 variants) were detected less frequently, while insertions (1%, 5/351 patients; 3%, 5/167 variants), stop loss variants (1%, 3/351 patients; 1%, 2/167 variants), and indels (1%, 2/351 patients; 1%, 2/167 variants) were rarely found.

While WGD naturally cause loss of the whole gene and thus all functional domains, a wide variety of non-WGD variants can be found with differing extents of impairments of the *NKX2-1* gene. Figure 19 depicts the distribution of all types of variants depending on gender and mode of inheritance. As Figure 19a indicates, the comparison of males and females revealed

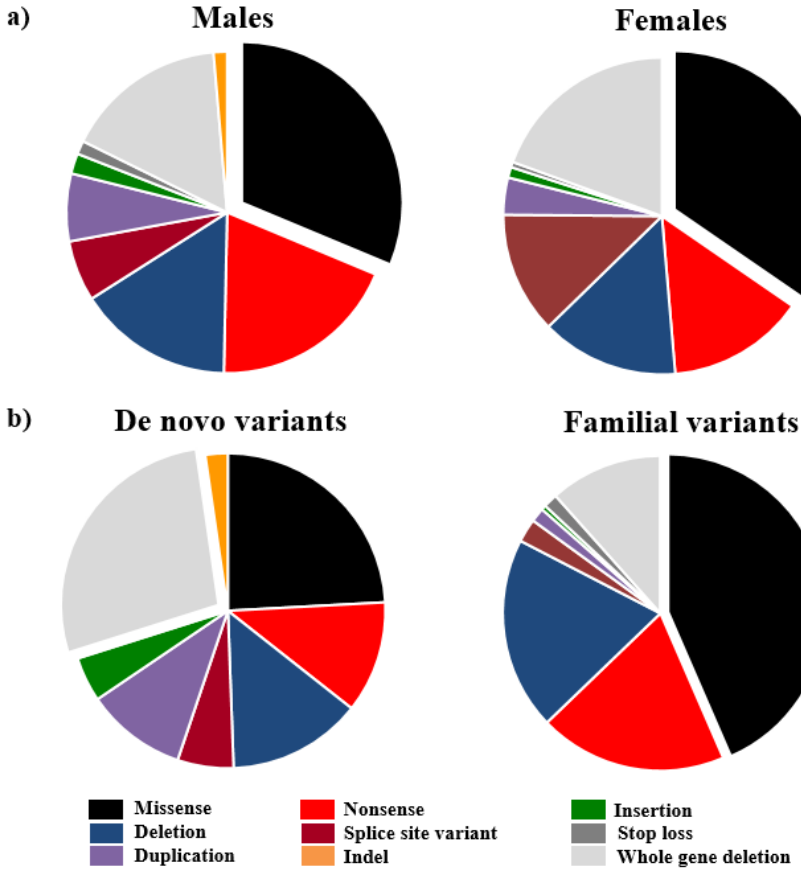

Figure 19. Distribution of types of variants depending on gender and mode of inheritance. *The diagrams are based on the number of patients with the respective variants. The numerous types of variants are coded in different colors (see legend at the bottom). a) Types of variants in 147 males and 185 females. As illustrated, missense variants were the leading group in both genders (males: 31%, 46/147 patients versus females: 35%, 64/185 patients), followed by nonsense variants in males (19%, 28/147 patients compared to females: 14%, 26/185 patients) and WGD in females (19%, 36/185 patients compared to males: 16%, 24/147 patients). The remaining variants were duplications (males: 7%, 10/147 patients versus females: 4%, 7/185 patients), splice site variants (males: 6%, 9/147 patients versus females: 12%, 23/185 patients), insertions (males: 2%, 3/147 patients versus females: 1%, 2/185 patients), indels (males: 1%, 2/147 patients versus females: 0%, 0/185 patients), and stop loss variants (males: 1%, 2/147 patients versus females: 1%, 1/185 patients). b) Types of variants in 86 patients with de novo and 230 with familial variants are shown. The major difference was that the vast majority of familial variants were missense variants (39%, 90/230 patients versus de novo: 24%, 21/86 patients), whereas the leading group in de novo variants was WGD (de novo: 28%, 24/86 patients versus familial: 10%, 24/230 patients). Moreover, in de novo variants, duplications (de novo: 10%, 9/86 patients versus familial: 1%, 3/230 patients), splice site variants (de novo: 6%, 5/86 patients versus familial: 2%, 5/230 patients), insertions (de novo: 5%, 4/86 patients versus familial: 0.4%, 1/230 patients), and indels (de novo: 2%, 2/86 patients versus familial: 0%, 0/230 patients) together accounted for approximately one quarter of cases (23%, 20/86 variants), while they were rare in familial variants. No large differences were found for other deletions (de novo: 14%, 12/86 patients versus familial: 18%, 41/230 patients) and nonsense variants (de novo: 12%, 10/86 patients versus familial: 18%, 40/230 patients).*

only slight differences, in particular for WGD (*males*: 16%, 24/147 patients; *females*: 19%, 36/185 patients), AOV (*males*: 52%, 77/147 patients; *females*: AOV 46%, 85/185 patients), and missense (*males*: 32%, 46/147 patients; *females*: 35%, 64/185 patients) variants. In contrast, there were considerable differences between patients with *de novo* and familial variants (see Figure 19b). While patients with familial cases for the most part carried missense variants (familial: 39%, 90/230 patients; *de novo*: 24%, 21/86 patients), in *de novo* cases, WGD (*de novo*: 28%, 24/86 patients; *familial*: 10%, 24/230 patients) was leading, followed by duplications (*de novo*: 10%, 9/86 patients; *familial*: 1%, 3/230 patients) and splice site variants (*de novo*: 6%, 5/86 patients; *familial*: 2%, 5/230 patients), which were rare in familial variants. In AOV, *de novo* (48%, 41/86 patients) and familial variants (50%, 116/230 patients) were found in equal abundance.

3.5.4.3 Additional variants of other genes

In 16% of patients (56/351) additional variants of overall 136 other genes were found[26,46,53,60,61,63,67,68,181,183,184,188-190,199,206-208,214,216,232,239,250,254]. The following genes were most frequently additionally deleted: *PAX9* (11%, 37/351 patients)[26,46,60,61,63,67,68,181,184,188-190,199,207,208,216,232,239,250], *NKX2-8* (8%, 27/351 patients)[26,46,61,63,67,68,188,199,207,208,216,239], *RALGAPA1 (*synonym *GARNL1;* 6%, 20/351 patients)[26,46,60,61,63,67,181,184,189,190,239], *SFTA3* (6%, 20/351 patients)[26,46,61,63,199,208,239], (6%, 20/351 patients)[26,46,61,63,67,68, 188,199,207,208,239], *MBIP* (5%, 16/351 patients)[26,46,61,63,67,188,199, 207,208,239], and *MIPOL1* (4%, 14/351 patients)[26,46,63,67,199,208,216,239]. Additional variants of other genes were primarily found in WGD (93%, 52/56 patients). However, they were rarely found in non-WGD (7%, 4/56 patients): in three patients with the deletion of c.469delC; p.(Arg157Alafs*9)[214] (Patients 278–280; Table 27, Part 3) and duplication of 16 additional genes, as well as in the insertion of c.1124_1125 p.(insAGGTGGATAC;pSer376Glyfs*66)[254] found in Patient 346 (Table 28, Part 2), with an additional missense variant in the SFTPC gene.

3.5.4.4 Whole gene deletions (WGD)

The majority of patients with WGD were found to have contiguous gene deletion syndromes (84%, 52/62 patients; 83%, 40/48 variants). The median deletion size of these variants was 3.3 Mb (IQR 0.9–11.4 Mb; information available in 93% of patients, 48/52 patients and 90% of variants, 36/40 variants), and the median number of additionally deleted genes was six (IQR 3–11.8; max. 48 genes in one patient). An analysis of cases with available genomic coordinates and those that could be identified (not possible if only cytobands were available in the original article) revealed additional proximal deletions ranging from the chromosomal coordinates hg19 chr14: 30631931 to 36.985.601 and additional distal deletion ranging from chromosomal coor-

dinates hg19: chr14:36.990.3544 to 50031170. Combined additional proximal and distal deletions to varying extents were found in four patients[26,46,60,184] (6%, 4/62 patients; 8%, 4/48 variants; Patients 11–12 and 30–31; Table 25, Part 1), and additional distal deletion without proximal deletion was found in 28 cases[26,61,63,67,184, 188,190,199,208,239] (45%, 28/62 patients; 50%, 24/48 variants; Patients 32–36, 38–47, 50–54 and 59–62; Table 25 Parts 1–2).

3.5.4.5 Non-WGD variants

3.5.4.5.1 Involvement of functional domains and phosphorylation sites

As mentioned earlier, Figure 18 illustrates the locations of all non-WGD variants, and Table 10 lists the statistics regarding the involvement of functional domains and phosphorylation sites. As presented in Figure 18, the vast majority of non-WGD variants are located on exon 3 (66%, 76/116 variants; 69%, 201/289 patients) and less frequently on exon 2 (24%, 28/116 variants; 18%, 53/289 patients) and intron 2 (10%, 11/116 variants; 12%, 35/289 patients). Interestingly, no variants were found on exon 1 or intron 1.

Closer examination of non-WGD variants revealed that nearly all of non-WGD variants involve at least one functional domain (96%, 111/116 variants; 97%, 281/289 patients). Among them, the HD and the C-TAD are affected in the majority of cases (both 68%, 79/116 variants), followed by the NK2-SD (59%, 68/116 variants) and the N-TAD (23%, 27/116 variants). The affected domains are also indicated for each variant in Tables 26–28. Moreover, non-WGD variants were found to impair the following three out of seven phosphorylation sites: 1) c.850_852;p.284 (52%, 60/116 variants), which is located in the NK2-SD domain, and 2) c.1069_1071;p.357 and 3) c.1096_1098;p.366 (64%, 74/116 non-WGD variants), which are both located in the C-TAD.

Furthermore, domain involvement was predominantly singular (overall 40%, 46/116 variants): either HD-only (22%, 26/116 variants), C-TAD-only (12%, 14/116 variants), N-TAD-only (3%, 3/116 variants), or NK2-SD-only (3%, 3/116 variants). Other combinations were as follows: two domains (NK2-SD and C-TAD: 10%, 12/116 variants), three domains (HD, NK2-SD and C-TAD: 25%, 29/116 variants), and four domains (HD, NK2-SD, C-TAD, and N-TAD: 21%, 24/116 variants). Interestingly, no variants were found to involve Domain I, the TN (on exon 1) or the four phosphorylation sites located in or adjacent to the TN (c.100_102;p.34, c.124_126;p.42, c.142_144;p.48, and c.157_159;p.53).

A subanalysis revealed that nearly all missense variants are single-domain variants (98%, 113/115 patients; 95%, 33/35 variants; see also Figure 27), which are mainly HD-only (86%, 99/115 patients; 74%, 26/35 variants). In contrast, AOV are rather multi-domain variants (83%,

144/174 patients; 80%, 65/81 variants). Among them, the C-TAD was found to be involved in 94% (164/174 patients; 96%, 78/81 variants), the NK2-SD in 83% (144/174 patients; 81%, 66/81 variants), the HD in 72% (126/174 patients; 68%, 55/81 variants), and the N-TAD in 29% (51/174 patients; 32%, 26/81 variants).

3.5.4.5.2 The gene regions c.96–565 and c.753–881

One third of the non-WGD variants (35%, 99/289 patients, 38%, 44/116 variants) are located in the gene region c.96–565 (see Figure 20a) and largely belong to the non-HD+ group (93%, 92/99 patients; 89%, 39/44 variants; shown in gray in Figure 20a), while the remaining minority are non-HD variants (7%, 7/99 patients; 11%, 5/44 variants; shown in white in Figure 20a). In comparison, the gene region c.753–881 comprises 6% of the non-WGD variants (18/289 pat-

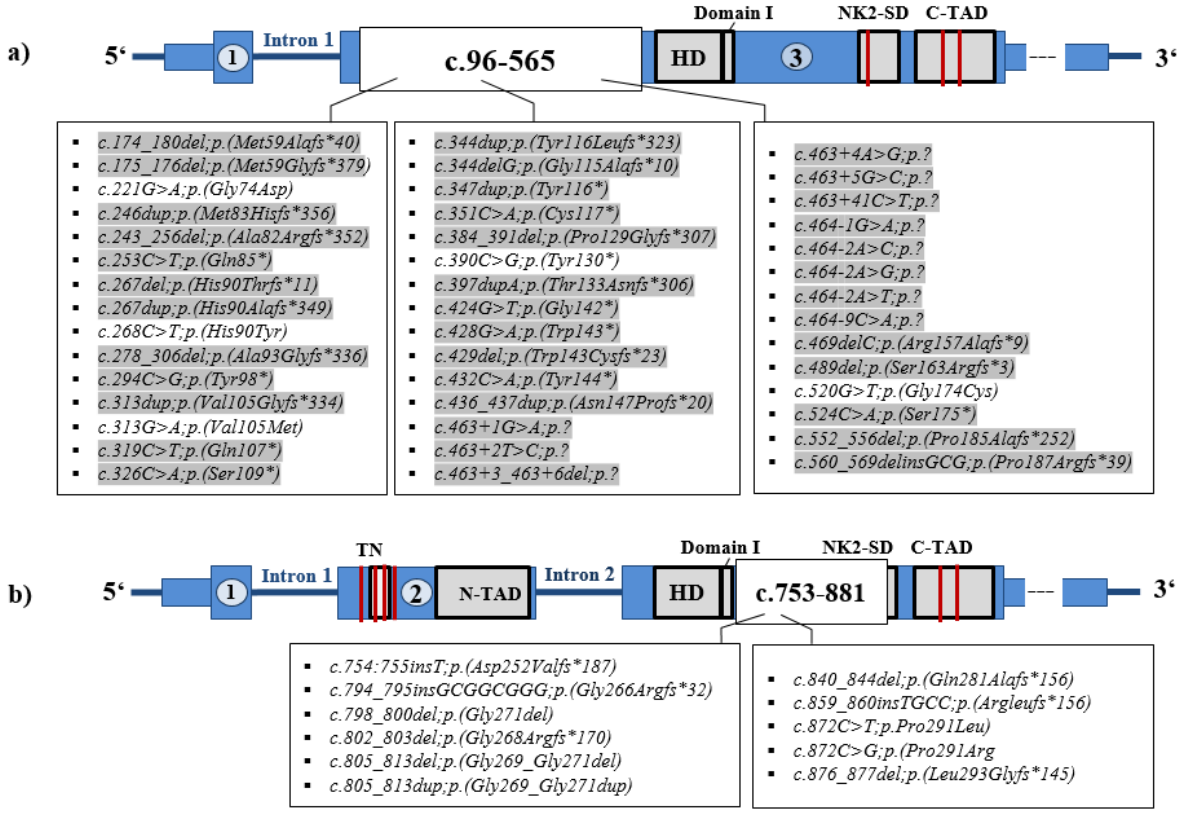

Figure 20. Variants located in the gene regions c.96–565 and c.753–881. *The schematic drawings depict the NKX2-1 gene and its different functional domains (gray boxes) and seven phosphorylation sites (red bars) for isoform 1 (NM_001079668.2). The exons and introns are shown in blue (exons indicated by circles with the numbers 1–3). The gene regions c.96–565 (a) and c.753–881 (b) are indicated by white boxes over the Nkx2-1 gene, and the variants located in the respective regions are listed in white boxes below. Variants involving at least one functional domain are marked with a gray background, and those without domain involvement have a white background. As illustrated in a), the gene region c.96–565 (also known as the transactivation site 1; see also ΔNH2 mutation in mice in Figure 3) encompasses the TN; the phosphorylation sites c.100_102;p.34, c.124_126;p.42, c.142_144;p.48, and c.157_159;p.53; and the N-TAD. Forty-four variants found in 99 patients are located in this region. The region c.753–881 (also known as the transactivation domain 2; see also ΔCOOH mutation in mice in Figure 3) comprises parts of the NK2-SD and the phosphorylation site c.850_852;p.284. Eleven variants found in 18 patients are located here. Abbreviations: C-TAD = C-terminal transactivation domain, HD = homeodomain, NK2-SD = NK2-specific domain, N-TAD = N-terminal transactivation domain, TN = Tinman-domain.*

ients and 9%, 11/116 of non-WGD variants; see Figure 20b), which are all non-HD variants (100%, 18/18 patients; 100%, 11/11 variants).

3.5.4.5.3 Effects of non-WGD variants on gene function

Information regarding the effects of non-WGD variants on gene function was available in less than half of the cases (46%, 53/116 variants; 40%, 117/289 patients; see Table 11 for details). Among them, functional impairments were identified in 75% (40/53 variants). The findings were derived from a variety of different functional studies that are summarized in Table 11. As the impairments were only detected whenever respective tests were performed, the numbers given here are not representative of either the whole patient population or the cases of this subanalysis. Therefore, only absolute and relative numbers for the studied variants are provided. A subanalysis of domain involvement revealed that all variants with reported functional impairments did at least impair one functional domain (100%, 40/40 variants; see Figures 21 and 22 for details). As demonstrated in Table 11 and Figures 21 and 22, the variants were mostly multi-domain variants (63%, 25/40 variants). Isolated domain variants were encountered less frequently: HD-only in 28% (11/40 variants), C-TAD-only in 8% (3/40 variants), and N-TAD-only in 3% (1/40 variants). The reason no NK2-SD-only variants were affected is that they were not investigated by the authors.

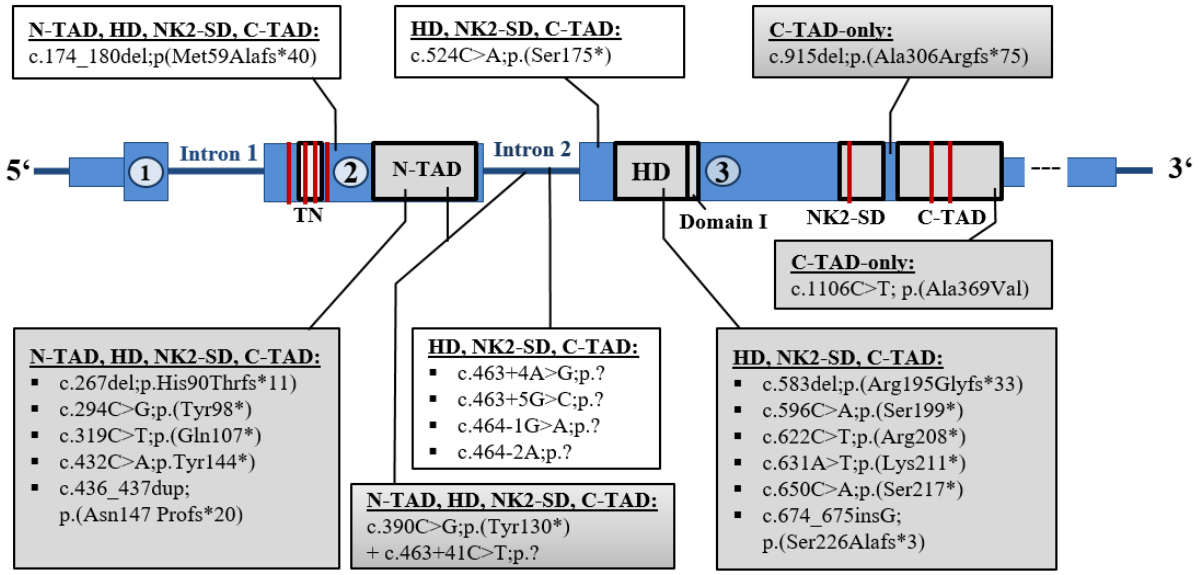

Figure 21: Protein-truncating variants. *The schematic drawing depicts the NKX2-1 gene and its different functional domains (gray boxes) and seven phosphorylation sites (red bars) for isoform 1 (NM_001079668.2). The exons and introns are shown in blue (exons indicated by circles with the numbers 1–3). Variants located in a domain are shown in gray boxes, and those outside of domains are in white boxes. The domains that are affected by the variants are presented in the boxes, and the respective variants are listed below them. As can be seen from this figure, all protein-truncating variants involve at least one functional domain, which is always the C-TAD. The variants are predominantly located in functional domains, and the majority of variants involve three or four domains. Abbreviations: C-TAD = C-terminal transactivation domain, HD = homeodomain, NK2-SD = NK2-specific domain, N-TAD = N-terminal transactivation domain, TN = Tinman-domain.*

Table 11: Impairments of gene function in non-WGD variants with available information

	Involved domains				Single-domain variants			Overall domain impairments				All
	4	3	2	1	HD-only	N-TAD-only	C-TAD-only	HD all	N-TAD all	NK2-SD all	C-TAD all	
Variants (n)	**n (%)**	**n (%)**	**n (%)**	**n (%)**	**n (%)**	**n (%)**	**n (%)**	**n (%)**	**n (%)**	**n (%)**	**n (%)**	**n (%)**
Information on gene function available[17,20,26,45,46,55,60,61,64,66,182,194,196,202,204,212,216-219,222-224,227,230,233,237,249,251-253,255]	18 (34%)	13 (24%)	5 (9%)	17 (32%)	11 (21%)	2 (4%)	4 (8%)	42 (79%)	20 (38%)	38 (72%)	42 (79%)	53 (100%)
Variants with functional impairments found[17,20,26,45,46,55,60,61,64,66,182,194,196,202,204,212,219,222-224,227,230,233,237,249,251-253,255]	11 (28%)	12 (30%)	2 (5%)	15 (75%)	11 (28%)	1 (3%)	3 (8%)	34 (85%)	12 (30%)	25 (63%)	28 (70%)	40 (100%)
Protein truncation[13,26,46,55,62-64,66,182,194,202,206,210,212,218,222-224,227,230,235,237,244,247,249]	8 (38%)	11 (53%)	-	2 (10%)	-		2 (10%)	19 (90%)	8 (38%)	19 (90%)	21 (100%)	21 (100%)
Impaired DNA binding[14,17,20,26,45,46,55,64,182,196,218,230,233,235,238,247,255]	3 (16%)	4 (21%)	1 (5%)	11 (58%)	8 (42%)		3 (16%)	15 (79%)	3 (16%)	8 (42%)	11 (58%)	19 (100%)
Impaired promoter binding[14,20,26,55,60,61,64,182,196,202,218,230,233,235,247,249,253,255]	3 (21%)	2 (14%)	1 (7%)	8 (57%)	5 (36%)	-	3 (21%)	10 (71%)	3 (21%)	6 (43%)	9 (64%)	14 (100%)
Impaired migration to the nucleus[168,180,188,203,216,229,238]	1 (25%)	1 (25%)	-	2 (50%)	1 (25%)		1 (25%)	3 (75%)	1 (25%)	2 (50%)	3 (75%)	4 (100%)
Dominant negative effect[14,20,26,55,60,61,64,182,196,219,230,235,247,255]	2 (22%)	1 (11%)	1 (11%)	5 (56%)	4 (44%)		1 (11%)	7 (77%)	2 (22%)	4 (44%)	5 (56%)	9 (100%)
Loss of function[26,46,55,60,182,196,202,230,247]	5 (42%)	3 (25%)	-	4 (33%)	2 (17%)	1 (8%)	1 (8%)	10 (83%)	6 (50%)	8 (67%)	10 (83%)	12 (100%)
Impaired protein synthesis[64,182,202,230,247,251]	2 (50%)	1 (25%)		1 (25%)	-		1 (25%)	3 (75%)	2 (50%)	3 (75%)	4 (100%)	4 (100%)
Impaired synergy with *PAX8* on the TG promoter[16,62,182]	1 (17%)	-	1 (17%)	4 (67%)	3 (50%)	-	1 (17%)	4 (67%)	1 (17%)	2 (33%)	3 (50%)	6 (100%)
PAX8 interaction could rescue promoter activation[20,230]	-		-	1 (100%)	-		1 (100%)	-	-	-	1 (100%)	1 (100%)
Intact *NKX2-1* could partially rescue the transcriptional activity (patient with simultaneous *NKX2-1* and *PAX8* variant)[204]				1 (100%)		1 (100%)	-		1 (100%)		-	1 (100%)
TAZ coexpression could rescue promoter activation[168,216,229]				1 (100%)		-	1 (100%)		-		1 (100%)	1 (100%)

Abbreviations and explanations: 1 = one domain (i.e. the sum of HD-only, N-TAD-only, and C-TAD-only); 2 = NK2-SD and C-TAD; 3 = HD, NK2-SD, and C-TAD; 4 = N-TAD, HD, NK2-SD, and C-TAD; C-TAD = C-terminal transactivation domain; HD = homeodomain; NK2-SD = NK2-specific domain; N-TAD = N-terminal transactivation domain; TAZ = tafazzin; TG = thyreoglobulin. Domain involvements that were found in all variants with the respective impairment are highlighted by gray backgrounds.

Overall, the majority of variants with impairments were found to involve the HD (85%, 34/40 variants), the C-TAD (70%, 28/40 variants), and the NK2-SD (63%, 25/40 variants). Among all impairments protein truncation was the most frequent one with 39% (21/53 variants; 48%, 58/117 patients; see Table 11 and Figure 21). Interestingly, all protein-truncating variants were found to impair the C-TAD (see Figure 21) and thus also the phosphorylation sites c.1069_1071;p.357 and c.1096_1098;p.366 (100%, 21/21 variants), which are located in the C-TAD. In addition, they vastly involved the HD and the NK2-SD domain, including its phosphorylation site c.850_852;p.284 (both 90%, 19/21 variants). Protein-truncating variants were mostly nonsense variants (33%, 7/21 variants)[26,46,62,210,218], deletions (19%, 4/21 variants)[64,202,222,230], or splice site variants (19%, 4/21 variants)[26,66,212,227]. Insertions (5%, 1/21

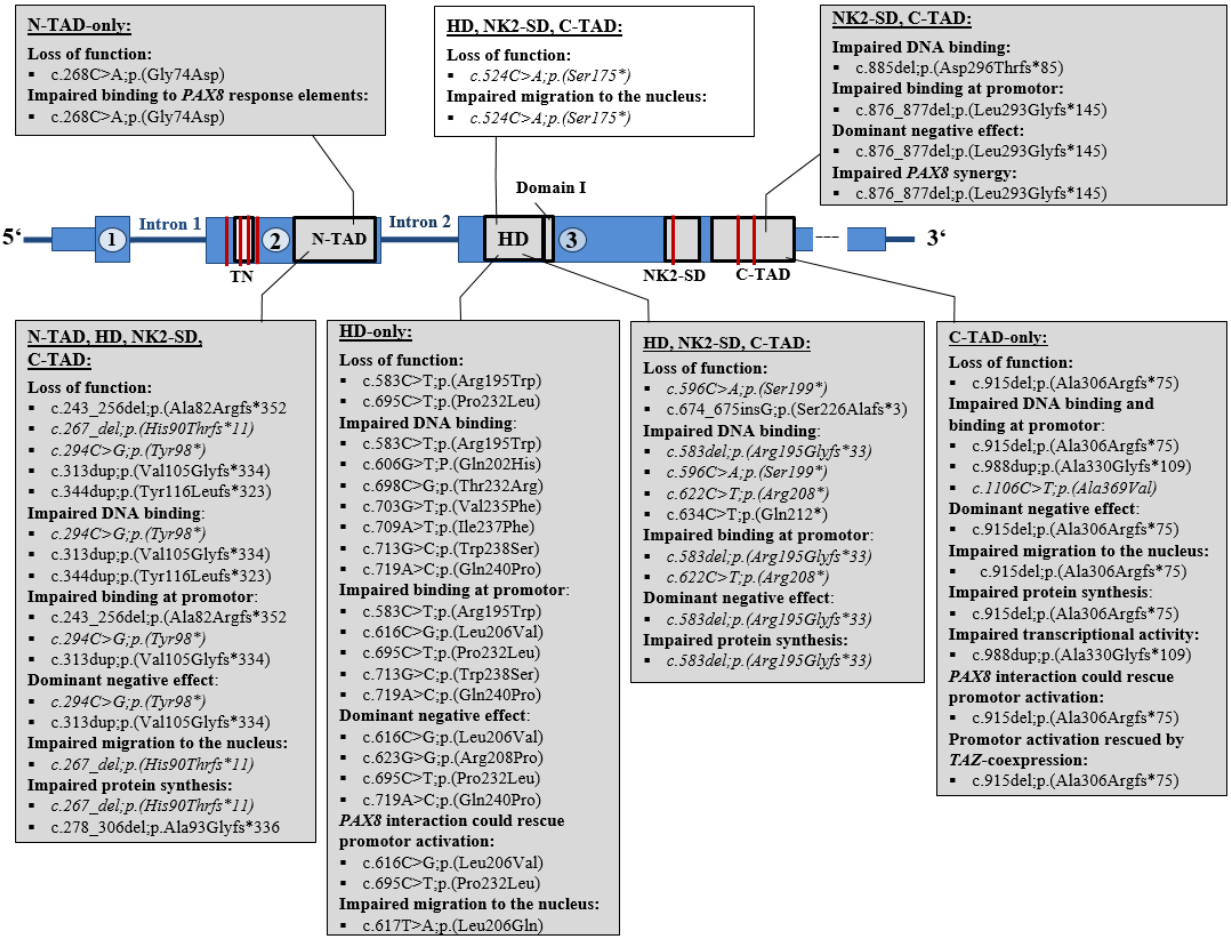

Figure 22. Results of functional studies regarding other effects on gene function. *The schematic drawing shows the NKX2-1 gene and its different functional domains (gray boxes) and seven phosphorylation sites (red bars) for isoform 1 (NM_001079668.2). The exons and introns are indicated in blue (exons denoted by circles with the numbers 1–3). Variants located in a domain are shown in gray boxes, and those outside of domains are in white boxes. The various impairments are indicated in bold letters in the boxes, and the variants causing the respective impairments are listed below them. Protein-truncating variants (cf. Figure 21) are shown in italics. It can be seen here that no impairments were found in variants that did not involve at least one functional domain. In contrast to protein-truncating variants, functional impairments were also described in HD-only and N-TAD-only variants. Abbreviations: C-TAD = C-terminal transactivation domain, HD = homeodomain, NK2-SD = NK2-specific domain, N-TAD = N-terminal transactivation domain, TAZ = tafazzin, TN = Tinman-domain.*

variants)[26,46], duplications (5%, 1/21 variants)[237], and missense variants (5%, 1/21 variants)[249] were rarely reported to cause protein-truncation. The C-TAD was further involved in all variants with impaired protein synthesis (100%, 4/4 variants; of these, 100% deletions, 4/4 variants)[64,182,202,230,247,251] and many of the other functional impairments (as highlighted in Table 11 and Figure 22). The second highest overall involvement of the HD (see Table 11 and Figure 22) was found in variants with loss of function (83%, 10/12 variants; of these 30% missense variants, 3/10 variants[60,196,204]), followed by impaired DNA binding (79%, 15/19 variants; of these, 46%, 7/15 missense variants[17,26,60,61,196,218]), dominant negative effect (77%, 7/9 variants; of these, 44% missense variants, 4/7 variants[60,61,219]), impaired protein synthesis (75%, 3/4 variants; of these, 100% deletions, 3/3 variants[64,202,251]), and impaired migration to the nucleus (75%, 3/4 variants; of these, 33% missense variants, 1/3 variants). Overall, in a subset of different types of variants, loss of function (23%, 12/53 variants) and dominant negative effects (17%, 9/53 variants) were described.

Furthermore, impaired synergy with *PAX8* on the thyreoglobulin (TG) promoter was found in the following variants: c.915del;p.(Ala306Argfs*75), impairing four domains (Patients 318–321; Table 28, Part 1)[20]; c.876_877del;p.(Leu293Glyfs*145) variant with C-TAD and NK2-SD involvement (Patients 311–2; Table 28, Part 1)[196]; c.313dup;p.(Val105Glyfs*334) with C-TAD-only involvement[20] (Patient 212; Table 27, Part 1); c.583C>T;p.(Arg195Trp)[196] (Patients 71– 73; Table 26, Part 1); c.616C>G;p.(Leu206Val)[60,61] (Patient 96; Table 26, Part 1); c.695C>T; p.(Pro232Leu) (Patient 131; Table 26, Part 2)[60] – the latter three each with HD-only involvement. Only in the c.915del;p.(Ala306Argfs*75) variant (Patients 318–9; Table 28, Part 1)[20], the presence of *PAX8* could partially rescue the activity on the TG promoter[20,60,61,196,230]. In addition, in one patient with a simultaneous *PAX8* variant and the *NKX2-1* variant c.268C>T;p.(His90Tyr), the absence of an intact *NKX2-1* gene was shown to result in a lack of *PAX8* transcriptional activity, whereas the presence of a wild-type *NKX2-1* gene partially rescued the *PAX8* transcriptional activity[204] (Patient 289; Table 28, Part 1). Moveover, the transcriptional co-activator *Tafazzin* (*TAZ*), which is co-expressed with *NKX2-1*, could rescue the transcriptional activity in the c.915del;p.(Ala306Argfs*75) variant (Patients 318–321; Table 28, Part 1)[168,216,229].

3.5.5 Phenotypes

3.5.5.1 Combination of organ manifestations

An analysis of cases with information on all three organs (298/351 patients) revealed a typical three-organ manifestation in 39% of cases (115/298 patients). The different combinations of organ manifestations and the cases lacking data on the thyroid gland (indicated in gray) are

illustrated in Figure 23. The second most frequent multi-organ combination was a brain/thyroid phenotype (28%, 82/298 patients), followed by a brain/lung phenotype (6%, 17/298 patients), and a lung/thyroid phenotype (4%, 13/298 patients). The remaining patients had single-organ phenotypes: 14% brain-only (43/298 patients), 5% lung-only (14/298 patients), and 5% thyroid-only (14/298 patients).

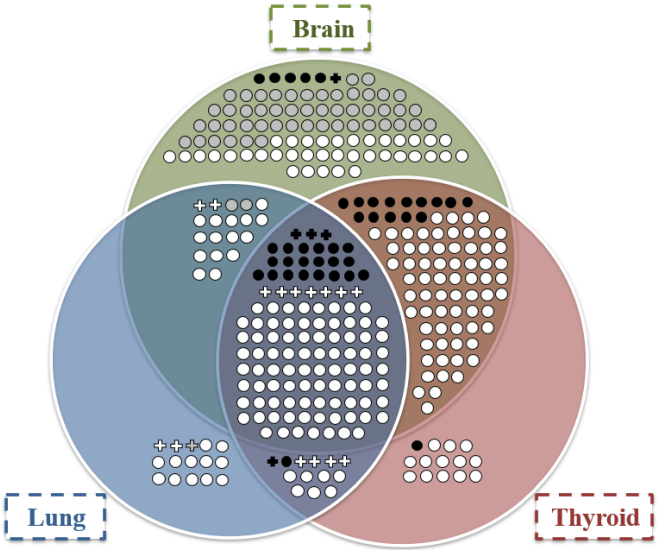

***Figure 23. Distribution of organ phenotypes in 351 patients**. Black dots and crosses (deaths) denote the 46 patients reviewed by Carré et al. (2009)[60]. White dots and crosses (deaths) denote cases published before 01/2009 but not reviewed by those authors and cases published between 01/2009 and May 2nd 2020, reviewed here. Gray dots and crosses denote cases where information on the thyroid phenotype was unavailable. As illustrated here, the majority of patients had brain involvement either in combination with other organ manifestations or alone, while single lung or thyroid phenotypes were rare.*

3.5.5.2 Brain phenotype

The majority of patients were found to have neurological symptoms (88%, 309/351 patients; see Table 12), which were mainly motor impairments (96%, 297/309 patients). Chorea (overall: 73%, 255/351 patients; without athetosis: 56%, 196/351 patients), general developmental delay (39%, 140/351 patients), and muscular hypotonia (34%, 120/351 patients) were encountered most frequently. The onset of motor symptoms or the diagnosis of a developmental delay was made at a median age of 1.1 years (IQR 0.5–2.5 years; see Figure 17) and seven patients were initially misdiagnosed with cerebral palsy before the *NKX2-1* variants were found[212,222,232,240,247,253]. Patients without neurological symptoms and known age (81%, 34/42 patients) were often infants under the age of four years (41%, 14/34 patients)[13,14,17,67,191,216,217,234,245,248,251]. In a quarter of patients with motor impairments (25%, 75/297 patients), information on the outcome was provided[53,55,60-62,64,182,184,188,193,196,206,208,215,222,223,225,226,228,233,235-237,240,242,243,247]. Among them the majority (53%, 40/75 patients) had an improvement in adolescence or adulthood[55,60-62,182,193,196,206,215,222,223,225,226,228,236,237,242,243,247] while 31% (23/75 patients) had stable symptoms[61,64,182,184,188,193,208,222,223,228,233,247] and 16% had progression (16%, 12/75 patients)[62,193,206,233,236,243,246]. Overall, five patients required a wheelchair[181,223,252], nine needed walking assistance intermittently[194,212,217,236], and five required leg orthotics[194] during childhood. Cognitive impairments were present in 29% of patients (101/351 patients; see Table 12)[13,14,20,26,46,47,55,60,61,64,181-

[184,189,193,194,196,199,202,206,208,213,214,216,217,219,221-223,225,229,230,232,235,239,241 -243,247,253], which mainly involved learning difficulties including attention deficit hyperactivity disorder (ADHD; 15%, 54/351 patients) and severe intellectual disability (12%, 41/351 patients). In contrast, psychiatric disorders were rare (3%, 9/351 patients).

Table 12: Neurological disorders and symptoms

Motor impairments/symptoms	n	n
Chorea		
Chorea without athetosis[13,14,46,47,53,55,60-65,68,182,188,193,194,196,199,201,206-210,214,215,217,218,220-226,231,233,235-237,241-244,247]	196	257
Choreoathetosis[20,26,45,46,60,62,65,66,184,194,198,202,208-212,219,222,228-230,232,238,240,246]	59	
Muscular hypotonia[13,14,26,45,47,55,60,61,63,66,68,181,184,189,193,196,201,202,206,209,214,215,220,225-227,232,236,239-241,243,244,246,247,250,252]		120
Muscular dystonia[26,46,60-64,193,196,199,206,210,217,218,223,224,227,233,236,241,243,244,247]		46
Ataxia[13,14,26,46,53,60-64,68,181,182,184,193,199,206,208,209,212,214,215,218-225,227-229,231,233,235,236,238,239,241-244,246,247]		89
Myocloni[60,61,63,184,189,193,196,199,201,205,214,220,226,232,233,235,236,241,243,250]		34
Spasticity[53,61,68,181,184,189,193,202,243]		17
Speech impairment		
Dysarthria[26,46,61-63,209,210,212,214,217,218,222,235,243]	46	50
Stammering[193,212,217,221]	7	
Apraxia[62,65,193,213,233,242]		9
Tics[61,193,226,257]		6
Intentional tremor[53,63,213,222,242]		5
Orofacial dyskinesia[214,215,235]		4
"Leg weakness"[217]		4
Restless legs syndrome[201]		3
Quadriplegia[14]		1
Other symptoms/disorders		**n**
Seizures[13,217,232,253,254]		7
Sleep disturbances[181]		4
Hypopallaesthesia[68]		3
Cluster headache[26,46]		3
Hearing loss[13,184,189]		3
Cortical blindness[181]		1

Cognitive impairments / developmental delay		n
General developmental delay[13,20,26,45,46,53,55,60,61,63-66,68,181,182,184,189,190,193,196,199,201,205,206,208,211-215,217-220,222,223,225-228,230,232,233,235-239,243,244,246,247,250]		140
Severe intellectual disability[14,26,46,47,55,60,61,64,181,183,184,189, 196,199,202,206,213,216,225,229,230,253]		41
Learning difficulties		
ADHD[15,62,168,170,179,180,191,194,197,201,209,219,221,222,230,231,233]	33	54
Learning difficulties (without ADHD)[61,214,217,221,239,243]	21	
Speech delay[13,53,60,61,196,201,202,208,214,222,236,247]		21
Microcephaly[13,181,183,184,189,243]		7
Short-term verbal/spatial memory deficit[194]		1
Autism[232]		1
Psychiatric disorders		**n**
Psychosis[210,217,243]		3
Depression[217]		2
Anxiety disorder[217]		2
Obsessive compulsive disorder[63]		2
Bipolar disorder[213]		1
Schizophrenia[210]		1
Panic disorder[217]		1
Self-injuring behavior[217]		1
Other symptoms/disorders		**n**
Optic atrophy[183]		1
Ischemic cerebral infarction[253]		1
Recurrent drop attacks[237]		1
Recurrent aseptic meningitis[220]		1
Vein of Galen thrombosis[13]		1
Glioblastoma multiforme[233]		1

Abbreviations: ADHD = attention deficit hyperactivity disorder, n = number of patients.

3.5.5.2.1 Brain autopsies and neuroimaging

In 71% of patients (5/7 patients) with postmortem studies, the brain was investigated. In one patient with a brain phenotype, the parents did not grant permission for brain autopsy[250] (Patient 13; Table 25, Part 1) and in the other, who also had a brain phenotype, brain autopsy was either not performed or no data were provided[196] (Patient 71; Table 26, Part 1).

Among those with brain autopsies, three had neurological manifestations (Patients 240 and 263; Table 27, Part 2; and Patient 324; Table 28, Part 2), and in only the patient with the c.464-2A>T;p.? splice site variant (Patient 263), pathologies were found, namely, loss of ChAT-positive and PV-containing striatal interneurons and efferent fibers, as well as atrophy and abnormal morphology of ChAT-positive striatal interneurons[65] (see Table 13 for details). In addition, one patient with the c.711C>G;p.(Ile237Met) missense variant and a lung/thyroid phenotype (Patient 141; Table 26, Part 2; see also Table 13) who died from respiratory failure had reduced brain weight and signs of hypoxic ischemic encephalopathy[234].

Table 13: Pathological brain autopsy results

Brain phenotype	Brain autopsy result
Isolated brain phenotype: choreoathetosis, clumsiness, developmental delay. Normal brain MRI.[65] (Patient 263)	Loss of ChAT-positive and PV-containing striatal interneurons and efferent fibers, atrophy and abnormal morphology of ChAT-positive striatal interneurons, loss of striatal calbindin-positive interneurons, less intense staining of calretinin-positive striatal neurons, reduced met-enkephalin-positive fibers, and reduced substance P positive fibers
Lung/thyroid phenotype[234] (Patient 141)	Reduced brain weight, indistinguishable from hypoxic ischemic encephalopathy

Abbreviations: ChAT = choline acetyl esterase, MRI = magnetic resonance imaging, PV = parvalbumin.

Neuroimaging was performed in 44% of cases (153/351 patients; see Figure 24) with MRI as the primary modality (85%, 130/153 patients)[13,20,26,45-47,53,55,60-62,64-69,181-184,188,190,193,194,199,201,202,205,207,209,211-214,217,218,220,221,223,225,227-233,238,239,242,243,246,247,250,251, 256]. Alternatively or additionally, computed tomography (CT; 8%, 12/153)[47,62,181,184,189,223,242,254], positron emission tomography (PET; 7%, 11/153 patients)[65,194,228,231], and single photon emission computer tomography (SPECT; 3%, 4/153 patients)[211,238] were used. Abnormal findings were reported in 27% of patients (41/153 patients; see Figures 24 and 25)[13,26,46,60,62,69,181,183,184,189,194,201,207,221,225,228,231-233,238,246,247,250,256].

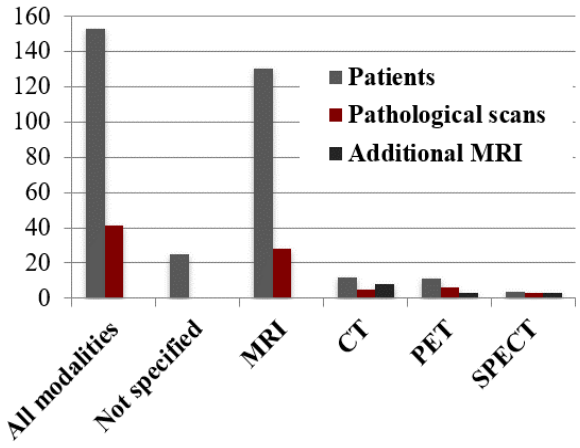

Figure 24. Neuroimaging. *The gray bars show the number of patients who had brain scans (any modality). The red bars denote the number of patients with abnormal findings in the respective categories. The black bars indicate how many of the patients had a combination of an MRI with another modality (CT, PET, or SPECT). As illustrated here, MRI (85%) was the leading modality, and pathologies were found in a minority of patients (27%). Abbreviations: CT = computed tomography, MRI = magnetic resonance imaging, PET = positron emission tomography and SPECT = single photon emission computed tomography.*

Figure 25 details the range of pathologies that were predominantly found to involve the basal ganglia (39%, 16/41 patients) and the cerebral cortex (34%, 14/41 patients), followed by pituitary findings (27%, 11/41 patients).

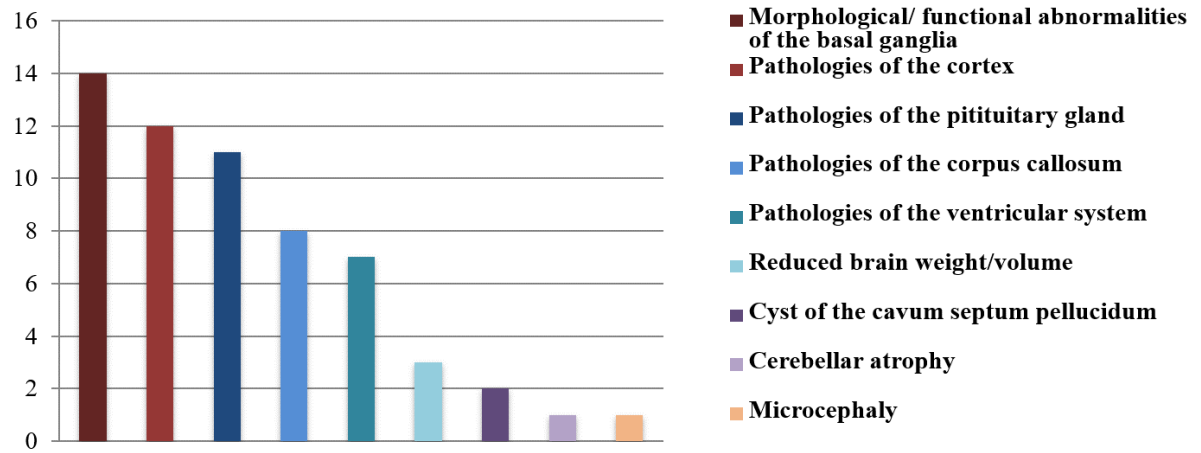

Figure 25. Pathological findings in brain scans*. The bars represent the number of patients with the various pathological findings in the brain scans. Each category of pathological findings is indicated by a different color (see legends on the right). Morphological and functional abnormalities of the basal ganglia include hypoplastic pallida[26,46], reduced volume[183,232,256], hypometabolism[194,195,231], hypoperfusion[238], or reduced dopamine binding[228] in PET scans. Pituitary findings range from hypoplasm[181,184,185,201,258], empty sella[62,195,243], Rathke pouches[26], and cysts[26,46,62] to masses[46] and duplication[207]. Cortical pathologies include cortical dysgenesis[250], atrophy[181,184,189,201,221,233], pachygyria[181,184,250], delayed myelination[181,183,184], and cortical hyperintensities[221]. Pathologies of the corpus callosum[234] include thinning[181], hypoplasm or agenesis[13,60,181,184], and truncation[250]. Findings of the ventricular system comprise asymmetry[181] and dilatation[183,184,189,194,250].*

3.5.5.2.2 Pharmacological treatments for chorea

In a subset of patients with chorea, pharmacological treatments were tried (28%, 72/257 patients with chorea). Information on the effects of these treatments was available in 83% (60/72 patients; no information for 12 patients[61,63,193,196,210,214,217,243]). A complete overview of the treatments and their effects is provided in Table 14. Predominantly dopaminergic drugs were used.

Among them, methylphenidate (100% improvement, 6/6 patients), tetrabenazine (73% improvement, 11/15 patients), and levodopa/carbidopa (67% improvement, 2/3 patients) had the highest success rates. Ropinirole was used in one patient and was found to be helpful. In contrast, levodopa without the use of or information on a combination with a dopa decarboxylase was only effective in a minority of cases (38% improvement, 12/32 patients). Additionally, all anticholinergic drugs used in the patients (see Table 14) were found to be highly effective (100% improvement, 6/6 patients). Chorea was also reduced by benzodiazepines (64% improvement, 9/14 patients). In addition, betablockers and levothyroxine substitution, which were each used in single patients, were found to improve their symptoms. Anticonvulsives (38%

improvement, 3/8 patients), antipsychotics (33% improvement, 3/9 patients), and ethanol consumption (50% improvement, 3/6 patients) were less effective.

Table 14: Pharmacological treatments for chorea

Drugs	Patients		Effect on chorea			Stopped for side effects (s.e.)	
			+	+/-	-		
Dopaminergic drugs	**(n)**	**Dosage**	***n (%)***	***n (%)***	***n (%)***	***s.e., if effect known***	***n (%)***
Levodopa[55,61,63,193,201,209,223,225,233,236,238,242]	32	1.5–20 mg/kg/d	12 (38%)	20 (62%)	1 (3%)	nausea (2)[228]	4 (13%)
Levodopa/carbidopa[182,244]	3	25/100 mg 3–6x/d or 3 mg/kg/d	2 (67%)	1 (33%)	-	digestive side effects (2)[182]	2 (67%)
Methylphenidate[61,68,215,235, 241]	6	54–72 mg/d	6 (100%)	-			
Ropinirole[55]	1	2 mg/d	1 (100%)				
Tetrabenazine[61,63,194,201,236,244,246]	15	25–75 mg/d or 0.35–1.5 mg/kg/d	11 (73%)	4 (27%)	-	gait freezing (1)[236]	6 (40%)
Anticholinergics							
Benztropine[62]	1	no data	1 (100%)	-		pregnancy as a contraindication (1)[62]	-
Biperiden[209]	1	no data	1 (100%)			-	
Trihexyphenidyl[61,63,201,209]	4	1 mg/d	4 (100%)				
Anticonvulsives							
Carbamazepine[61]	2	1,000 mg/d or 20 mg/kg/d	-	2 (100%)	-		
Topiramate[61,193]	3	100 mg/d	3 (100%)	-		depression (2)[193], fatigue (2)[193]	2 (67%)
Valproate[61,63]	2	1,000 mg/d	-	2 (100%)	-		
Levetiracetam[215]	1	2x 500 mg/d		1 (100%)	-	gait instability (1)[215]	1 (100%)
Antipsychotics							
Fluphenazine[236]	1	no data	1 (100%)	-			
Haloperidol[61,223,228]	4	no data	-	3 (75%)	1 (25%)	-	
Olanzapine[55]	1	5 mg/d		1 (100%)	-		
Quetiapine[194]	1	75 mg/d	1 (100%)	-			
Risperidone[236]	1	no data	-	1 (100%)	-	somnolence[236]	1 (100%)
Sulpiride[63]	1	200 mg/d		1 (100%)	-		
Benzodiazepines							
Clobazam[61]	1	0.5 mg/kg/d	-	1 (100%)	-		1 (100%)
Clonazepam[63,193,215,223,236]	13	0.5 mg 1–2x/d	8 (62%)	5 (38%)	-	eye itching (2)[215]	3 (23%)
Diazepam[217]	1	40 mg/d	1 (100%)	-			
Beta-blockers							
Propranolol[55]	1	60 mg/d	1 (100%)	-		asthma (1)[55]	1 (100%)
Other treatments							
Ethanol consumption[194,211, 223,233]	6	no data	3 (50%)	3 (50%)	-		
Levothyroxine[66]	1	no data	1 (100%)	-			

Abbreviations and symbols: + = improvement, +/- = no improvement, - = worsening, n = number of patients, s.e. = side effects

3.5.5.3 Lung phenotype

A pulmonary phenotype was found in 46% of patients (161/351 patients; see Table 15). Among them, IRDS (51%, 82/161 patients) was the most frequent manifestation, followed by recurrent pulmonary infections (40%, 64/161 patients), bronchial hyperreactivity/asthma (16%, 26/161 patients), and ILD (15%, 24/161 patients). Lung cancer was rarely found (2%, 3/161 patients). One fifth of patients (21%, 34/161 patients) with a pulmonary phenotype, mainly those with IRDS, needed mechanical ventilation[14,17,20,60,63,68,189,190,196,201,208,212,213,217,234,236,247,250,251,254],

27% (44/161 patients) required supplemental oxygen[13, 14,20,26,46,191,192,196,213,227,239,247, 254] (median duration of treatment 1.5 years, IQR 0.2–4.8 years); 4% (6/161 patients) needed extracorporeal membrane oxygenation (ECMO)[13,17,213] and 2% (4/161 patients) were treated with nitric oxide inhalation[189,234,250,251], and 1% (2/161 patients) with non-invasive ventilation[14,239]. Some of the patients also received steroids (15%, 24/161 patients)[13,14,17,20,68,189,192,196,213,227,234, 247,251,254], surfactants (5%, 8/161 patients)[17,61,189,196,234,250, 251], hydroxychloroquine (6%, 10/161 patients)[13,14,61,196,213,251], azathioprine (4%, 7/161 patients)[14,213] and, bronchodilators (2%, 3/161 patients)[20, 191,192,227].

Table 15: Pulmonary symptoms and disorders

Disorders of infancy	**n**	**End-stage respiratory failure**	**n**
IRDS[13,14,17,20, 26,45-47,60,68,181,182,184,189,190,196, 201,205,209,212,213, 216,217,226,234,236, 239,247,250-254]	82	**Death of pulmonary cause**[13,14,17,47,181, 184,189,196,210,234,250,251,254]	18
Chronic lung disease of prematurity/bronchopulmonary dysplasia[250,254]	2	**Lung transplantation**[13]	6
		Neoplastic disorders	**n**
Bronchopulmonary malformations: laryngeal cleft[13], lung sequestration[46]	2	**Lung cancer**[38,47,178]	3
		Benign lung tumor[228]	1
Interstitial lung disease (ILD)	**n**	**Other disorders**	**n**
ILD (all)[11,12,38,45,170,21068,206,247]	14	**PAH**[11,207,210236]	10
NEHI[165,166]	5	**Alveolar proteinosis**[13]	1
Pulmonary fibrosis[11,12,194,217254]	5	**(Recurrent) pneumothoraces**[13,254]	2
Bronchiectasis[68]	1	**Lung perforation**[184]	1
Inflammatory disorders	**n**	**Unspecific symptoms/conditions**	**n**
Recurrent pulmonary infections[13,20,26,46,55,60,62-64,67,68, 181,183,184,189,191-193,208,213,219,225,227,229,232,238,239]	64	**Pulmonary dysfunction** (not further specified)[40,166191]	7
Bronchial hyperreactivity/asthma[14,26,46,55,61,64, 193,210,219, 227,236,241,243]	26	**Exertional dyspnea** (not further specified)[12,191]	7

Abbreviations: ILD = interstitial lung disease, IRDS = infant respiratory distress syndrome, n = number of patients, NEHI = neuroendocrine hyperplasia of infancy, PAH = pulmonary arterial hypertension.

Lung CTs were performed in 13% of the patients with a lung phenotype (21/161 patients; Table 16)[13,20,61,191,192,196,213,247,250,254], and biopsies were carried out in 6% (10/161 patients; see Tables 16 and 17)[13,20,47,191,192,196,250]. All patients with lung autopsies (85% of patients with autopsies, 6/7 patients[13,17,47,234,250]; no data in one patient[65]) had a lung phenotype and pathological findings. The main observations were signs of ILD and/or infiltrates (Tables 16–17). In addition, on the histological level impairments of surfactant were revealed by pathological SP-B, -C, -D, pro-SP-B, and/or pro-SP-C[13,234] staining. Moreover, weak or absent staining of *NKX2-1*[13,196,250] in type II pneumocytes was found (see Tables 16–17).

3.5.5.4 Kaplan-Meier analysis for survival

Kaplan-Meier analysis was performed for patients with known age at last visit and information on survival or death (n = 246; Figure 26). The comparison of patients with or without a lung phenotype, irrespective of the type of variant, showed a reduced survival rate for patients with a lung phenotype (n = 131, $p = 0.000$). The mean estimated survival for all patients was

Table 16: Pathological lung computed tomography (CT) and biopsy results

Lung Pheno.	Lung CT result	Lung biopsy result
IRDS[250] (Patient 13)	*At 2.75 months:* multiple discrete 1- to 2 mm cysts predominantly posteriorly, atelectasis	*At 2.75 months*: type II pneumocytes with *NKX2-1* staining, diminished SP-A and ABCA3 staining, numerous lamellar bodies, increased alveolar muscularization
NEHI[191,192] (Patient 63)	*At 7 months, and at 3,8,12,17 and25 years:* diffuse ground-glass opacities	*At 7 months*: peribronchiolar lymphocytic aggregates. Increased neuroendocrine cells in Bombesin staining.
IRDS, ILD[196] (Patient 71)	*At 1 month:* bilateral ground-glass opacities with alveolar condensations	*At 6 months*: macrophage accumulation, hyperplasia and decreased *NKX2-1* staining of type II pneumocytes
IRDS, fibrosis[255] (Patient 212)	*Age unknown*: extensive lung destruction	*Age unknown:* emphysema and interstitial fibrosis
ILD[14]*	*11/13 patients (age unknown)*:* ground-glass opacities, atelectasis infiltrates, cysts, fibrosis	*2/3 patients (age unknown):* septal wall thickening, increased macrophages, hyperplasia of pneumocytes type II; proSPC staining not detected (Patients 81–83)
Lung pheno.	**Lung biopsy result**	
Unspecific[192]	*At 11 years:* patchy chronic bronchiolitis, increased alveolar macrophages, patchy lymphoid aggregates. Increased neuroendocrine cells (bombesin staining). Normal staining for *NKX2-1*, SP-A/-B, pro-SP-C, and ABCA3 (Patient 66)	
IRDS[47]	*At 11 months*: pulmonary alveolar proteinosis (Patient 324)	
NEHI[191,192]	*At 63 years*: subtle mosaicism in the lower and middle lobes and lingula (Patient 64)	
NEHI[191,192]	*At 33 years*: subtle mosaicism throughout all lobes, small pleural nodules (Patient 67)	
NEHI[191,192]	*At 62 years*: subtle mosaicism at the bases, nodule (Patient 68)	
IRDS[213]	*At 4 months*: atelectasis *At 20 months*: diffuse ground-glass, ill-defined, scattered, nodular opacities (Patient 274)	
IRDS,ILD[247]	*At 1 month*: cystic fibrosis, ground-glass opacities (Patient 320)	
IRDS,ILD[254]	*At 1 month*: bronchopulmonary dysplasia, pneumonia (Patient 346)	

*Abbreviations and symbols: *not exactly known which patients of the publication (the report provided a summary of all CTs). ABCA3 = ATP-binding cassette subfamily A member 3, ILD = interstitial lung disease, IRDS = infant respiratory distress syndrome, NEHI = neuroendocrine hyperplasia of infancy, pheno. = phenotype, SP-A = surfactant protein A.*

Table 17: Pathological lung autopsy results

Lung phenotype	Lung autopsy result
Triad: IRDS and lung cancer[47] (Patient 324)	Large-cell lung cancer; widespread metastases including the myocardium, emphysema, interstitial fibrosis, type II pneumocyte hyperplasia
Lung/thyroid: IRDS, ECMO, mechanical ventilation[17] (Patient 140)	Impaired pulmonary architecture, low alveolar counts, hyperplasia of type II pneumocytes
Triad: IRDS, chronic lung disease of prematurity / bronchopulmonary dysplasia, mechanical ventilation[250] (Patient 13)	Emphysema, bronchopneumonia, *NKX2-1* deletion found in the tissue
Triad: IRDS, interstitial pneumopathy, mechanical ventilation[196] (Patient 71)	Hyperplasia of type II pneumocytes; reduced *NKX2-1* labeling in alveolar type 2 epithelial cells; reduced SP-B, SP-C, pro-SP-B, and pro-SP-C labeling
Lung/thyroid: IRDS, ECMO, mechanical ventilation[234] (Patient 141)	Type II pneumocyte hyperplasia, alveolar remodeling, normal *NKX2-1* staining, nearly absent pro-SP-C labeling
Triad: IRDS, recurrent infections, lung transplantation, pulmonary arterial hypertension[13] (Patient 240)	Weak *NKX2-1* and SP-D staining

Abbreviations: ECMO = extracorporeal membrane oxygenation, IRDS = infant respiratory distress syndrome, SP-B/C/D = surfactant protein A/B/C/D, pro-SP-B/C = pro surfactant protein B/C.

64 years (IQR 59.7–68.6 years, 95% confidence interval). The estimated survival for patients with a lung phenotype (n = 131), at 51.1 years (IQR 45.6–56.6 years, 95%, confidence interval) was significantly lower than that of patients without pulmonary manifestations (n = 115), at 71 years (IQR 65.7–76.3 years, 95% confidence interval).

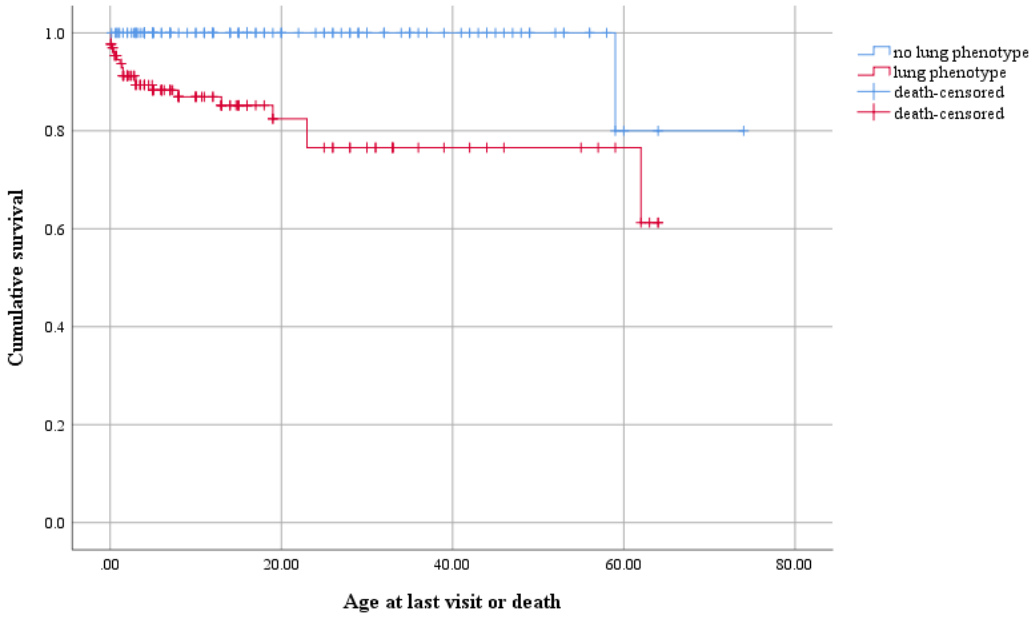

Figure 26. Kaplan-Meier analysis for survival*. Overall survival is depicted for all patients with known age or age at death (n = 246) divided into two groups: those with a lung phenotype (red, n = 131) and those without pulmonary manifestations (blue, n = 115). The analysis shows a significantly lower estimated survival for patients with a lung phenotype. Abbreviations: n = number of patients.*

3.5.5.5 Thyroid phenotype

Thyroid manifestations were reported in 64% of patients (224/351 patients; see Table 18 for symptoms and disorders and Table 8 for cases with lacking data on the thyroid). The majority had hypothyroidism (93%, 208/224 patients) – often congenital (45%, 94/208 patients) – and thyroid dysgenesis (23%, 47/208 patients), which was primarily thyroid hypoplasia (64%, 30/47 patients). Thyroid carcinomas were rarely found in the patients (1%, 3/224 patients).

Table 18: Disorders and symptoms of the thyroid and parathyroid gland

Hypothyroidism	n	Thyroid dysgenesis	n
Hypothyroidism (all)[13,14,17,20,26,45-47,55,60-64,66-68,181,182,184,188-194,196,199-204,206-210,212-214,216,217,219,225-227,229,230,232-236,238-248,250-254]	208	**Agenesis**[14,26,46,60,203,206,209,214,216,226,248]	11
Congenital hypothyroidism[13,14,17,20,26,45,46,55,60,63,67,68,181,182,184,188-190,194,196,200-204,206,208,209,212-214,216,217,219,226,227,229,230,232,234,235,239,240,244,245,247,248,250-254]	94	**Hypoplasia**[20,26,46,60,204,214,216,217,232,239,246,250,253]	30
Diagnosis in infancy (not neonatal)[13,47,61,64,66,68,202,209,210,212,217,225,236,238,241,242]	30	**Ectopy**[14,26,202]	9
		Other disorders	**n**
Diagnosis in adulthood[55,61,62,64,68,182,212,214,219,232,233,236,243]	19	**Goiter**[228,249]	10
Compensated hypothyroidism[17,26,47,55,60,61,63,64,66-68,182,184,188,189,193,194,196,201,206,208,212,216,217,219,225,226,229,230,232,240,242,248]	61	**Hypoparathyroidism**[62,68]	3
Transient hypothyroidism / Euthyroid sick syndrome[13,14,182,188,193,208,250,254]	7	**Papillary thyroid carcinoma**[249]	3

Abbreviations: n = number of patients.

In 43% of those with hypothyroidism (90/208 patients), levothyroxine substitution was reported[13,20,26,45-47,55,61,62,64,66-68,182,184,189,190,193,194,196,201,202,204,208-210,212-214,216,217,219,225-227,229,230,232-234,238-240,242 244,247,251-254]. Among patients with thyroid autopsies (71% of patients with postmortem studies, 5/7 patients; in two cases, no data were available on the thyroid[47,65]), all had a thyroid phenotype, and 60% (3/5 patients) presented pathological findings, which included hemiagenesis, reduced labeling of *NKX2-1*, TG, and colloid depletion (see Table 19). No abnormalities were found in two patients[13,17].

Table 19: Pathological thyroid autopsy results

Thyroid phenotype	Thyroid autopsy result
Triad: congenital hypothyroidism, congenital hypothyroidism/euthyroid sick syndrome, hemiagenesis[250] (Patient 13)	Hemiagenesis
Triad: congenital hypothyroidism[196] (Patient 71)	Reduced *NKX2-1* staining in the epithelium of follicles and reduced thyreoglobulin labeling
Lung/thyroid phenotype: congenital hypothyroidism[234] (Patient 141)	Colloid depletion, *NKX2-1* staining intact

3.5.5.6 Additional clinical features

Additional clinical features were found in 20% of the patients (71/351 patients; 30%, 49/166 variants). Short stature (7%, 26/351 patients; 12%, 20/166 variants), fever of unknown origin (6%, 20/351 patients; 8%, 13/166 variants), retarded skeletal age (5%, 18/351 patients; 7%, 11/166 variants), facial dysmorphisms (4%, 15/351 patients; 7%, 12/166 variants) and hypo-/oligodontia (4%, 15/351 patients; 7%, 11/166 variants) were described most frequently. All additional clinical features and disorders reported in the patients are detailed in Table 20.

Table 20: Additional clinical features

Additional clinical feature	Patients n (%)	Variants n (%)	References
All patients with *NKX2-1* variants	351 (100%)	166 (100%)	see Tables 25–28
All patients with additional clinical features	71 (20%)	49 (30%)	20,26,46,60,61,64,67,68,181,183,184,188-190,194,198,199,205,206,208-210,214,223,226,233,236,239,243,250,254
Short stature	26 (7%)	20 (12%)	20,26,46,64,181,184,210,223,236,239,243
Fever of unknown origin	20 (6%)	13 (8%)	26,46
Retarded skeletal age	18 (5%)	11 (7%)	20,26
Hypo-/oligodontia	15 (4%)	12 (7%)	26,46,60,61,68,183,184,188,190,199,208,239
Facial dysmorphism			26,46,181,183,184,189,198,205,206,214,226,239
Hyperphagia	12 (3%)	9 (5%)	26,46,68
Congenital heart defects	9 (3%)	8 (5%)	26,46,194,254
Hypophagia	7 (2%)	6 (4%)	26,46
Hyperextendable joints	4 (1%)	4 (2%)	209,214,233,239
Hydronephrosis	2 (<1%)	2 (1%)	67,183
Megabladder			67,194
Maldescensus testis			67,183
Hypogammaglobulinemia	1 (<1%)		239
Developmental dysplasia of the hip			239
Polydaktylie			206
Vesicouretral reflux			181

Abbreviations: n = number of patients.

3.5.6 Analysis of genotype-phenotype correlations

3.5.6.1 The disease severity score

After analysis of all organ manifestations in the patients (cf. Section 3.5.5), the following disease severity score (see Table 21) was developed including only those symptoms that could – *on the basis of current knowledge about the syndrome* – be plausibly related to *NKX2-1* deficiency.

Table 21: Clinical disease severity score

Combinations of organ manifestations							**Point**
Organ triad (brain, lung and thyroid)							3
Two organ manifestations (brain-lung, brain-thyroid or lung-thyroid)							2
Isolated organ manifestation (brain-only, lung-only or thyroid-only)							1
Neurological phenotype	**Points**			**Thyroid phenotype**			**Point**
a) Motor impairments				Hypothyroidism (if congenital or diagnosed in early childhood, then 1 extra point)			1 or 2
Chorea (with or without athetosis)	2						
Muscular hypotonia	2			Thyroid agenesis			3
Muscular dystonia	2			Thyroid hypoplasm			2
Ataxia	2			Thyroid carcinoma			3
Spasticity/ paralysis	2			Thyroid ectopy			1
Impairments of speech	2			Goiter			1
Apraxia	2			Hypoparathyroidism			1
Epilepsy/seizures	2			**Lung phenotype**			**Point**
Myocloni	1			Infant respiratory distress syndrome			2
Orofacial dyskinesia	1			Chronic lung disease of prematurity / bronchopulmonary dysplasia			2
Tics	1						
Intentional tremor	1			Interstitial lung disease			2
Recurrent drop attacks	1			Lung transplantation			3
Restless legs syndrome	1			Malignant pulmonary neoplasm			3
Subjective “leg weakness” (no paralysis)	1			Pulmonary cause of death			3
Hypopallaesthesia	1			Benign pulmonary neoplasm			2
b) Impairments of development				Pulmonary alveolar proteinosis			2
General developmental delay	2			Pulmonary arterial hypertension			2
Speech delay	2			Bronchiectasis			2
c) Cognitive impairments				Bronchial hyperreactivity / asthma / “exertional dyspnea”			1
Severe intellectual disability/ severe psychomotor delay	3			Recurrent pulmonary infections			1
Memory deficit	2						
Learning difficulties	1						
d) Psychiatric disorders							
Each disorder	1						
Scores							
Overall Score	**Points**	**Brain score**	**Points**	**Thyroid score**	**Points**	**Lung score**	**Point**
Mild	**1–5**	Mild	**1–2**	Mild	**1–2**	Mild	**1**
Moderate	**6–10**	Moderate	**3**	Moderate	**3**	Moderate	**2**
Severe	**11–15**	Severe	**4–5**	severe	**4–5**	severe	**3–4**
Very severe	**16/>16**	Very severe	**6/>6**	Very severe	**6/>6**	Very severe	**5/>5**

The disease severity score is composed of three organ-specific scores (brain, lung, and thyroid) and combinations of organ manifestations (triad, two-organ, or single-organ manifestation) with one to three points for each manifestation, resulting in an overall score (mild, moderate, severe, and very severe). Three points were reserved for the organ triad and very severe organ impairments such as thyroid agenesis, lung or thyroid carcinoma, end-stage lung failure, severe intellectual disability, and ultimately death due to *NKX2-1* deficiency. Two points were attributed to two organ manifestations and severe impairments, and finally, one point was given to isolated organ manifestations and all other mild to moderate impairments. Each patient was scored according to this disease severity score, and the results of the overall scores are presented for each patient in Tables 25–28. The overall scores for all patients revealed mild (30%, 106/351 patients) or moderate overall severity (31%, 110/351 patients) in the majority of patients, followed by severe (23%, 80/351 patients) and very severe disease (16%, 55/351 patients). The median overall disease severity score for the whole patient population was moderate: 9 points (IQR 5–13). Males were found to have significantly higher overall and brain scores compared to females (see Table 22). In addition, patients with *de novo* variants had significantly higher overall and organ scores compared to patients with familial variants (see Table 22).

Table 22: Disease severity scores for gender and mode of inheritance

Criteria	n	Overall score	Brain score	Lung score	Thyroid score
All patients	351	**9** (IQR 5–13)	**4** (IQR 2–8)	**0** (IQR 0–2)	**1** (IQR 0–2)
Males	147	**10** (IQR 5–15)	**5** (IQR 2–8)	**0** (IQR 0–2)	**1** (IQR 0–2)
Females	185	**8** (IQR 5–12)	**4** (IQR 2–7)	**0** (IQR 0–2)	**1** (IQR 0–2)
Males vs. females (*p*-value)		**0.007**	**0.013**	0.328	0.103
Mode of Inheritance	**n**	**Overall score**	**Brain score**	**Lung score**	**Thyroid score**
De novo variants	86	**11** (IQR 8–15)	**6** (IQR 3–8)	**1** (IQR 0–3)	**2** (IQR 1–2)
Familial variants	230	**7** (IQR 3–11)	**4** (IQR 2–7)	**0** (IQR 0–1)	**1** (IQR 0–1)
De novo vs. familial variants (*p*-value)		**0.000**	**0.009**	**0.000**	**0.000**

Abbreviations and explanations: n = number of patients. Significant results (p < 0.05) are indicated in bold letters and on gray backgrounds. If there is a significant p-value for the comparison of two groups using the two-tailed unpaired Student's t-test, then the p-value is indicated on the handside of the respective group that has a higher score (i.e. it is written either on the left or right side of the column). P-values of comparisons without a significant difference between the groups are shown in the middle of the column.

3.5.6.2 Analysis of genotypic groups based on the impaired gene region

3.5.6.2.1 Statistics and overview of genotypic groups based on the impaired gene region

Table 23 summarizes the statistics on the frequency of different variants, demographics, modes of inheritance, and organ manifestations for the genotypic groups based on the impaired gene region, namely, the WGD, HD, non-HD+, and non-HD groups (see Table 6a for group definitions). For each category, absolute and relative numbers, and if one-way ANOVA was performed, then *p*-values for the differences between all groups before the subanalysis (see Table 23) were provided.

Table 23: Statistics for genotypic groups based on the impaired gene region

Criteria	WGD n (%)	HD n (%)	Non-HD+ n (%)	Non-HD n (%)	All n (%)	ANOVA (*p-value*)
Total number of variants	48 (29%)	41** (25%)	41 (25%)	37 (22%)	167 (100%)	
Total number of patients	62 (18%)	133 (38%)	92 (26%)	64 (18%)	351 (100%)	
Demographics*						
Info on sex available	*60 (97%)*	*130 (98%)*	*87 (95%)*	*55 (86%)*	*332 (95%)*	
Females	*36 (60%)*	*70 (54%)*	*50 (57%)*	*29 (53%)*	*185 (56%)*	
Information on age available	*45 (73%)*	*82 (62%)*	*71 (77%)*	*48 (75%)*	*246 (71%)*	
Age at last visit (median +IQR, years)	***8** (IQR 3.8–16)*	***14.5**(IQR 4.7–30)*	***12** (IQR 3–30)*	***10** (IQR 4.1–22)*	***11.5** (IQR 4–28)*	
Lethal cases (mortality rate)	4 (6%)	10 (8%)	3 (3%)	4 (6.3%)	21 (6%)	
Death due to the lung phenotype	4 (100%)	8 (80%)	2 (67%)	4 (100%)	18 (86%)	
Age at death (median +IQR, years)	**2.1**(IQR 0.6–10.5)	**4** (IQR 1.2–20)	**1.4** (1,1,59)	**0.4** (IQR 0.1–17.4)	**1.5** (IQR 0.5–16)	
Mode of inheritance*						
Info on mode of inh. available	*50 (81%)*	*126** (95%)*	*83 (90%)*	*58 (91%)*	*317 (90%)*	
Familial variants	*24 (48%)*	*111** (88%)*	*57 (69%)*	*38 (65%)*	*230** (73%)*	**0.000**
De novo variants	*24 (48%)*	*16** (13%)*	*26 (31%)*	*20 (35%)*	*86** (27%)*	**0.000**
Parental mosaicism	*2 (4%)*		-		*2 (1%)*	
Members of affected families	21 (34%)	108 (81%)	59 (64%)	33 (52%)	221 (63%)	
Pat. without affected fam. members	41 (66%)	25 (19%)	33 (36%)	31 (48%)	130 (37%)	
Types of variants*						
Deletion of only *NKX2-1*	10 (16%)		-		10 (3%)	
Additional deletion of other genes	52 (84%)	-	3 (3%)	1 (2%)	56 (16%)	
Missense variants	-	99** (74%)	-	16 (25%)	115 (33%)	
Nonsense variants		23 (17%)	30 (33%)	4 (6%)	57 (16%)	
Other deletions		10 (8%)	16 (17%)	30 (47%)	56 (16%)	
Splice site variants		-	35 (38%)	-	35 (10%)	
Duplications			11 (12%)	6 (9%)	17 (5%)	
Stop loss variants		-		3 (5%)	3 (1%)	
Insertions		1 (0.8%)	-	4 (5%)	5 (1%)	
Indels		-	1 (1%)	1 (2%)	2 (1%)	
Frameshift variants		10 (8%)	37 (40%)	35 (55%)	82 (23%)	
Combination of organ manifestations*						
Information on thyroid available	*52 (84%)*	*98 (74%)*	*92 (100%)*	*56 (88%)*	*298 (85%)*	
Triad	30 (48%)	31 (23%)	36 (39%)	18 (28%)	115 (33%)	**0.002**
Brain/thyroid	15 (24%)	29 (22%)	26 (28%)	12 (19%)	82 (23%)	0.535
Brain/lung	5 (8%)	6 (5%)	4 (4%)	4 (6%)	19 (5%)	0.719
Lung/thyroid	1 (1.6%)	7 (5%)	2 (2%)	3 (5%)	13 (4%)	0.488
Isolated brain	11 (18%)	46 (35%)	22 (24%)	14 (22%)	93 (26%)	**0.047**
Isolated thyroid	-	1 (1%)	-	13 (20%)	14 (4%)	**0.000**
Isolated lung		13 (10%)	2 (2%)	-	15 (4%)	**0.001**
Brain phenotype*						
Overall brain phenotype	61 (98%)	112 (84%)	88 (96%)	48 (75%)	309 (88%)	**0.000**
Motor impairments	56 (90%)	112 (84%)	84 (91%)	45 (70%)	297 (85%)	**0.002**
Info. on age at onset available	*27 (44%)*	*25 (22%)*	*32 (36%)*	*23 (48%)*	*107 (34%)*	
Age at onset (median +IQR, years)	***1** (IQR 0.3–2)*	***1.1** (IQR 0.6–2)*	***1.8**(IQR0.8-4.8)*	***1** (IQR 0.5–3)*	***1.1** (IQR 0.5–2.5)*	
Chorea without athetosis	21 (34%)	95 (71%)	48 (52%)	32 (50%)	196 (56%)	**0.000**
Choreoathetosis	15 (24%)	10 (8%)	27 (29%)	7 (11%)	59 (17%)	**0.000**
Muscular hypotonia	27 (44%)	34 (26%)	35 (38%)	24 (38%)	120 (34%)	0.052
Cognitive impairments (overall)	28 (45%)	35 (26%)	25 (27%)	13 (20%)	101 (29%)	**0.012**
Developmental delay (overall)	31 (50%)	33 (25%)	50 (54%)	26 (41%)	140 (39%)	**0.000**
Severe intellectual disability	16 (26%)	9 (7%)	9 (10%)	7 (11%)	41 (12%)	**0.001**
Learning difficulties	5 (8%)	27 (20%)	15 (16%)	7 (11%)	54 (15%)	0.109
Microcephaly	6 (10%)	-	1 (1%)	-	7 (2%)	**0.000**
Thyroid phenotype*						
Overall thyroid phenotype	46 (74%)	68 (51%)	64 (69%)	46 (72%)	224 (64%)	**0.001**
Overall hypothyroidism	46 (74%)	64 (48%)	59 (64%)	39 (61%)	208 (59%)	**0.003**
Congenital hypothyroidism	21 (34%)	21 (16%)	28 (30%)	24 (38%)	94 (27%)	**0.003**
Thyroid dysgenesis	8 (13%)	13 (10%)	15 (16%)	11 (17%)	47 (13%)	0.397
Lung phenotype*						
Overall lung phenotype	36 (58%)	57 (43%)	44 (48%)	24 (38%)	161 (46%)	0.105
IRDS	17 (27%)	22 (17%)	20 (22%)	23 (36%)	82 (23%)	**0.020**
Recurrent pulmonary infections	19 (31%)	21 (16%)	16 (17%)	8 (13%)	64 (18%)	**0.038**
Bronchial hyperreactivity / asthma	1 (2%)	16 (12%)	8 (9%)	1 (2%)	26 (7%)	0.869
Interstitial lung disease	5 (8%)	10 (8%)	6 (7%)	3 (5%)	24 (7%)	**0.014**
Additional clinical features*						
Overall additional clinical features	30 (48%)	16 (12%)	16 (17%)	9 (14%)	71 (20%)	**0.000**
Dysmorphic features	10 (16%)	-	2 (2%)	3 (5%)	15 (4%)	**0.000**
Hypo-/oligodontia	14 (23%)	1 (1%)		-	15 (4%)	**0.000**

*(Table 23 continued): Abbreviations, explanations, and symbols: *data are provided for the number of patients; **category with a patient that has two different variants and is therefore counted in two different categories. Calculations that were based on the total number of patients with available information are given in cursive. ANOVA = analysis of variance, HD = homeodomain, info = information, inh. = inheritance, +IQR = median numbers are provided with interquartile ranges, IRDS = infant respiratory distress syndrome, n = number of patients, nHD = non-HD, nHD+ = nHD affecting the HD, mem. = members, pat. = patients, WGD = whole gene deletions. Significant results of ANOVA for the comparison of all groups (p <0.05) are indicated by bold letters and gray backgrounds.*

Tables 25 to 28 list the genotypes, phenotypes, and overall disease severity scores for all 351 patients coded by color (green for brain, blue for lung, and red for thyroid manifestations), allowing for the identification of visual patterns. The patients were grouped according to their respective genotypic groups based on the gene region, namely, WGD (Table 25, Parts 1 and 2), HD variants (Table 26, Parts 1 and 2), non-HD+ variants (Table 27, Parts 1–3), and non-HD variants (Table 28, Parts 1 and 2) for better comparability.

As can be seen in Tables 23 and 25, WGD often resulted in an organ triad (48%, 30/62 patients) and additional clinical features (48%, 30/62 patients). Patients with HD, non-HD+, and non-HD variants were found in similar abundance (see Table 23), but there were many differences in the types of variants and organ phenotypes (see Tables 26–28 for details). While HD variants were predominantly missense variants (74%, 99/133 patients, see Table 23), often resulting in an isolated brain phenotype (35%, 46/133 patients; see Tables 23 and 26), non-HD+ variants were mostly splice site variants (38%, 35/92 patients, see Table 23) and found to often lead to an organ triad (39%, 36/92 patients; see Tables 23 and 27). In contrast, half of the non-HD variants were deletions (47%, 30/64 patients, see Table 23) and caused numerous combinations of organ manifestations (see Tables 23 and 28). Overall, 82% of all patients with *NKX2-1* variants (287/351; sum of all WGD and non-WGD with HD involvement i.e. the HD and non-HD+ groups) were found to have an impairment of the HD (see Tables 23 and 25–27).

A subanalysis of non-WGD (see Figure 18 and Tables 26–28 for all variants) revealed HD involvement in 78% (225/289 patients) and the following differences in the involvement of functional domains. 1) The HD group predominantly involves HD-only variants (74%, 99/133 patients; 66%, 27/41 variants; see Table 23) that are all missense variants (see Figure 27a and Table 30 for the composition of the missense group), and the remaining groups involve three domains (24%, 34/133 patients; 34%, 14/41 variants; see Table 23). 2) The non-HD+ group is composed of three domains (48%, 44/92 patients; 39%, 16/41 variants; see Table 23) and four domain variants (52%, 48/92 patients; 61%, 25/41 variants; see Table 23) and does not involve missense variants (see Figure 27a and Table 30 for the composition of the missense group). 3) The non-HD group, as shown in Table 23, comprises the two-domain variants that involve the

NK2-SD and the C-TAD (28%, 18/64 patients; 32%, 12/37 variants); the single-domain variants C-TAD-only (47%, 30/64 patients; 38%, 14/37 variants), N-TAD-only (8%, 5/64 patients; 8%, 3/37 variants), and NK2-SD-only (5%, 3/64 patients; 8%, 3/37 variants); and all variants without domain involvement (13%, 8/64 patients; 14%, 5/37 variants). In contrast to the non-HD+ group, missense variants are found in the non-HD group (25%, 16/64 patients; 26%, 9/34 variants; see also Figure 27 for the composition of the missense group). Overall, phosphorylation site involvement is high in the non-HD group (75%, 48/64 patients; 85%, 29/34 variants; see Table 23).

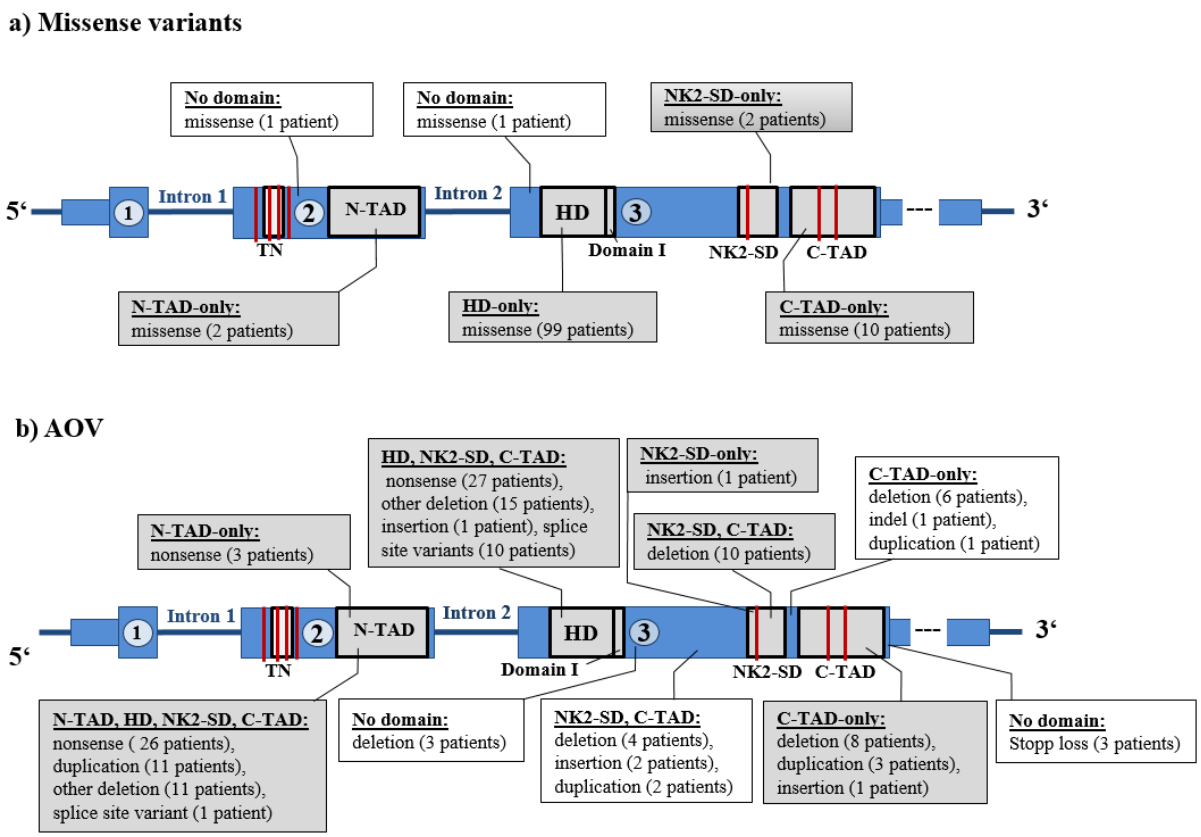

Figure 27. Missense variants and all other variants (AOV). *The schematic drawings depict the NKX2-1 gene and its different functional domains (gray boxes) and seven phosphorylation sites (red bars) for isoform 1 (NM_001079668.2). The exons and introns are shown in blue (exons indicated by circles with the numbers 1–3). Missense variants are shown in (a) and AOV in (b). Variants located in a domain are shown in gray boxes, and those outside of domains are in white boxes. The domains that are affected by the variants are shown in the boxes, and the respective variants are listed below them. While missense variants involve either one domain or no domain (a), AOV are mostly multi-domain variants (b). Abbreviations: C-TAD = C-terminal transactivation domain, Domain I = inhibitory domain, HD = homeodomain, NK2-SD = NK2-specific domain, N-TAD = N-terminal transactivation domain, and TN = Tinman domain.*

3.5.6.2.2. Disease severity for genotypic groups based on the impaired gene region

Table 24 presents the results of the subanalysis of severity scores for the genotypic groups based on the impaired gene region. Overall, HD variants were found to have the lowest scores in all

categories (see Table 26 for the overall scores for each patient with a HD variant). The differences were statistically significant compared to WGD and non-HD+ for the overall, brain, and thyroid scores and for the thyroid score compared to non-HD. No significant differences could be detected between the other groups for any of the scores. Moreover, for the lung scores, there were no significant differences between any of the groups.

Table 24: Disease severity scores for genotypic groups based on the impaired gene region

Region of the *NKX2-1* gene	**n**	**Overall score**	**Brain score**	**Lung score**	**Thyroid score**
WGD	62	**11** (IQR 6–15)	**5** (IQR 2–8)	**1** (IQR 0–3)	**1** (IQR 0–2)
HD	133	**8** (IQR 3–11)	**3** (IQR 2–7)	**0** (IQR 0–2)	**1** (IQR 0–1)
Non-HD+	92	**9** (IQR 7–14)	**6** (IQR 3–8)	**0** (IQR 0–2)	**1** (IQR 0–2)
Non-HD	64	**8** (IQR 5–14)	**4** (IQR 0–8)	**0** (IQR 0–3)	**1** (IQR 0–2)
p-values					
WGD vs. HD		**0.001**	**0.034**	0.142	**0.001**
WGD vs. non-HD+		0.338	0.960	0.050	0.687
WGD vs. non-HD		0.919	0.467	0.204	0.854
HD vs. non-HD+		**0.005**	**0.019**	0.603	**0.003**
HD vs. non-HD		0.127	0.794	0.357	**0.000**
Non-HD+ vs. non-HD		0.450	0.098	0.208	0.276

Abbreviations: HD = homeodomain, n = number of patients, non-HD+= non-HD variants affecting the HD, non-HD = non-HD variants that do not affect the HD, vs. = versus, WGD = whole gene deletions. Significant results ($p < 0.05$) are indicated in bold letters and on gray backgrounds. If there is a significant p-value for the comparison of two groups using the two-tailed unpaired Student's t-test, then the p-value is indicated on the handside of the respective group that has a higher score (i.e. it is written either on the left or right side of the column). P-values of comparisons without a significant difference between the groups are listed in the middle of the column.

Table 25: Genotypes and phenotypes in WGD (Part 1)

Patient number	Deletion of cytobands: 14q… + (including partial deletion)	Hg19 base pair coordinates, if known: NC_000014.8:g…del	Deletion size (Mb)	Overall score	Gender	Age at last visit or death (years)	Triad	Brain/ lung phenotype	Brain/ thyroid pheno.	Thyroid/ lung pheno.	Brain-only phenotype	Brain phenotype (any)	Chorea	Athetosis	Muscular hypotonia	Muscular dystonia	Ataxia	Myocloni	Spasticity	Speech impairment	Developmental delay	Speech delay	Intellectual disability	Learning difficulties	Microcephaly	Lung phenotype (any)	IRDS	Recurrent resp. infect.	Interstitial lung disease	PAH	Asthma/ bronch. hyper.	Pulm. cause of death	Thyroid phenotype (any)	Hypothyroidism	Thyroid agenesis	Thyroid hypoplasia	Thyroid ectopy	Hypoparathyroidism	Mode of inheritance	Deletion of other genes
1[181]	**11.2q13**	no data	8.8-9.2	11	f	6																										?						D	+
2[181]			?	6	m	1																										?						D	+
3[181]			6.9-8.7	13	f	2.4																											C						D	+
4[181]	**11.2q21.1**		7.4-11.5	3	f	0.2																										?						D	+
5[238]	**12q13**		2.6	11	f	5																																F	?
6[183]	**12q13.1**		3.5	6	m	2.5		X																														D	+
7[189]	**12q13.3**		17.9	19	f	3	X																								X		C					F	+
8[189,184]				20	f	13																									X		C					F	+
9[198]	**12q21.2**		11.1	6	?	?																										?						?	?
10[14,198]			?	9	f	7																																?	?
11[60]	**12q21.2**	**30631931_46074637** #1	15.4	18	f	6																											C					?	+
12[26]	**12q21.3**	**32696760_47874795** #2	15.17	16	m	?																											C					?	+
13[250]	**12q21.3**	no data	?	19	m	0.4	X																								X		C					F	+
14[181]	**12q22**		?	10	f	1.2		X																		*					X	?						D	+
15/6[53]			19, 20.6	3, 3	f	56,53																										?						F	+
17/8[53]			16.4, 19	3, 3	m,f	29,26																										?						F	+
19[53]			20.6	12	f	4																										?						F	+
20[14]	**13**		?	15	m	19																																?	?
21/2[207]			1.2	5, 5	f	18,?a																																F	+
23[68]	**13.1**		0.9	18	m	44																											C					F	+
24[68]				18	m	8																																F	+
25[68]				8	m	?,a																																F	+
26[13]	**13.1q13.2**		?	15	m	2.8																											C					?	?
27[13]	**13.1q21.1**		?	13	f	15																																D	?
28[232]			3.68	13	m	8																											C					?	+
29[216]			1.1	16	m	0.1																																D	+
30[24,41]	**13.1q21.1**	**33376720_4259706** #3	9.2	25	m	16																											C					D	+
31[184]		**34600510_41578688** #4	6.98	9	f	30																																D	+
32[26]	**13.1q13.3**	**35001682_37176663** #5	2.17	5	f	?																																?	+
33[184,19]	**13.1q21.2**	**35007710_44901392** #6	9.9	9	f	15																											C					F	+
34[239]	**13.2q21.1**	**35404289_38723530**	3.32	19	f	10																											C					D	+
35[67]		**35443407_3952738** #7	4.08	5	f	1.3																											C					D	+
36[63]	**13.2q21.1**	**35581654_40301792**	4.7	12	m	6																											C					D	+
37[13]	**13.2q21.1**	-	?	17	f	7																											C					D	?
38[61]		**35862118_42103227** #8	6.2	6	f	20																																D	+
39[26]	**13.2q21.1**	**36068376_38433164** #9	2.36	16	m	?																																F	+
40[184]		**36129971_39816401** #10	3.69	8	f	7																											C					D	+
41[26]	**13.2q13.3**	**36151329_37483407** #11	1.33	3	f	?																																?	+
42[61]	**13.2q21.3**	**36165220_50031170** #12	13.8	6	f	12																																D	+
43[199]	**13.2-q21.1**	**36257988_39532584** #13	3.27	15	?	?																																D	+

*Abbreviations and symbols: *not further specified, ? = no information available, + = Additional deletion of other genes, a = adult, bronch. hyper. = bronchial hyperreactivity, C = congenital hypothyroidism, D = de novo variant, f = female, F = familial variant, infect. = infections, IRDS = infant respiratory distress syndrome, m = male, Mb = Megabytes, PAH = pulmonary arterial hypertension, pheno. = phenotype, pulm. = pulmonary, resp. = respiratory, x = death. Original descriptions of the variant in the respective publications (gray background): #1 = 29701682_45144380 and 14q13, #2 = 31766511_46944545, #3 = 32446471_41666818 and 14q11.2q13.3, #4 = 33670261_40648438 and 14q13.3, # 5= 34071433_36246141, #6 = 3407077461_43971142 and 14q13-21, #7 = 36058432_36825446, #8 = 34931869_41172977 and 14q13.2q21.2, #9 = 35138127_37502915, #10 = 35199722_38886152 and 14q13.3, #11 = 35221080_36553158, #12 = 35234971_49100920, #12 = 35878308_36227579 and 14q13.3, #13 = 35327739_38602335*

Table 25: Genotypes and phenotypes in WGD (Part 2)

Patient number	Deletion of cytobands: 14q… + (including partial deletion)	Hg19 base pair coordinates, if known: NC_000014.8:g…del	Deletion size (Mb)	Overall score	Gender	Age at last visit or death (years)	Triad	Brain/ lung phenotype	Brain/ thyroid phenotype	Thyroid/ lung phenotype	Brain-only phenotype	Brain phenotype (any)	Chorea	Athetosis	Muscular hypotonia	Muscular dystonia	Ataxia	Myocloni	Spasticity	Speech impairment	Developmental delay	Speech delay	Intellectual disability	Learning difficulties	Microcephaly	Lung phenotype (any)	IRDS	Recurrent resp. infect.	Interstitial lung disorder	PAH	Asthma/ bronch. hyper.	Pulm. cause of death	Thyroid phenotype (any)	Hypothyroidism	Thyroid agenesis	Thyroid hypoplasia	Thyroid ectopy	Hypoparathyroidism	Mode of inheritance	Deletion of other genes
44/5	**13.2q13.3**[69,188,208]	**36463113_37638965** #14	1.17	5	f	16, ?a			■			■	■																										**F**	+
46[61]	**13.3**	**36808557_37157828** #15	0.3	9	f	17			■			■	■		■	■	■			■				■															**D**	+
47[26]	**13.3q21.1**	**36814106_39156509** #16	2.34	20	m	?	■					■									■					■	■	■			■								**F**	+
48[26]			2.34	8	m	8		■				■											■			■			■										**?**	**?**
49[15]	**13.3q21.1**	no data	?	15	m	3.5	■					■			■						■	■			■	■	■			■				**C**					**D**	**?**
50[208]	**13.3**	**36852260_37735951** #17	0.88	13	f	14.9	■					■	■	■							■	■				■				■				**C**					°	+
51[208]				12	f	7.3	■					■	■	■							■	■				■				■									°	+
52[26]	**13.3**	**36919999_37684890** #18	0.76	6	f	?	■					■	■	■												*													**D**	+
53[63]		**36924171_37283221**	0.36	9	f	11					■	■	■		■	■					■																		**?**	+
54[184,208]	**13.3**	**36940484_37656846** #19	0.717	10	m	5	■					■					■				■			■		■		■										**D**	+
55[15]		no data	?	11	f	10	■					■					■				■					■	■	■											**?**	**?**
56[206]				11	m	?, a			■			■				■	■																■	■					**F**	+
57[206]				9	m	?, a	■					■	■													■			■				■	■					**F**	+
58[206]				10	f	7			■			■	■		■		■																■	■					**D**	+
59[26]		**36972991_37072789** #20	0.099	15	m	?	■					■	■	■	■	■				■	■					■		■					■						**F**	+
60[26]				15	m	?	■					■			■	■					■		■			■		■					■	■			■		**F**	+
61[26]				18	m	?	■					■	■	■	■	■				■	■		■			■		■					■	■					**F**	+
62[26]		**36988681_37755695** #21	0.76	5	m	?			■			■	■	■																								**?**	+

*Abbreviations and symbols: *not further specified, °parental genetic mosaicism, ? = no information available, + = deletion of other genes, a = adult, bronch. hyper. = bronchial hyperreactivity, C = congenital hypothyroidism, D = de novo variant, f = female, F = familial variant, infect. = infections, IRDS = infant respiratory distress syndrome, m = male, Mb = megabytes, PAH = pulmonary arterial hypertension, pulm. = pulmonary, resp. = respiratory, x = death. Original descriptions of the variants in the publications (gray background): #14 = 35532864_36708716 and 14q13.1, #15 = 35878308_36227579, #16 = 35883857_3822620, #17 = 35922011_36805702, #18 = 35989750_36754641, #19 = 36010235_36726597, #20 = 36042742_36142540, #21 =36058432_36825446*

Table 26: Genotypes and phenotypes in variants of the HD (Part 1)

Patient number	Variant (hg19)		Overall score	Gender	Age at last visit or death (years)	Triad	Brain/ lung pheno.	Brain/ thyroid pheno.	Thyroid/ lung pheno.	Brain-only pheno.	Lung-only phenotype	Thyroid-only pheno.	Brain phenotype	Chorea	Athetosis	Muscular hypotonia	Muscular dystonia	Ataxia	Myocloni	Speech impairment	Developmental delay	Speech delay	Intellectual disability	Learning difficulties	Lung phenotype (any)	IRDS	Recurrent resp. inf.	Inter. lung disorder	Asthma/ bronch. hyper.	Lung transplantation	PAH	Pulm. cause of death	Thyroid pheno. (any)	Hypothyroidism	Thyroid agenesis	Thyroid hypoplasia	Hypoparathyroidism	Mode of inheritance	NK2-SD impaired	C-TAD impaired	Phosph. site(s) imp.
63[191,192]	**c.572G>T**; p.(Arg191Leu) #22	mis	4	f	25																															F			
64[191,192]			6	f	63																															F			
65-6[192]			1	f	?, a																			*												F			
67-9[191,192]			3	2m,1f	33,62,0.3																															F			
70[14]			4	f	10																															D			
71[196]	**c.583C>T**; p.(Arg195Trp) #23		14	f	1.5	X																								X		C				D			
72[13]			18	m	3	X																								X		C				D			
73[14]			11	f	1.5	X																								X						D			
74[64]	**c.583delC**; p.(Arg195Glyfs*33) #24	del	12	m	13																															F	!	!	!
75[64]			7	f	?																															F	!	!	!
76-7[64]			7, 5	m, f	22, 17																															F	!	!	!
78[64]			8	f	14																															F	!	!	!
79[13]	**c.590T>C**;p.(Leu197Pro)	mis	11	m	17																															?			
80[13]	**c.592T>C**; p.(Phe198Leu)		9	f	19						X																				X					F			
81[13]			8	f	23						X																				X					F			
82[13]			4	f	33																											?					F			
83[13]			7	f	5						X																				X					F			
84[13]	**c.594C>G**;p.(Phe198Leu)		8	f	16																															F			
85[26,46]	**c.596C>A**;p.(Ser199*) #25	non	21	m	13																											C				?	!	!	!
86[200]	**c.596C>G**;p.(Ser199Trp) #26	mis	3	?	?																											C				D			
87[240]	**c.605A>T**;p.(Gln202Leu) #27		8	f	5																											C				F			
88[233]	**c.606G>T**;p.(Gln202His) #28		13	f	32																											C				?			
89[63]	**c.612delC**;p.(Tyr204*) #29	del	13	m	1.5																											C				F	!	!	!
90[26]	**c.612C>G**; p.(Tyr204*) #30	non	20	m	?																															?	!	!	!
91[26]			17	?	?																															?	!	!	!
92[209]	**c.613G>T**; p.(Glu205*) #31		10	f	4																															F	!	!	!
93[209]			16	m	8	X																										C				F	!	!	!
94-5[209]			5	f	?																															F	!	!	!
96[52,60,61]	**c.616C>G**;p.(Leu206Val) #32	mis	11	m	16																															D			
97[217]	**c.617T>A**; p.(Leu206Gln)		11	f	55																			*												F			
98[217]			17	m	31																			*												F			
99[217]			9	f	33																															F			
100[217]			8	f	29																															F			
101[217]			10	f	4																															F			
102[217]			10	f	2.5																															F			
103[217]			18	f	13																											C				F			
104[217]			7	f	0.5																															F			
105[217]			9	m	9																															F			
106[217]			11	m	3.5																															F			
107[218]	**c.622C>T**;p.(Arg208*) #33	non	11	f	15																															F	!	!	!
108[219]	**c.623G>C**; p.(Arg208Pro) #34	mis	11	f	7																											C				F			
109[219]			9	m	28																															F			
110[259]	**c.626G>C**; p.(Arg209Pro)		15	f	13																															D			
111[242]			9	m	24																															F			
112[242]			7	m	20																															F			
113[242]			8	f	0.6																															F			
114[62]	**c.631A>T**; p.(Lys211*) #35	non	14	f	49																															F	!	!	!
115[62]			9	f	26																															F	!	!	!
116[26]	**c.634C>T**;p.(Gln212*) #36		6	f	?																			*												D	!	!	!

*Abbreviations and symbols: *not further specified, ? = no information available, ! and gray background = the variant affects the respective domain, ! and black background = the variant encompasses the respective domain, a = adult, bronch. hyper. = bronchial hyperreactivity, C = congenital hypothyroidism, C-TAD = C-terminal transactivation domain, D = de novo variant, del = deletion, f = female, F = familial variant, imp. = impaired, inf. = infections, inter. = interstitial, IRDS = infant respiratory distress syndrome, m = male, mis = missense, NK2-SD = NK2-specific domain, non = nonsense, PAH = pulmonary arterial hypertension, pheno. = phenotype, phos. = phosphorylation, pulm. = pulmonary, resp. = respiratory, x = death. Original descriptions of the variants in the respective publications (gray background): #22 = c.572G>T;p.Arg243His, #23 = c.493C>T;p.R165W, #24 = c.493delC;p.Arg165Glyfs*, #25 = C2519/c.596C>A;p.Ser169X, #26 = c.596C>A;p.Ser199W, #27 = c.515A>T;pQ172L, #28 = c.516G>T;p.Gln172His, #29 = c.612delC;p.Y204X, #30 = c.522C>T;p.Tyr174X, #31 = c.523G>T;p.E175X, #32 = c.526C>G;p.Leu176Val, #33 = R178X, #34 = 533G>C;p.R178P, #35 = c.727A>T;p.K211X, #36 = c.544C>T;p.Gln182X*

Table 26: Genotypes and phenotypes in variants of the HD (Part 2)

Patient number	Variant (hg19)		Overall score	Gender	Age at last visit or death (years)	Triad	Brain/ lung phenotype	Brain/ thyroid pheno.	Thyroid/ lung pheno.	Brain-only phenotype	Lung-only phenotype	Brain phenotype (any)	Chorea	Athetosis	Muscular hypotonia	Muscular dystonia	Ataxia	Myocloni	Spasticity	Speech impairment	Developmental delay	Speech delay	Intellectual disability	Learning difficulties	Lung phenotype (any)	IRDS	Recurrent resp. infect.	Interstitial lung disorder	Asthma/ bronch. hyper.	Lung cancer	PAH	Pulm. cause of death	Thyroid phenotype (any)	Hypothyroidism	Thyroid agenesis	Thyroid hypoplasia	Thyroid ectopy	Hypoparathyroidism	Mode of inheritance	NK2-SD impaired	C-TAD impaired	Phosph. site(s) impaired
117[210]	**c.650C>A**; p.(Ser217*) #37	non	10	m	35			■				■	■	■		■				■																		F	!	!	!
118[210]			7	m	?	■						■	■											■					■				■	■					F	!	!	!
119[210]			6	m	?			■				■	■																				■	■					F	!	!	!
120[210]			10	f	62		X					■	■											■						■									F	!	!	!
121[61,193]	**c.671T>G**; p.(Leu224Arg)	mis	18	f	59	■						■	■		■	■	■		■		■			■					■				■	■					F			
122[61,193]			19	f	39	■						■	■		■	■	■	■	■		■			■					■				■	■					F			
123[61,193]			18	f	36	■						■	■			■	■	■	■		■			■					■				■	■					F			
124[61,193]			22	m	64	■						■	■		■	■	■		■	■	■			■		■			■				■	■					F			
125[61,193]			16	m	44					■		■	■		■	■	■	■	■		■			■															F			
126[61,193]			14	m	14		■					■	■		■		■		■		■			■					■										F			
127[61,193]			12	m	11		■					■	■				■		■		■			■					■										F			
128[61,193]			14	m	8	■						■	■				■		■		■			■					■										F			
129[46]	**c.674_675insGG**; p.(Ser226Alafs*3) #38	ins	16	m	3	■						■	■	■	■					■	■				■		■						■	C		■			D	!	!	!
130[201]	**c.689T>C**;p.(Leu230Pro)	mis	13	?	16			■				■	■		■						■	■											■	C					D			
131[60]	**c.695C>T**;p.(Pro232Leu) #39		13	f	1.5			■				■	■		■								■										■	C		■			D			
132[26]**	**c.698C>G**; p.(Thr233Arg) #40		4	f	?				■															■				■					■			■		F			
133[26]			4	f	?				■															■				■					■			■		F			
134[26]			4	f	?				■															■		■							■			■		F			
135[26]			16	m	?	■						■	■		■				■	■				■	■	■							■	■		■	■	F			
136[26]	**c.703G>T**; p.(Val235Phe) #40-43		23	f	?	■						■	■		■		■		■	■		■		■	■	■							■	C		■			D			
137-8[26,238]			8,7	f.m	6,7			■				■	■	■			■							■									■	C					F			
139[24,41]			23	m	15	■						■	■		■		■		■	■		■		■	■	■							■	C		■			D			
132[24]**			9	f	?				■															■				■					■				■	F			
140[17]	**c.709A>T**;p.(Ile237Phe) #44		9	f	0.1				X															■	■						X	■	C					D			
141[234]	**c.711C>G**;p.(Ile237Met) #45		6	f	0.1				X															■	■						X	■	C					?			
142[188]	**c.713G>T**; p.(Trp238Leu) #46		3	f	?			■																								■	C					F			
143-154[188]			3	10f,2m	?					■		■	■																			?						F			
155[233]	**c.713G>C**;p.(Trp238Ser) #47		15	m	12			■				■	■	■	■	■	■	■														■						?			
156[14]	**c.714G>A**;p.(Trp238*)	non	11	f	4.9	■						■	■		■									■	■							■	C					F	!	!	!
157[14]			13	m	0.8	■						■			■							■		■	■				■			■	C					F	!	!	!
158[201]	**c.714delG**; p.(Trp238Cysfs*9)	del	10	m	35			■				■	■		■						■											■						F	!	!	!
159[201]			11	m	8			■				■	■		■						■	■										■						F	!	!	!
160-1[201]			10,9	m, f	4, 3			■				■	■		■						■											■						F	!	!	!
162[52,53]	**c.719A>C**;p.(Gln240Pro) #48	mis	13	f	16			■				■	■		■								■									■	C		■			D			
163-84[188]	**c.727C>T**;p.(Arg243Ser) #49		3	15m,7f	?					■		■	■																			?						F			
185[220]	**c.727C>T**; p.(Arg243Cys)		5	f	3					■		■	■								■																	F			
186[220]			10	f	30					■		■	■		■		■	■			■																	F			
187[61]	**c.728G>C**; p.(Arg243Pro) #50		9	m	12					■		■	■		■	■							■															F			
188[61]			8	m	37					■		■	■		■	■		■			■																	F			
189[14]	**c.728G>A**;p.(Arg243His) #51		4	m	0.3						■														■	■													D			
190[61]	**c.732C>A**; p.(Tyr244*) #52	non	8	m	12			■				■	■		■									■								■	■					F	!	!	!
191-2[61]			5	m	28,29					■		■	■		■																							F	!	!	!
193[61]			6	f	53			■				■	■				■																■	■					F	!	!	!
194[221]	**c.745C>T**; p.(Q249*) #53		8	m	30					■		■	■		■	■			■					■														F	!	!	!
195[221]			3	f	59					■		■												■														F	!	!	!

*Abbreviations and symbols: *not further specified, **patient with two different variants, ? = no information available, ! and gray background=the variant affects the respective domain, ! and black background= the variant encompasses the respective domain, a = adult, bronch. hyper. = bronchial hyperreactivity, C = congenital hypothyroidism, C-TAD = C-terminal transactivation domain, D = de novo variant, del = deletion, f = female, F = familial variant, infect. = infections, IRDS = infant respiratory distress syndrome, ins = insertions, m = male, mis = missense, NK2-SD = NK2-specific domain, non = nonsense, PAH = pulmonary arterial hypertension, pheno. = phenotype, phos. = phosphorylation, pulm. = pulmonary, resp. = respiratory, x = death. Original descriptions of the variants in the respective publications (gray background): #37 = c.650C>A;p.S217X, #38 = 2595insGG, #39 = P202L, #40 = c.608C>G;p.Thr203Arg+ c.613G>T;p.Val205Phe, #41 = c.613G>T;Val205Phe, #42 = c.613G>T;V205P, #43 = G2662T, #44 = Ile207Phe, #45 = c.621C>G;p.(Ile207Met), #46 = c.G713G>T;p.W238L, #47 = c.623G>C;p.Trp208Ser, #48 = Q210P, #49= C727A;p.R243S, #50 = c.728G>A;p.R243P, #51 = c.728G>A;p.Arg243His, #52 = c.642C>A, #53 = c.745C>T; p.Q249X*

Table 27: Genotypes and phenotypes in non-HD variants affecting the HD (non-HD+; Part 1)

Patient number	Variant (hg19)		Overall score	Gender	Age at last visit or death (years)	Triad	Brain/ lung phenotype	Brain/ thyroid pheno.	Thyroid/ lung pheno.	Brain-only phenotype	Thyroid-only pheno.	Lung-only phenotype	Brain pheno. (any)	Chorea	Athetosis	Muscular hypotonia	Muscular dystonia	Ataxia	Myocloni	Spasticity	Speech impairment	Developmental delay	Speech delay	Intellectual disability	Learning difficulties	Lung pheno. (any)	IRDS	Recurrent resp. infect.	Inter. lung disorder	Asthma/ bronch. hyper.	Lung cancer	PAH	Pulm. cause of death	Thyroid pheno. (any)	Hypothyroidism	Thyroid agenesis	Thyroid hypoplasia	Thyroid ectopy	Mode of inheritance	Exon/ intron	NK2-SD impaired	N-TAD impaired	C-TAD impaired	Phosph. site(s) imp.
196[222]	**c.174_180del**; p.(Met59Alafs*40) #54	del	6	m	49					■			■					■			■																	F	Exon 2	!	!	!	!
197[222]			9	f	15					■			■	■				■			■	■																F		!	!	!	!
198[222]			7	f	10					■			■	■	■			■				■																F		!	!	!	!
199[14]	**c.175_176del**; p.(Met59Glyfs*379) #55		4	f	3.5							■													■			■										F		!	!	!	!
200[14]			3	m	33							■													■			■										F		!	!	!	!
201[26]	**c.246dup**; p.(Met83Hisfs*356) #56	dup	21	f	?	■							■	■	■	■	■	■			■	■				■	■	■							■			D		!	!	!	!
202[26]	**c.253C>T**;p.(Gln85*) #57	non	5	f	?			■					■	■	■																							D		!	!	!	!
203[14]	**c.243_256del**; p.(Ala82Argfs*352) #58	del	9	?	2			■					■	■	■	■		■					■															D		!	!	!	!
204[14]	**c.267del**; p.(His90Thrfs*11) #59		10	?	4			■					■							■				■											C			D		!	!	!	!
205[14]	**c.267dupG**; p.(His90Alafd*349) #60	dup	9	f	42	■							■	■												■			■						C			D		!	!	!	!
206[251]	**c.278_306del**; p.(Ala93Glyfs*336) #61	del	9	m	0.8				X																	■							X		C			D		!	!	!	!
207[26]	**c.294C>G**; p.(Tyr98*) #62-64	non	6	m	?	■							■	■	■											*												F		!	!	!	!
208[55]			14	f	19	■							■	■		■								■		■			■						C			D		!	!	!	!
209[55]			8	m	5			■					■	■								■												C			D		!	!	!	!
210[55]			5	f	47			■					■	■																							D		!	!	!	!
211[235]			12	m	15					■			■	■					■		■	■		■														?		!	!	!	!
212[20]	**c.313dupG**; p.(Val105Glyfs*334) #65-66	dup	16	m	14.7	■							■	■	■							■				■	■	■	■						C		■	D		!	!	!	!
213-15[223]	**c.319C>T**; p.(Gln107*) #67	non	12	m	42,41,39					■			■	■	■		■	■				■			■													F		!	!	!	!
216-7[223]			3	m	64,36					■			■	■																								F		!	!	!	!
218[26]	**c.326C>A**; p.(Ser109*) #68		14	f	?			■					■	■	■	■		■				■														■		F		!	!	!	!
219[26]			18	f	?	■							■	■	■	■	■	■			■	■				■	■											F		!	!	!	!
220[26]			15	f	?	■							■	■		■	■					■				■		■								■		F		!	!	!	!
221[14]	**c.344dup**; p.(Tyr116Leufs*323) #69-70	dup	11	f	2.2		■						■	■		■								■		■	■											D		!	!	!	!
222[14]			15	m	3	■							■	■		■								■		■	■								C			D		!	!	!	!
223[45]			13	m	6	■							■	■	■	■						■				■	■								C			D		!	!	!	!
224[236]	**c.344delG**; p.(Gly115Alafs*10)	del	15	f	2		■						■	■		■						■	■			■	■		■		■							?		!	!	!	!
225[61]	**c.347dup**; p.(Tyr116) #71	dup	14	f	6		■						■	■		■		■			■	■			■	■			■										F		!	!	!	!
226[61]			12	f	30	■							■	■		■	■								■	■			■										F		!	!	!	!
227[61]			10	f	57	■							■	■											■	■				■									F		!	!	!	!
228[24,41]	**c.351C>A**; p.(Cys117*) #72-73	non	22	m	1.5	■							■	■	■	■					■	■		■		■	■	■							C	■		F		!	!	!	!
229[26]			13	m	?	■							■	■		■					■	■				■								*				F		!	!	!	!
230[13]	**c.384_391del**; p.(Pro129Glys*307) #74	del	11	f	5	■							■	■		■						■				■		■										D		!	!	!	!
231[236]	**c.397dupA**; p.(Thr133Asnfs*306)	dup	9	m	7			■					■	■		■						■																?		!	!	!	!
232[252]	**c.424G>T**;p.(Gly142*) #75	non	9	m	0.1	■							■			■										■	■								C			D		!	!	!	!

*Abbreviations and symbols: *not further specified, ? = no information available, ! and greay background = the variant affects the respective domain, ! and black background = the variant encompasses the respective domain, a = adult, bronch. hyper. =bronchial hyperreactivity, C = congenital hypothyroidism, C-TAD = C-terminal transactivation domain, D = de novo variant, del = deletion, dup = duplications, f = female, F = familial variant, imp. = impaired, infect. = infections, IRDS = infant respiratory distress syndrome, ins = insertions, inter. = interstitial, m = male, NK2-SD = NK2-specific domain, non = nonsense, N-TAD = N-terminal transactivation domain, PAH = pulmonary arterial hypertension, pheno. = phenotype, phosph. = phosphorylation, pulm. = pulmonary, resp. = respiratory, x = death. Original descriptions of the variants in the respective publications (gray background): #54 = c.84_90delCCGAATC, #55 = c.175_176del, #56 = c.157insC;p.Met53Hisfx355, #57 = c.160C>T;p.Gln55X, #58 = c.243_256delGGCCATGCAGCAGC, #59 = c.267delG, #60 = c.267dupG, #61 = c.278_308del;p.Ala93GlyfsX54, #62 = c.204C>G;p.Tyr68X, #63 = c.294C>G;p.Y98*, #64 = c.204C>G;p.Y68*, #65 = c.223dupG;p.Val75Glyfs*334, #66 = c.224insG, #67 = c.319C>T;p.Q107*, #68 = c.236C>A;p.Ser79X, #69 = c.344dup, #70 = c.255insG;p.G86fs*322, #70 = c.257dupA, #72 = C1302A, #73 = c.261C>A;p.Cys87X, #74 = c.384_391del;p.Pro129GlysX307, #75 = c.G334>T; p.Gly112X*

Table 27: Genotypes and phenotypes in non-HD+ (Part 2)

Patient number	Variant (hg19)		Overall score	Gender	Age at last visit or death (years)	Triad	Brain/ lung phenotype	Brain/ thyroid phenotype	Thyroid/ lung phenotype	Brain-only phenotype	Brain phenotype (any)	Chorea	Athetosis	Muscular hypotonia	Muscular dystonia	Ataxia	Myocloni	Spasticity	Apraxia	Speech impairment	Developmental delay	Speech delay	Intellectual disability	Learning difficulties	Lung phenotype (any)	IRDS	Recurrent resp. inf.	Interstitial lung disorder	Asthma/ bronch. hyper.	PAH	Benign lung tumor	Pulm. cause of death	Thyroid phenotype (any)	Hypothyroidism	Thyroid agenesis	Thyroid hypoplasia	Goiter	Mode of inheritance	Intron/ exon	NK2-SD impaired	N-TAD impaired	C-TAD impaired	Phosph. site(s) impaired
233[63]	**c.428G>A**; p.(Trp143*) #76-77	non	18	f	3	■																										C				F	Exon 2	!	!	!	!
234[26]			14	m	?	■																										C				?		!	!	!	!
235[243]			17	m	46	■																														F		!	!	!	!
236[243]			20	f	20					■																										F		!	!	!	!
237[26]	**c.429del**; p.(Trp143Cysfs*23) #78	del	15	m	?			■																												D		!	!	!	!
238[206]	**c.432C<A**; p.(Tyr144*) #79-82	non	8	m	2			■																								C				?		!	!	!	!
239[63]			7	f	6					■																										D		!	!	!	!
240[224]			14	f	1.4	X																									X					?		!	!	!	!
241[13]			8	f	4					■																										D		!	!	!	!
242[237]	**c.436_437dup**; p.(Asn147Profs*20) #83	dup	6	m	12					■																										?		!	!	!	!
243[225]	**c.463+1G>A**;p.?	spl	16	f	6	■																														D		!		!	!
244[226]	**c.463+2T>C**;p.? #84		17	m	2	■																										C				D	Intron 2				
245[226]			9	f	26	■																														D		!		!	!
246[14]	**c.463+2T>C**;p.?		9	f	28	■																										C				F		!		!	!
247[14]			16	m	2	■																														F		!		!	!
248[61]	**c.463+3_463+6del**; p.? #85		9	f	5			■																												F		!		!	!
249[61]			11	f	4	■																														F		!		!	!
250[61]			7	f	28			■																												F		!		!	!
251[26]	**c.463+4A>G**;p.? #86		6	m	?	■																			*											D		!		!	!
252[26]	**c.463+5G>C**;p.? #87		5	m	?			■																												?		!		!	!
253[66]**	**c.463+41C>T**;p.? #88		9	?	1.1			■																												F		!	!	!	!
254[227]	**c.464-1G>A**;p.? #89		15	m	1.6	■																										C				?		!		!	!
255[211]	**c.464-2A>C**;p.? #90		4	f	14					■																										F		!		!	!
256[211]			4	f	?,a					■																										F		!		!	!
257[60]	**c.464-2A>G**;p.? #91		16	?	16	■																										C				F		!		!	!
258[60]			9	?	16				■																											F		!		!	!
259[212]			9	f	?,a			■																												F		!		!	!
260[212]			7	f	52			■																												F		!		!	!
261[212]			11	f	4	■																										C				F		!		!	!
262[212]			15	m	8	■																										C				F		!		!	!
263[65]	**c.464-2A>T**;p.?		7	f	59					X																										F		!		!	!
264-7[65]			3	f	?					■																										F		!		!	!
268[228]	**c.464-9C>A**; p.(Ile155Thrfs*286)		10	m	64		■																													F		!		!	!
269[228]			7	f	58			■																												F		!		!	!
270-1[228]			7, 5	m, f	25,32					■																										F		!		!	!
272-3[228]			7	f, m	27,25			■																												F		!		!	!
274[213]			10	f	2	■																										C				F		!		!	!
275[213]			8	f	0.6	■																										C				F		!		!	!
276[213]			7	f	?,a			■																								C				F		!		!	!
277[213]			7	f	?,a			■																												F		!		!	!

*Abbreviations and symbols: *not further specified, **patient with two different variants, ? = no information available, ! and gray background = variant affects the respective domain, ! and black background = the variant encompasses the respective domain, a = adult, bronch. hyper. = bronchial hyperreactivity, C = congenital hypothyroidism, C-TAD = C-terminal transactivation domain, D = de novo variant, del = deletion, dup = duplications, f = female, F = familial variant, infect. = infections, IRDS = infant respiratory distress syndrome, m = male, NK2-SD = NK2-specific domain, non = nonsense, N-TAD = N-terminal transactivation domain, PAH = pulmonary arterial hypertension, phos. = phosphorylation, pulm. = pulmonary, resp. = respiratory, spl = splice site variant, x = death. Original descriptions of the variants in the respective publications (gray background): #76 = c.428G>A;p.W143X, #77 = c.338G>A;p.Trp113X, #78 = c.339delG;p.Trp113CysfsX16, #79 = c.342C>G;p.Y114fs, #80 = Y114, #81 = c.432C>A;p.Tyr144X, #82 = c.432C>A;p.Y144X, #83 = c.437_438dupGC;p.N147Pfs*20, #84 = c.373+2T>C, #85 = c.373+1_373+4del, #86 = c.373+4A>G/r.373ins4;p.IleSerfsX126, #87 = c.373+5G>C/r.338_373del35;p.Trp143CysfsX138, #88 = c.463+41C>T and Y130X, #89 = c.374-1G>A, #90 = c.376-2A>C, #91 = c.376-2A>G*

Table 27: Genotypes and phenotypes in non-HD+ (Part 3)

Patient number	Variant (hg19)		Overall score	Gender	Age at last visit or death (years)	Triad	Brain/ lung phenotype	Brain/ thyroid phenotype	Thyroid/ lung phenotype	Brain-only phenotype	Brain phenotype (any)	Chorea	Athetosis	Muscular hypotonia	Muscular dystonia	Ataxia	Myocloni	Spasticity	Apraxia	Speech impairment	Developmental delay	Speech delay	Intellectual disability	Learning difficulties	Lung phenotype (any)	IRDS	Recurrent resp. infections	Interstitial lung disorder	Asthma/ bronch. hyper.	PAH	Benign lung tumor	Pulm. cause of death	Thyroid phenotype (any)	Hypothyroidism	Thyroid agenesis	Thyroid hypoplasia	Goiter	Mode of inheritance	Intron/ exon	NK2-SD impaired	N-TAD impaired	C-TAD impaired	Phosph. site(s) impaired
278[214]	**c.469delC**; p.(Arg157Alafs*9)	del	20	f	3.3			■																									C				F	Exon 3	!		!	!
279[214]			12	f	?, a			■																													F		!		!	!
280[214]			10	f	?, a			■																													F		!		!	!
281[61]	**c.489del**; p.(Ser163Argfs*3) #92		10	f	18	■																														D		!		!	!
282[244]	**c.524C>A**; p.(Ser175*) #93	non	14	f	2.8			■																									C				D		!		!	!
283[194,195]			7	m	5	■																		*								C				F		!		!	!
284[194,195]			5	f	26			■																													F		!		!	!
285[194,195]			8	m	56			■																													F		!		!	!
286[13]	**c.552_556del**; p.(Pro185Alafs*252) #94	del	14	m	1.5	■																											C				D		!		!	!
287[229]	**c.560_569delinsCGC**; p.(Pro187Argfs*29) #95	ind	13	m	14	■																											C				D		!		!	!

*Abbreviations and symbols: *not further specified, ? = no information available, ! and gray background = the variant affects the respective domain, ! and black background = the variant encompasses the respective domain, a = adult, bronch. hyper. = bronchial hyperreactivity, C = congenital hypothyroidism, C-TAD = C-terminal transactivation domain, D = de novo variant, del = deletion, f = female, F = familial variant, ind = indel, infect. = infections, IRDS = infant respiratory distress syndrome, m = male, NK2-SD = NK2-specific domain, non = nonsense, PAH = pulmonary arterial hypertension, phosph. = phosphorylation, pulm. = pulmonary, resp. = respiratory, x = death. Original descriptions of the variants in the respective publications (gray background): #92 = c.399delC, #93 = C609A;p.S145X, #94 = c.552_556del;p.Pro185AlafsX250, #95 = c.470_479delinsGCG;p.P157fsX196*

Table 28: Genotypes and phenotypes in non-HD variants without impairment of the HD (non-HD; Part 1)

Patient number	Variant (hg19)		Overall score	Gender	Age at last visit or death (years)	Triad	Brain/ lung phenotype	Brain/ thyroid pheno.	Thyroid/ lung phenotype	Brain-only phenotype	Thyroid-only phenotype	Brain phenotype (any)	Chorea	Athetosis	Muscular hypotonia	Muscular dystonia	Ataxia	Myocloni	Apraxia	Speech impairment	Developmental delay	Speech delay	Intellectual disability	Learning difficulties	Lung phenotype (any)	IRDS	Recurrent resp. infect.	Interstitial lung disorder	Asthma/ bronch. hyper.	Benign lung tumor	PAH	Lung transplantation	Pulm. cause of death	Thyroid phenotype (any)	Hypothyroidism	Thyroid agenesis	Thyroid hypoplasia	Goiter	Mode of inheritance	Exon/ intron	N-TAD impaired	NK2-SD impaired	C-TAD impaired	Phosph. site(s) impaired
288[203]	**c.221G>A**;p.(Gly74Asp) #96	mis	6	?	?						■																							C				?	Exon 2					
289[204]	**c.268C>T**;p.(His90Tyr)#97		5	?	?						■																							C			■	D		!				
290[216]	**c.313G>A**; p.(Val105Met) #98		5	m	5						■																							C			■	D		!				
291[201]	**c.390C>G**; p.(Tyr130*)	non	15	?	7	■						■	■		■						■	■			■	■								C				D		!				
292[236]			17	f	31	■						■	■			■	■	■			■	■			■	■								C				F		!				
293[236]			16	f	2	■						■	■		■		■				■				■	■					■							F		!				
253[66]**				?	1.1			■				■	■	■	■						■																F		!				
294[63]	**c.520G>T**; p.(Gly174Cys) #99	mis	11	m	14			■				■	■		■	■	■																					D						
295[236]	**c.754_755insT**; p.(Asp252Valfs*187)	ins	18	m	8	■						■	■		■						■	■		■	■	■	■		■									D			!	!	!	
296[245]	**c.794_795insGCGGCGGG**; p.(Gly266Argfs*32) #100		3	f	3						■																							C				D			!	!	!	
297[205]	**c.798_800del**; p.(Gly271del) #101	del	5	?	?					■		■						■			■			■															F						
298[205]			8	?	?		■					■						■			■			■	■	■													F						
299[205]			5	?	?					■		■						■			■			■															F						
300[26]	**c.802_803del**; p.(Gly268Argfs*170) #102		17	m	?		■					■	■		■				■		■				■	■	■									■			F		!	!	!	
301[26]			19	f	?		■					■	■	■	■	■	■				■				■	■	■									■			F		!	!	!	
302[216]	**c.805_813del**;p.(Gly269_ Gly271del) #103		6	f	0.1				■																■	■								C				F		!	!	!	
303[13]	**c.805_813dup**;p.(Gly269_ Gly271dup) #104	dup	11	m	11.7		X					■									■	■			■	■							X						?		!	!	!	
304[13]			9	m	2.8		■					■			■						■				■	■	■												D		!	!	!	
334**	**c.818_838del**; p.(Gly273Gln279) #119	del	9	m	0.1				X																■	■								C				?		!	!	!	
305-6[206]	**c.840_844del**; p.(Gln281Alafs*156) #105		6	?	?		■					■	■																					C				F		!	!	!	
307[206]			13	m	3			■				■	■																					C				F		!	!	!	
308[246]	**c.859_860insTGCC**; p.(Argleufs*156)	ins	14	f	7			■				■	■	■	■		■				■													C		■		D		!	!	!	
309[63]	**c.872C>T**;p.(Pro291Leu)	mis	15	m	12		■					■	■	■	■	■		■			■				■	■	■												D		!	!	!	
310[201]	**c.872C>G**;p.(Pro291Arg)		10	?	5			■				■	■		■			■		■	■																		D		■		!	
311[196,53]	**c.876_877del**;p.(Leu293 Glyfs*145) #106-7	del	23	f	5			■				■	■		■	■					■	■	■		■	■			■					C				D		!	!	!	
312[14]			13	f	10	■						■	■		■								■																D		!	!	!	
313[26]	**c.885del**; p.(Asp296Thrfs*85) #108		16	m	?	■						■	■		■					■	■			■				■						C		■		D		!	!	!	
314[26]	**c.886_887del**; p.(Asp296fs*142) #109		6	f	?	■						■	■	■																								D		■	!	!	
315-6[188]	**c.908del**; p.(Gly303Valfs*78) #110		3	f,m	?					■		■	■	■																				?				F		!	!	!	
317[188]			3	m	?					■		■	■																					?				F		!	!	!	
318[230]	**c.915del**; p.(Ala306Argfs*75) #111-12		11	f	11		■					■	■	■							■		■											C				F			!	!	
319[230]			5	f	4		■					■	■																										F			!	!	
320[247]			22	f	5	■						■	■		■	■	■				■	■		■	■	■	■								C				D			!	!	
321[205]			12	f	19		■					■	■			■	■				■			■											C				F			!	!	

*Abbreviations and symbols: *not further specified, **patient with two different variants, ? = no information available, ! and gray background = variant affects the respective domain, ! and black background = the variant encompasses the respective domain, a = adult, bronch. hyper. = bronchial hyperreactivity, C = congenital hypothyroidism, C-TAD = C-terminal transactivation domain, D = de novo variant, del = deletion, dup = duplications, f = female, F = familial variant, infect. = infections, IRDS = infant respiratory distress syndrome, ins = insertion, m = male, mis = missense, non = nonsense, NK2-SD = NK2-specific domain, N-TAD = N-terminal transactivation domain, PAH = pulmonary arterial hypertension, pheno. = phenotype, phosph. = phosphorylation, pulm. = pulmonary, resp. = respiratory, spl = splice site variant, x = death. Original descriptions of the variants in the respective publications (gray background): #96 = c.G131A;p.G44D, #97 = c.268C>T;p.H60W, #98 = c.313G>A;V105M, #99 = c.520G>T;p.G174C, #100 = c.793-794insGGGCGGCGG;p.264-265insGGG +c.760-887del127;p.A254Tfs*85, #101 = G266del, #102 = c.712delGG;p.Gly238ArgfsX170, #103 = c.793-801del;p.G264-266del, #104 = c.804-812dupCGGCGGGGG;p.Gly269-271dupGlyGlyGly, #105 = c.750delTCAGC, #106 = c.786-787del;p.L263fs, #107 = c.876-877del, #108 = c.793delA;p.Lys265LysfsX87, #109 = c.796delGA; p.Asp266Argfs142X, #110 = ΔG908;p.G30fsX77, #111 = c.825delC, #112 = c.915_916ins;p.Ala303ArgfsX132.*

Table 28: Genotypes and phenotypes in non-HD (Part 2)

Patient number	Variant (hg19)		Overall score	Gender	Age at last visit or death (years)	Triad	Brain/ thyroid phenotype	Lung/ thyroid phenotype	Brain-only phenotype	Thyroid-only phenotype	Brain phenotype (any)	Chorea	Athetosis	Muscular hypotonia	Muscular dystonia	Ataxia	Myocloni	Seizures	Speech impairment	Developmental delay	Speech delay	Intellectual disability	Learning diff./ ADHD	Lung phenotype (any)	IRDS	Interstital lung disorder	Recurrent resp. infections	PAH	Lung cancer	Lung transplantation	Pulm. cause of death	Thyroid phenotype (any)k	Hypothyroidism	Thyroid agenesis	Thyroid hypoplasia	Goiter	Papillary thyroid carcinoma	Mode of inheritance	Exon/ intron	NK2-SD impaired	C-TAD impaired	Phosph. site(s) impaired
322[205]	**c.915del**; p.(Ala306Argfs*75) #111-12	del	13	f	13																										C						F	Exon 3		!	!
323[205]			5	f	43																																F			!	!
324[47]	**c.949dup**; p.(Gln371Profs*122) #113	dup	16	m	23	X																								X							D			!	!
325[26]	**c.956_973delinsCTACA**; p.(Gln319Profs*58) #114	del	23	m	?																																D			!	!
326[231]	**c.978_1056del**; p.(Ala330Serfs*25) #115		5	m	5																										?						F			!	!
327[231]			5	m	11																										?						F			!	!
328-9[231]			5,3	f,m	15,45																										?						F			!	!
330[232]	**c.980_986del**; p.(Ala327Glyfs*52) #116		5	m	32																																F			!	!
331[232]			7	m	3.6																																F			!	!
332[253]	**c.988dup**;p.(Ala330Glyfs*109) #117	dup	13	m	12																										C						D			!	!
333[63]	**c.1022C>T**;p.(Ala341Val) #118	mis	6	f	5																																D			!	!
334[13]**	**c.1036_1049del**; p.(His346Alafs*88) #119	del	9	m	0.1																										C						?		!	!	!
335[13]	**c.1045dup**;p.(His349Profs*90) #120	dup	18	m	6.3																										C						?			!	!
336[248]	**c.1054G>A**; p.(Gly352Ser) #121	mis	6	f	0.8																										C						F			!	!
337[248]			3	m	0.8																										C						F			!	!
338[13]	**c.1092_1108del**; p.(Ser366Glyfs*67) #122	del	11	m	0.5																										C						?			!	!
339[249]	**c.1106C>T**; p.(Ala369Val) #123	mis	5	f	60																																F			!	!
340[249]			2	f	49																																F			!	!
341-2[249]			5	f	44,74																																F			!	!
343-4[249]			2	m,f	44,35																																F			!	!
345[249]			2	f	34																																F			!	!
346[254]	**c.1124_1125insAGGTGGATAC**; p.(Ser376Glyfs*66)	ins	18	m	0.2	X																								X		C					F				!
347[13]	**c.1157_1163dup**;p.(Thyr398Leufs*52)	dup	19	f	4.5																											C					?			!	!
348[63]	**c.1161C>G**;p.(Tyr387*) #124	non	13	f	16																																D			!	!
349[63]	**c.1204T>A**;p.*402Argext*63) #125	stl	10	f	18																																F				
350[215]	**c.1206A>C**; p.(*402Cysext*63) #126-7		3	m	10																										?						F				
351[215]			11	m	?																																F				

*Abbreviations and symbols: *not further specified, **patient with two different variants, ? = no information available, ! and gray background = the variants affects the respective domain, ! and black background = the variant encompasses the respective domain, a = adult, ADHD = attention deficit hyperactivity disorder, bronch. hyper. = bronchial hyperreactivity, C = congenital hypothyroidism, C-TAD = C-terminal transactivation domain, D = de novo variant, del = deletion, diff. = difficulties, dup = duplications, f = female, F = familial variant, IRDS = infant respiratory distress syndrome, ins = insertion, m = male, mis = missense, NK2-SD = NK2-specific domain, non = nonsense, PAH = pulmonary arterial hypertension, phosph. = phosphorylation, pulm. = pulmonary, resp. = respiratory, stl = stopp loss variant, x = death. Original descriptions of the variants in the respective publications (gray background): #111 = c.825delC, #112 = c.915_916ins;p.Ala303ArgfsX132, #113 = 859_860insC, #114 = c.866_883delinsCTACA;p.Gln289ProfsX5, #115 = c.978_1056del;S330fsX24, #116 = c.890_896del; p.Ala327Gly*52, #117 = c.986_987insG;p.A329GfsX108, #118 = c.1022C>T;p.A341V, #119 = c.1034_1047del;p.His346Alafs + c.818_838del21;p.273_279delGCPQQQQ, #120 = c.1044_1045insC;p.His349ProfsX90, #121 = c.G1054A;p.G352S, #122 = c.1092_1108del17; p.Ser366GlyfsX67, #123 = c.1016C>T;p.A339V, #124 = c.1157_1163dupACTACGG;p.Thr389LeufsX52, #125 = c.1161C>G;p.Y387X, #126 = c.1204T>A;p.X402RextX63, #127 = c.1206A>C*

3.5.6.2.3 ANOVA for genotype-phenotype correlation in genotypic groups based on the impaired gene region

Table 29 presents detailed results of the comparison of the four genotypic groups based on the impaired gene region for possible associations with certain phenotypes using one-way ANOVA. Results of the subanalyses between the groups are only given for categories that were previously identified to have significant differences in the overall *p*-values among the four groups (see Table 23 for the respective *p*-values). The main observations of the post hoc analysis with Tukey HSD of the ANOVA analysis are as follows:

1) In most of the categories no relevant differences between WGD and non-HD+ were found.
2) *Inheritance:* HD variants show an association with familial variants when compared to the other groups.
3) *Organ combinations:* The isolated lung phenotype and the isolated thyroid phenotype are significantly more frequent in HD variants and in non-HD variants respectively, compared to all other groups.

Table 29: P-values of post hoc analysis with Tukey HSD of analysis of variance (ANOVA) for genotypic groups based on the impaired gene region

Category	***p-values***					
Inheritance	**WGD vs. HD**	**WGD vs. nHD+**	**WGD vs. nHD**	**HD vs. nHD+**	**HD vs. nHD**	**nHD+ vs. nHD**
Familial variant	**0.000**	**0.010**	0.050	**0.003**	**0.003**	0.985
De novo variant	**0.000**	0.430	0.751	**0.024**	**0.015**	0.972
Organ combination						
Triad	**0.003**	0.615	0.068	0.058	0.903	0.461
Brain-only	0.062	0.827	0.952	0.278	0.228	0.992
Thyroid-only	0.993	1.000	**0.000**	0.990	**0.000**	**0.000**
Lung-only	**0.008**	0.910	1.000	**0.026**	**0.007**	0.907
Brain phenotype						
Overall brain phenotype	**0.019**	0.952	**0.000**	**0.039**	0.221	**0.000**
Motor impairments	0.678	0.998	**0.009**	0.455	0.051	**0.002**
Chorea, no athetosis	**0.000**	0.095	0.236	**0.017**	**0.019**	0.992
Choreoathetosis	**0.016**	0.824	0.174	**0.000**	0.926	**0.011**
Cognitive impairments	**0.033**	0.071	**0.011**	0.999	0.815	0.783
Developmental delay	**0.004**	0.945	0.686	**0.000**	0.130	0.289
Severe intellectual disability	**0.001**	**0.012**	**0.043**	0.895	0.821	0.996
Microcephaly	**0.000**	**0.001**	**0.000**	0.935	1.000	0.961
Thyroid phenotype						
Overall thyroid phenotype	**0.009**	0.933	0.993	**0.022**	**0.022**	0.991
Overall hypothyroidism	**0.003**	0.587	0.418	0.072	0.305	0.978
Congenital hypothyroidism	**0.037**	0.964	0.966	0.066	**0.006**	0.752
Lung phenotype						
IRDS	0.333	0.843	0.665	0.798	**0.014**	0.162
Recurrent pulmonary inf.	0.059	0.154	**0.041**	0.990	0.180	0.862
Additional features						
Overall additional features	**0.000**	**0.000**	**0.000**	0.728	0.985	0.950
Facial dysmorphism	**0.000**	**0.000**	**0.006**	0.844	0.392	0.858
Hypo-/oligodontia	**0.000**	**0.000**	**0.000**	0.991	0.993	1.000

Abbreviations: HD = homeodomain, inf. = infections, IRDS = infant respiratory distress syndrome, nHD = non-HD variants that do not affect the HD, nHD+ = non-HD variants affecting the HD, vs. = versus, WGD = whole gene deletion. Significant results ($p < 0.05$) are indicated in bold letters and on gray backgrounds. If there is a significant p-value for the comparison of two groups, then the p-value is indicated on the handside of the respective group that has shown an association (i.e. it is written either on the left or right side of the column). P-values of comparisons without a significant difference between the groups are listed in the middle of the column..

4) *Brain phenotype:* WGD were found to be associated with the majority of neurological manifestations compared to HD and non-HD variants and with severe intellectual disability and microcephaly in comparison to non-HD+ variants. Chorea without athetosis, as the only manifestation, showed an association with HD variants when compared to all other groups.
5) *Lung phenotype:* No large differences were found between the groups.
6) *Thyroid phenotype:* The overall thyroid phenotype occurred significantly more frequently in any of the other groups (i.e. WGD, non-HD+, and non-HD) compared to HD variants. Also congential hypothyroidism was signifantly less frequent in HD variants compared to WGD and non-HD variants.
7) *Additional clinical features:* The frequency of overall additional clinical features, facial dysmorphism, and hypo-/oligodontia was significantly higher in WGD compared to the other groups.

3.5.6.3 Analysis of genotypic groups based on the type of variant

3.5.6.3.1 Statistics and overview of genotypic groups based on the type of variant

Table 30 summarizes the statistics regarding the frequency of different variants, demographics, modes of inheritance, and organ manifestations for the genotypic groups based on the type of variant (WGD, missense, and AOV; see Table 6b for definitions of the groups). In this context, it is important to note that there were certain major differences in the composition of the groups regarding the involvement of functional domains (see Table 30 and Figure 27).

As illustrated in Figure 27a, missense variants either impair one domain (namely, the one they are located in) or are located outside of domains and thus do not impair any of them. Furthermore, as listed in Table 30, all HD-only variants are missense variants (100%, 99/99 variants), although missense variants are less frequent in the C-TAD-only (33%, 10/30 variants) and N-TAD-only groups (40%, 2/5 patients). Furthermore, missense variants often result in single-organ manifestations, mainly the isolated brain or isolated thyroid phenotype. In contrast, AOV (see Figure 27b) for the most part involve multiple functional domains and frequently lead to multiple organ manifestations.

Table 30: Statistics for genotypic groups based on the type of variant

	WGD	Missense	AOV	All	ANOVA
Criteria	**n (%)**	**n (%)**	**n (%)**	**n (%)**	***(p-value)***
Total number of variants	48 (29%)	36** (22%)	83 (50%)	167 (100%)	
Total number of patients	62 (18%)	115 (33%)	174 (50%)	351 (100%)	
Demographics*				**All**	
Information on sex available	*60 (97%)*	*110 (96%)*	*162 (93%)*	*332 (95%)*	
Females	*36 (60%)*	*64 (58%)*	*85 (53%)*	*185 (56%)*	
Information on age available	*45 (73%)*	*70 (61%)*	*131 (75%)*	*246 (71%)*	
Age at last visit (median +IQR, years)	***8** (IQR 3.8–16)*	***14.5** (IQR 5–33)*	***11** (IQR 3.6–28)*	***11.5** (IQR 4–28)*	
Lethal cases (mortality rate)	4 (6%)	8 (7%)	9 (5%)	21 (6%)	
Death due to lung phenotype	4 (100%)	7 (88%)	7 (78%)	18 (86%)	
Age at death (median +IQR, years)	**2.1**(IQR 0.6-10.5)	**2.3** (IQR 0.5–15.5)	**1.4** (IQR 0.4–41)	**1.5** (IQR 0.5–16)	
Mode of inheritance*				**All**	
Information on the mode of inheritance	*50 (81%)*	*110** (96%)*	*157 (90%)*	*317 (90%)*	
Familial variants	*24 (48%)*	*90** (82%)*	*116 (74%)*	*230** (73%)*	**0.000**
De novo variants	*24 (48%)*	*21** (19%)*	*41 (26%)*	*86** (27%)*	**0.009**
Parental mosaicism	*2 (4%)*	-		*2 (1%)*	
Member of an affected family	21 (34%)	86 (75%)	114 (66%)	221 (63%)	
Patients without affected family members	41 (66%)	29 (25%)	60 (34%)	130 (37%)	
Impairment of functional domains*				**non-WGD variants**	
All patients with non-WGD variants		115 (100%)	174 (100%)	289 (100%)	
Overall impairment of the HD (non-WGD)		99 (86%)	126 (73%)	225 (78%)	
4 domains (HD, C-TAD, NK2-SD, N-TAD)		-	48 (28%)	48 (17%)	
3 domains (HD, C-TAD, NK2-SD)			78 (44%)	78 (27%)	
2 domains (C-TAD, NK2-SD)			18 (10%)	18 (6%)	
1 domain (N-TAD-/HD-/NK2-SD-/C-TAD-only)		113 (98%)	24 (14%)	137 (47%)	
HD-only		99 (86%)	-	99 (34%)	
C-TAD-only		10 (9%)	20 (11%)	30 (10%)	
N-TAD-only		2 (2%)	3 (2%)	5 (2%)	
NK2-SD-only		2 (2%)	1 (1%)	3 (1%)	
No domain		2 (2%)	6 (3%)	8 (3%)	
Combination of organ manifestations*				**All**	
Information on thyroid available	*52 (84%)*	*80 (70%)*	*166 (95%)*	*298 (85%)*	
Triad	30 (48%)	17 (15%)	68 (39%)	115 (33%)	**0.000**
Brain/thyroid	15 (24%)	17 (15%)	50 (29%)	82 (23%)	**0.023**
Brain/lung	5 (8%)	6 (5%)	8 (5%)	19 (5%)	0.583
Lung/thyroid	1 (2%)	7 (6%)	5 (3%)	13 (4%)	0.233
Isolated brain	11 (18%)	42 (37%)	40 (23%)	93 (26%)	**0.009**
Isolated thyroid	-	13 (11%)	1 (1%)	14 (4%)	**0.000**
Isolated lung		13 (11%)	2 (1%)	15 (4%)	**0.000**
Brain phenotype*				**All**	
Overall brain phenotype	61 (98%)	82 (71%)	166 (95%)	309 (88%)	**0.000**
Motor impairments	56 (90%)	82 (71%)	159 (91%)	297 (85%)	**0.000**
Information on onset available	*27 (44%)*	*19 (17%)*	*61 (35%)*	*107 (34%)*	
Age at onset (median +IQR, years)	***1** (IQR 0.3–2)*	***1.5** (IQR 0.5–2.4)*	***1.3** (IQR 0.5–3)*	***1.1** (IQR 0.5–2.5)*	
Chorea without athetosis	21 (34%)	74 (68%)	101 (58%)	196 (56%)	**0.000**
Choreoathetosis	15 (24%)	4 (3%)	40 (23%)	59 (17%)	**0.000**
Muscular hypotonia	27 (44%)	23 (20%)	70 (40%)	120 (34%)	**0.000**
Cognitive impairments (overall)	28 (45%)	28 (24%)	45 (26%)	101 (29%)	**0.007**
Developmental delay (overall)	31 (50%)	25 (22%)	84 (48%)	140 (39%)	**0.000**
Severe intellectual disability	16 (26%)	5 (4%)	20 (11%)	41 (12%)	**0.000**
Learning difficulties	5 (8%)	24 (21%)	25 (14%)	54 (15%)	0.069
Microcephaly	6 (10%)	-	1 (1%)	7 (2%)	**0.000**
Thyroid phenotype*				**All**	
Overall thyroid phenotype	46 (74%)	54 (47%)	124 (71%)	224 (64%)	**0.000**
Overall hypothyroidism	46 (74%)	44 (38%)	118 (68%)	208 (59%)	**0.000**
Congenital hypothyroidism	21 (34%)	20 (17%)	53 (30%)	94 (27%)	**0.018**
Thyroid dysgenesis	8 (13%)	13 (11%)	26 (15%)	47 (13%)	0.670
Lung phenotype*				**All**	
Overall lung phenotype	36 (58%)	43 (37%)	82 (47%)	161 (46%)	**0.028**
IRDS	17 (27%)	17 (15%)	48 (28%)	82 (23%)	**0.030**
Pulmonary infections	19 (31%)	15 (13%)	30 (17%)	64 (18%)	**0.013**
Bronchial hyperreactivity / asthma	1 (1.6%)	11 (10%)	14 (8%)	26 (7%)	0.468
Interstitial lung disease	5 (8%)	10 (9%)	9 (5%)	24 (7%)	0.142
Additional clinical features*				**All**	
Overall additional clinical features	30 (48%)	8 (7%)	33 (19%)	71 (20%)	**0.000**
Dysmorphic features	10 (16%)	-	5 (3%)	15 (4%)	**0.000**
Hypo-/oligodontia	14 (23%)	1 (1%)	-	15 (4%)	**0.000**

*(Table 30 continued): Abbreviations, explanations, and symbols: *data are provided for the number of patients; **category with a patient who has two different variants and is therefore counted in two different categories. Calculations that were based on the total number of patients with available information are given in cursive. ANOVA = analysis of variance, C-TAD = C-terminal transactivation domain, HD = homeodomain, +IQR = median numbers are provided with interquartile ranges, IRDS = infant respiratory distress syndrome, n = number of patients, nHD = non-HD, nHD+ = nHD affecting the HD, NK2-SD = NK2-specific domain, N-TAD = N-terminal transactivation domain, WGD = whole gene deletion. Significant results of ANOVA ($p < 0.05$) are indicated in bold letters and on gray backgrounds.*

3.5.6.3.2 Disease severity scores for genotypic groups based on the type of variant

A subanalysis of disease severity scores for the three genotypic groups based on the type of variant and the composition of the AOV group and its respective subscores are detailed in Table 31. As shown, there were no significant differences between WGD and AOV. However, the comparison of both WGD and AOV with missense variants revealed significantly lower scores for the overall, brain, and thyroid scores in missense variants. In contrast, no significant differences were found for the lung phenotype between the groups.

Table 31: Disease severity scores for genotypic groups based on the type of variant

Type of variant		n	Overall score	Brain score	Lung score	Thyroid score
WGD		62	**11** (IQR 6–15)	**5** (IQR 2–8)	**1** (IQR 0–3)	**1** (IQR 0–2)
Missense		115	**5** (IQR 3–11)	**2** (IQR 0–7)	**0** (IQR 0–1)	**0** (IQR 0–1)
AOV:	**n**	174	**10** (IQR 6–14)	**5** (IQR 2–8)	**1** (IQR 0–2)	**0** (IQR 0–2)
Nonsense	57		***10** (IQR 7–14)*	***6** (IQR 3–10)*	***0** (IQR 0–2)*	***1** (IQR 0–2)*
Other deletions	56		***9** (IQR 5–12)*	***5** (IQR 2–7)*	***0** (IQR 0–2)*	***1** (IQR 0–2)*
Splice site variants	35		*7 (IQR 6–10)*	***4** (IQR 2–6)*	***0** (IQR 0–2)*	***1** (IQR 0–2)*
Duplications	17		***13** (IQR 10–16)*	***6** (IQR 4–8)*	***2** (IQR 2–5)*	***1** (IQR 0–3)*
Insertions	5		***16** (IQR 9–18)*	***8** (IQR 1–9)*	***1** (IQR 0–8)*	***2** (IQR 2–4)*
Stop loss	3		*3, 10, 11*	*2, 9, 10*	*0, 0, 0*	*0, 0, 0*
Indels	2		*13 and 23*	*1 and 3*	*7 and 13*	*2 and 4*
p-values						
WGD vs. missense			**0.000**	**0.004**	0.094	**0.016**
WGD vs. AOV			0.371	0.909	0.194	0.412
AOV vs. missense			**0.000**	**0.000**	0.529	**0.009**

Abbreviations: AOV = all other variants (i.e. not WGD or missense), IQR = interquartile range, n = number of patients, vs. = versus, WGD = whole gene deletions. Significant results ($p < 0.05$) are indicated in bold letters and on gray backgrounds. If there is a significant p-value for the comparison of two groups using the two-tailed unpaired Student's t-test, then the p-value is indicated on the handside of the respective group that has a higher score (i.e. it is written either on the left or right side of the column). P-values of comparisons without a significant difference between the groups are listed in the middle of the column.

3.5.6.3.3 ANOVA for genotype-phenotype correlation in genotypic groups based on the type of variant

Table 32 displays detailed results of the comparison of the three groups based on the type of variant for association with certain phenotypes using one-way ANOVA. Here, only the results of manifestations are shown, where significant differences in the overall *p*-values among the three groups (see Table 30 for the respective *p*-values) were identified. In summary, the following main observations were made using post hoc analysis with Tukey HSD of ANOVA:

1) In most categories, no significant differences between WGD and AOV were detectable.

Table 32: P-values of post hoc analysis with Tukey HSD of ANOVA for genotypic groups based on the type of variant

Criteria	*p-values*								
Mode of inheritance	**WGD**	**vs**	**Missense**	**WGD**	**vs.**	**AOV**	**Missense**	**vs.**	**AOV**
Familial variant			**0.000**			**0.000**		0.090	
De novo variant	**0.007**			**0.044**				0.555	
Organ combination									
Triad	**0.000**				0.348				**0.000**
Brain/Thyroid		0.331			0.746				**0.017**
Brain-only	**0.018**				0.696		**0.028**		
Thyroid-only			**0.001**		0.977		**0.000**		
Lung-only			**0.001**		0.981		**0.000**		
Brain phenotype									
Overall brain	**0.000**				0.785				**0.000**
Motor impairments	**0.002**				0.977				**0.000**
Chorea, no athetosis			**0.000**			**0.003**		0.530	
Choreoathetosis	**0.001**				0.973				**0.000**
Muscular hypotonia	**0.004**				0.880				**0.001**
Cognitive impairments	**0.010**			**0.011**				0.957	
Developmental delay	**0.001**				0.967				**0.000**
Severe intellectual disability	**0.000**			**0.006**				0.142	
Microcephaly	**0.000**			**0.000**				0.934	
Thyroid phenotype									
Overall thyroid	**0.001**				0.906				**0.000**
Overall hypothyroidism	**0.000**				0.630				**0.000**
Congenital hypothyroidism	**0.047**				0.859				**0.037**
Lung phenotype									
Overall lung	**0.023**				0.295			0.232	
IRDS		0.138			1.000				**0.031**
Recurrent pulmonary infections	**0.010**					**0.048**		0.633	
Additional features									
Overall additional features	**0.000**			**0.000**					**0.023**
Facial dysmorphism	**0.000**			**0.000**				0.439	
Hypo-/oligodontia	**0.000**			**0.000**				0.919	

Abbreviations: AOV = all other variants, IRDS = infant respiratory distress syndrome, n = number of patients, vs. = versus, WGD = whole gene deletions. Significant results ($p < 0.05$) are indicated in bold letters and on gray backgrounds. If there is a significant p-value for the comparison of two groups using the two-tailed unpaired Student's t-test, then the p-value is indicated on the handside of the respective group that has a higher score (i.e. it is written either on the left or right side of the column). P-values of comparisons without a significant difference between the groups are shown in the middle of the column.

2) *Inheritance:* Comparisons of the three groups revealed a significantly higher frequency of *de novo* variants in WGD and of familial cases in missense variants.
3) *Organ combinations:* Comparisons of both WGD and AOV with missense variants revealed an association of the organ triad with WGD and AOV. In turn, missense variants were found to be associated with thyroid-only and lung-only phenotypes.
4) *Brain phenotype:* Both WGD and AOV exhibited an association with the overall brain phenotype, motor impairments, choreoathetosis, muscular hypotonia, and developmental delay when compared to missense variants. Furthermore, cognitive impairments, severe intellectual disability, and microcephaly were found to be associated with WGD rather than the two other groups. Only chorea without athetosis showed a significantly higher frequency in missense variants compared to WGD and AOV.

5) *Thyroid phenotype:* The overall thyroid phenotype, hypothyroidism, and congenital hypothyroidism were found to be associated with WGD and AOV rather than missense variants.
6) *Lung phenotype:* WGD were found to have an association with the overall lung phenotype and recurrent pulmonary infections compared to missense variants. In addition, AOV displayed an association with IRDS in comparison to missense variants.
7) *Additional clinical features:* Overall, additional clinical features, facial dysmorphism, and hypo-/oligodontia were associated with WGD when compared to the two other groups.

3.5.6.4 Subanalysis of other regional subgroups

3.5.6.4.1 Disease severity scores depending on the impaired functional groups

A subanalysis of disease severity scores for variants with different numbers of affected functional domains is detailed in Table 33. The most important findings are as follows. 1) There were no significant differences between WGD and variants with four domain impairments. 2) HD-only variants have significantly lower overall scores than variants with four and three domains, lower thyroid scores in comparison to variants with impairment of four domains, three domains or two domains, and C-TAD-only and N-TAD-only variants. 3) There are no significant differences among the single-domain variants – C-TAD-only, N-TAD-only, and NK2-SD-only – for any of the scores. 4) Variants with no domain involvement have significantly lower lung scores than variants with three domains, two domains, HD-only, and C-TAD-only. A subanalysis of the variants with no domain involvement (see Table 28) revealed lung involvement in only 13% (1/8 patients; Patient 298 with a brain-lung phenotype; Table 28, Part 1). The other no-domain variants were found to be 63% brain-only (5/8 patients; Patients 297 and 299; Table 28, Part 1; and Patients 349–351; Table 28, Part 2), 13% thyroid-only (1/8 patients; Patient 288; Table 28, Part 1) and 13% brain/thyroid (1/8 patients; Patient 298; Table 28, Part 1) phenotypes. In addition, all scores of no-domain variants were lower compared to three domains, and the overall score also reduced compared to WGD.

Table 33: Disease severity scores for number of affected functional domains

Affected domains		n	Overall score	Brain score	Lung score	Thyroid score
WGD		62	**11** (IQR 6–15)	**5** (IQR 2–8)	**1** (IQR 0–3)	**1** (IQR 0–2)
4: HD, NK2-SD, C-TAD, N-TAD		48	**11** (IQR 8–14)	**6** (IQR 4–9.5)	**0.5** (IQR 0–2)	**1** (IQR 0–2)
3: HD, NK2-SD, C-TAD		78	**9** (IQR 6.8–12)	**4** (IQR 2–7)	**0** (IQR 0–2)	**1** (IQR 1–2)
2: NK2-SD, C-TAD		18	**9** (IQR 5.3–16.3)	**3** (IQR 2–8.3)	**1.5** (IQR 0–3.3)	**2** (IQR 0–2)
HD-only		99	**6** (IQR 3–11)	**2** (IQR 2–7)	**0** (IQR 0–2)	**0** (IQR 0–1)
C-TAD-only		30	**5.5** (IQR 5–13)	**3** (IQR 0–7)	**0** (IQR 0–2.3)	**1.5** (IQR 1–2)
N-TAD-only		5	**15** (IQR 5–16.5)	**8** (IQR 0–9.5)	**2** (IQR 0–3)	**2** (IQR 1–4)
NK2-SD-only		3	10, 14, 15	7, 8, 10	0, 0, 3	0, 1, 4
No domain		8	**7** (IQR 5–10.8)	**4** (IQR 2.5–8.8)	**0** (IQR 0–0)	**0** (IQR 0–0.8)
p-values						
4 domains vs.	**WGD**		0.882	0.191	0.275	0.471
	3 domains		**0.048**	**0.012**	0.157	0.605
	2 domains		0.552	0.111	0.190	0.496
	HD-only		**0.000**	**0.001**	0.965	**0.021**
	C-TAD-only		0.112	**0.005**	0.528	0.139
	N-TAD-only		0.822	0.541	0.564	0.081
	NK2-SD-only		0.495	0.443	0.835	0.61
	No domain		**0.038**	0.335	0.089	0.360
3 domains vs.	2 domains		0.621	0.824	0.057	0.690
	HD-only		**0.019**	0.104	0.166	**0.000**
	C-TAD-only		0.488	0.481	0.715	0.566
	N-TAD-only		0.292	0.784	0.200	0.077
	NK2-SD-only		0.161	0.054	0.825	0.691
	No domain		**0.000**	**0.032**	**0.011**	**0.013**
2 domains vs.	HD-only		0.076	0.553	0.271	**0.007**
	C-TAD-only		0.488	0.481	0.715	0.566
	N-TAD-only		0.668	0.781	0.852	0.175
	NK2-SD-only		0.470	**0.026**	0.531	0.904
	No domain		0.136	0.840	**0.007**	0.215
HD-only vs.	**C-TAD-only**		0.293	0.781	0.528	**0.001**
	N-TAD-only		0.120	0.516	0.672	**0.002**
	NK2-SD-only		0.096	0.087	0.867	0.156
	No domain		0.890	0.519	**0.008**	0.939
C-TAD-only vs.	N-TAD-only		0.384	0.469	0.940	0.347
	NK2-SD-only		0.087	0.077	0.769	0.941
	No domain		0.347	0.447	**0.037**	0.109
N-TAD-only vs.	NK2-SD-only		0.669	0.282	0.645	0.582
	No domain		0.205	0.910	0.065	0.111
No domain vs.	NK2-SD-only		**0.021**	0.171	0.535	0.478

Abbreviations: C-TAD = C-terminal transactivation domain, HD = homeodomain, n = number of patients, NK2-SD = NK2-specific domain, N-TAD = N-terminal transactivation domain, vs. = versus, WGD = whole gene deletions. Significant results ($p < 0.05$) are indicated in bold letters and on gray backgrounds. If there is a significant p-value for the comparison of two groups using the two-tailed unpaired Student's t-test, then the p-value is indicated on the handside of the respective group that has a higher score (i.e. it is written either on the left or right side of the column). P-values of comparisons without a significant difference between the groups are listed in the middle of the column.

3.5.6.4.2 Gene regions c.96–565 and c.753–881: phenotypes and disease severity scores

As shown in Table 34, there were no significant differences between the two gene regions for the frequency of any organ involvements and manifestations except for IRDS, which was significantly more often found in variants of the gene region c.753–881. The main difference between the two groups is, as previously outlined in Section 3.5.4.5.2, the involvement of the HD (93% in the gene region c.96–565 and 0% in the gene region c.753–881), while the distribution of AOV and missense is similar (see Table 34).

Table 34: Statistics for variants of the gene regions c.96–565 and c.753–881

Criteria	c.96–565 n (%)	c.753–881 n (%)	All n (%)	Student's t-test *(p-values)*
Total number of variants	44 (100%)	11 (100%)	55 (100%)	
Total number of patients	99 (100%)	18 (100%)	117 (100%)	
Genotypic groups based on the impaired gene region				
HD	-			
Non-HD+	92 (93%)	-	92 (79%)	
Non-HD	7 (7%)	18 (100%)	25 (21%)	
Genotypic groups based on the type of variant				
Missense	4 (4%)	2 (11%)	6 (5%)	
AOV	95 (96%)	16 (89%)	111 (95%)	
Organ manifestations*				
Information on thyroid available	99 (100%)	18 (100%)	117 (100%)	
Triad	39 (39%)	5 (28%)	44 (38%)	0.340
Brain/thyroid	27 (27%)	5 (28%)	32 (27%)	0.965
Brain/lung	4 (4%)	4 (22%)	8 (7%)	0.094
Lung/thyroid	2 (2%)	1 (6%)	3 (3%)	0.545
Isolated brain	22 (22%)	2 (11%)	24 (21%)	0.212
Isolated thyroid	3 (3%)	1 (6%)	4 (3%)	0.591
Isolated lung	2 (2%)	-	2 (2%)	0.547
Brain phenotype*				
Overall brain phenotype	92 (92%)	16 (89%)	108 (92%)	0.558
Motor impairments	88 (99%)	15 (83%)	103 (88%)	0.508
Chorea without athetosis	51 (51%)	9 (50%)	60 (51%)	0.907
Choreoathetosis	27 (27%)	2 (11%)	29 (25%)	0.078
Muscular hypotonia	38 (38%)	10 (56%)	48 (41%)	0.176
Cognitive impairments (overall)	25 (25%)	4 (22%)	29 (25%)	0.786
Developmental delay (overall)	53 (53%)	12 (67%)	65 (56%)	0.304
Severe intellectual disability	9 (9%)	3 (17%)	12 (10%)	0.434
Learning difficulties	15 (15%)	4 (22%)	19 (16%)	0.459
Microcephaly	1 (1%)	-	1 (1%)	0.672
Pituitary abnormalities	2 (2%)	-	2 (2%)	0.547
Thyroid phenotype*				
Overall thyroid phenotype	71 (71%)	12 (67%)	83 (71%)	0.685
Overall hypothyroidism	66 (66%)	12 (67%)	78 (67%)	1.000
Congenital hypothyroidism	32 (32%)	7 (39%)	39 (33%)	0.591
Thyroid dysgenesis	18 (18%)	4 (22%)	22 (19%)	0.690
Lung phenotype*				
Overall lung phenotype	47 (47%)	10 (56%)	57 (49%)	0.532
IRDS	23 (23%)	10 (56%)	33 (28%)	**0.019**
Pulmonary infections	16 (16%)	5 (28%)	21 (18%)	0.323
Bronchial hyperreactivity / asthma	8 (8%)	1 (6%)	9 (8%)	0.687
Interstitial lung disease	6 (6%)	1 (6%)	7 (6%)	0.934
Additional clinical features*				
Overall additional clinical features	16 (16%)	6 (33%)	22 (19%)	0.728
Dysmorphic features	2 (2%)	3 (17%)	5 (4%)	0.127
Hypo-/oligodontia	-			

*Abbreviations and symbols: *data are provided for the number of patients. IRDS = infant respiratory distress syndrome, n = number of patients. Significant results of ANOVA ($p < 0.05$) are indicated in bold letters and on gray backgrounds.*

In line with the similarities in the frequencies of organ manifestations (see Table 34) also the comparison of disease severity scores did not reveal any significant differences between the two gene regions – neither for the overall, nor for the organ-specific scores (see Table 35).

Table 35: Disease severity scores of variants of the gene regions c.96–565 and c.753–881

Affected regions	n	Overall score	Brain score	Lung score	Thyroid score
c.96–565 (incl. TN, N-TAD)	99	**9** (IQR 7–14)	**6** (IQR 2–8)	**0** (IQR 0–2)	**1** (IQR 0–2)
c.753–881(incl. Domain I, partly NK2-SD)	18	**11** (IQR 6–16)	**6** (IQR 4–9)	**2** (IQR 0–3)	**2** (IQR 0–2)
p-values					
c.96–565 vs. c.753–881		0.430	0.735	0.055	0.897

Abbreviations: Domain I = inhibitory domain, incl. = includes, n = number of patients, NK2-SD = NK2-specific domain, N-TAD = N-terminal transactivation domain, vs. = versus, TN = Tinman-domain, WGD = whole gene deletions. P-values of comparisons using the two-tailed unpaired Student's t-test without a significant difference (significant when p < 0.05) between the groups are shown in the middle of the column.

3.5.6.4.3 Disease severity scores in WGD with additional deletion of other genes

A comparison of disease severity scores of WGD with additional proximal (chromosomal coordinates hg19: < chr14:36.985.602) and distal (chromosomal hg19: > chr14:36.990.354) deletion to those with only additional distal deletion (see Section 3.5.4.4. for the respective patients with additional deletion of other genes) revealed significantly higher overall disease severity scores in WGD with additional proximal and distal deletion (Table 36).

Table 36: Disease severity scores in WGD with additional proximal and/or distal deletion of other genes

Affected gene regions		n	Overall score	Brain score	Lung score	Thyroid score
WGD with add. proximal and distal deletion		4	**17** (IQR 11–23)	**10** (IQR 5–13)	**2** (IQR 0–5)	**3** (IQR 1–5)
WGD with only add. distal deletion		24	**9** (IQR 5–15)	**6** (IQR 2–9)	**0** (IQR 0–1)	**1** (IQR 1–2)
p-values						
WGD with proximal and distal deletion vs.	WGD with only distal deletion		**0.026**	0.142	0.281	0.164

Abbreviations: add. = additional, n = number of patients, vs. = versus, WGD = whole gene deletions. Significant results (p < 0.05) are indicated in bold letters and on gray backgrounds. If there is a significant p-value for the comparison of two groups using the two-tailed unpaired Student's t-test, then the p-value is indicated on the hand-side of the respective group that has a higher score (i.e. it is written either on the left or right side of the column). P-values of comparisons without a significant difference between the groups are shown in the middle of the column.

3.5.6.5 Additional clinical features: frequencies in WGD, non-WGD variants and variants with additional deletion of other genes

Table 37 details the range of additional clinical features (see also Section 3.5.5.6 and Table 20 for the respective variants and Table 25 for the affected patients). Overall, they were more frequently found in patients with non-WGD variants (58%, 41/71 patients) that rarely had deletion of additional genes (5%, 2/41 patients). In contrast, WGD with additional clinical features nearly always involved other genes (97%, 29/30 patients). Among all patients with additional clinical features, the following genes were most frequently found to also be deleted: *PAX9* (31%, 22/71 patients), *NKX2-8* (21%, 15/71 patients), *RALGAPA1* (18%, 13/71 patients), *SLC25A21* (15%, 11/71 patients), and *SFTA3* (14%, 10/71 patients).

Significant differences are noticeable between the various additional clinical features in terms of the involvement of other genes. The differences between WGD and non-WGD, which are

displayed in detail in Table 37 and described next, have not been statistically analyzed due to different extents of deletion of other genes that might play a role in the phenotypes; therefore, the observations are descriptive. The main findings are as follows: 1) short stature, retarded skeletal age, fever of unknown origin and hyperphagia were often found in non-WGD variants that did not involve any other genes; 2) hypophagia was only found in non-WGD without involvement of other genes; 3) hypo-/oligodontia was nearly exclusively encountered in patients with WGD and additional deletion of *PAX9*; and 4) facial dysmorphism was often found in variants with concomitant deletion of *PAX9* and *RALGAPA1*.

Table 37: Additional clinical features and additional deletion of other genes

Additional clinical feature	Additional features		WGD		Non-WGD		Most frequent additionally deleted genes among all patients with the feature and additionally deleted genes
	Patients n (%)	+ n (%)	Patients n (%)	+ n (%)	Patients n (%)	+ n (%)	n (%)
All patients with additional clinical features	71 (100%)	**31 (44%)**	30 (43%)	**29 (97%)**	41 (58%)	**2 (5%)**	***PAX9*** *(71%, 22/31),* ***NKX2-8*** *(48%, 15/31),* ***RALGAPA1*** *(48%, 13/31),* ***SLC25A21*** *(35%, 11/31),* ***SFTA3*** *(32%, 10/31)*
Short stature	26 (37%)	**8 (31%)**	8 (31%)	**8 (100%)**	18 (69%)	-	***RALGAPA1*** *(75%, 6/8),* ***NKX2-8*** *(63%, 5/8),* ***PAX9*** *(63%, 5/8),* ***SFTA3*** *(63%, 5/8)*
Fever of unknown origin	20 (28%)	**8 (40%)**	4 (20%)	**4 (100%)**	16 (80%)		***NKX2-8*** *(63%, 5/8),* ***SFTA3*** *(50%, 4/8)*
Retarded skeletal age	18 (25%)	**6 (33%)**	6 (33%)	**6 (100%)**	12 (67%)		***NKX2-8*** *(100%, 6/6),* ***SFTA3*** *(100%, 6/6)*
Hypo-/oligodontia	15 (21%)	**14 (93%)**	14 (93%)	**14 (100%)**	1 (7%)		***PAX9*** *(93%, 13/14),* ***NKX2-8*** *(64%, 9/14),* ***SLC25A21*** *(57%, 8/14)*
Facial dysmorphism	15 (21%)	**10 (67%)**	10 (67%)	**9 (90%)**	5 (33%)	**1 (20%)**	***RALGAPA1*** *(70%, 7/10),* ***PAX9*** *(70%, 7/10)*
Hyperphagia	12 (17%)	**5 (42%)**	5 (42%)	**5 (100%)**	7 (58%)	-	***NKX2-8*** *(100%, 5/5),* ***SFTA3*** *(80%, 4/5)*
Congenital heart defects	9 (13%)	**4 (44%)**	3 (33%)	**3 (100%)**	6 (67%)	**1 (17%)**	***SFTA3*** *(75%, 3/4)*
Hypophagia	7 (10%)	-			7 (100%)	-	-
Hyperextendable joints	4 (6%)	**1 (25%)**	1 (25%)	**1 (100%)**	3 (75%)	**1 (33%)**	*multiple genes (25%, 1/4)*
Hydronephrosis	2 (3%)	**1 (50%)**	2 (100%)	**1 (50%)**	-		*multiple genes (50%, 1/2)*
Megabladder	2 (3%)	**1 (50%)**	1 (50%)	**1 (100%)**	1 (50%)	-	*multiple genes (50%, 1/2)*
Osteoporosis	2 (3%)	**2 (100%)**	2 (100%)	**2 (100%)**	-		***NKX2-8*** *(100%, 2/2),* ***PAX9*** *(100%, 2/2),* ***SLC25A21*** *(100%, 2/2)*
Hypogammaglobulinemia	1 (1%)	**1 (100%)**	1 (100%)	**1 (100%)**			*multiple genes (100%, 1/1)*
Developmental dysplasia of the hip	1 (1%)	**1 (100%)**	1 (100%)	**1 (100%)**			
Cryptorchism	1 (1%)	**1 (100%)**	1 (100%)	**1 (100%)**			
Hexadaktylie	1 (1%)	**1 (100%)**	1 (100%)	**1 (100%)**			
Maldescensus testis	1 (1%)	**1 (100%)**	1 (100%)	**1 (100%)**			
Vesicouretral reflux	1 (1%)	**1 (100%)**	1 (100%)	**1 (100%)**			

Symbols and Abbreviations: += deletion of additional genes present, n = number of patients, non-WGD = variants that are not WGD, WGD = whole gene deletions.

A subanalysis of patients with additional deletion of *RALGAPA1* (5%, 19/351 patients; Patients 1, 3–4, 11–12, 30–36 and 38–42; Table 25, Part 1), which is known to be associated with a phenotype involving developmental delay, including speech delay, severe intellectual disability, microcephaly, muscular hypotonia, and seizures[260], showed developmental delay in 63% (12/19 patients), muscular hypotonia in 53% (10/19 patients), severe intellectual disability in 37% (7/19 patients), and microcephaly in 21% (4/19 patients) of patients with *NKX2-1* variants. However, no seizures or speech delay was found in these patients.

3.5.6.6 Disease severity scores and phenotypes in multigenerational families

In this study, 38 multigenerational families were identified, and the clinical severity scores of the family members were compared. As indicated in Table 38, in half of the families (53%, 20/38 families; highlighted on gray backgrounds in Table 38), there was at least a two-point increase in the overall disease severity score[55,61,64,68,182,206,209,212-214,217,219, 221-223,226,231,232,243], partially in conjunction with an increase in the number of organ manifestations[55,61,68,213]. However, there were also 18 families with similar (29%, 11/38 patients)[61,64,193, 201,207,213,217,222,223,242,249] or less severe (18%, 7/38 patients)[62,64,65,191,192,194, 209,217,220, 236] manifestations compared to the previous generation/s.

Table 38: Multigenerational families: phenotypes and disease severity scores

Variant	First generation		Second generation		Third generation	
	Phenotype	Score*	Phenotype	Score*	Phenotype	Score*
14q13[207]	brain/thyroid	5	brain/thyroid	5		
14q13.1[68]	**brain/thyroid**	**8**	**triad**	**18**	**triad**	**18**
14q13.3[206]	**triad**	**9**	**brain/thyroid**	**11**		
c.174_180del;p.(Met59Alafs*40)[222]	**brain-only**	**6**	**brain-only**	7, **9**		
c.294C>G;p.(Tyr98*)[55]	**brain/thyroid**	**5**	**triad**	**8, 14**		
c.319C>T;p.(Gln107*)[223]	**brain-only**	**3**	**brain-only**	3, **12**		
c.390C>G;p.(Tyr130*)[236]	triad	17	triad	16		
c.428G>A;p.(Trp143*)[243]	**triad**	**17**	**brain-only**	**20**		
c.463+2T>C;p.?[226]	**triad**	**9**	**triad**	**17**		
c.463+3_463+6del;p.?[61]	**brain/thyroid**	**7**	**brain/thyroid**	**9**		
			triad	**11**		
c.464-2A>G;p.?[212]	**brain/thyroid**	**7**	**brain/thyroid**	**9**	**triad**	**11, 15**
c.464-2A>T;p.?[65]	brain-only	7	brain-only	3		
c.464-9C>A;p.(Ile155Thrfs*286)[213]	brain/thyroid	7	brain/thyroid	7		
c.464-9C>A;p.(Ile155Thrfs*286)[213]	**brain/thyroid**	**7**	**triad**	8, **10**		
c.469delC;p.(Arg157Alafs*9)[214]	**brain/thyroid**	**10**	**brain/thyroid**	**12**	**brain/thyroid**	**20**
c.524C>A;p.(Ser175*)[194]	brain/thyroid	8	brain/thyroid	5	triad	7
c.572G>T;p.(Arg191Leu)[191,192]	thyroid/lung	6	lung-only	4		
	lung-only	1	lung-only	1		
c.583delC;p.(Arg195Glyfs*33)[64]	**triad**	**7**	**triad**	5, 7, 8, **12**		
c.613G>T;p.(Glu205*)[209]	**brain/thyroid**	**5**	**triad**	**16**		
c.617T>A p.;(Leu206Gln)[217]	brain/lung	11	triad	17	brain/thyroid	10
c.617T>A p.;(Leu206Gln)[217]	**brain/lung**	**11**	brain/lung	9	thyroid/lung	7
					triad	**18**
c.617T>A p.;(Leu206Gln)[217]	brain/lung	11	brain-only	8	brain/thyroid	9
					triad	11
c.623G>C;p.(Arg208Pro)[219]	**triad**	**9**	**triad**	**11**		
c.626G>C;p.(Arg209Pro)[242]	brain/thyroid	7	brain/thyroid	8		
c.631A>T;p.(Lys211*)[62]	brain/thyroid	14	triad	9		
c.650C>A;p.(Ser217*)[210]	brain/lung	10	brain/thyroid	6	triad	7
					brain/thyroid	10
c.671T>G;p.(Leu224Arg)[61,193]	triad	18	triad	18	triad	14
c.714delG;p.(Trp238Cysfs*9)[201]	brain/thyroid	10	brain/thyroid	10, 11		
c.727C>T;p.(Arg243Cys)[220]	brain-only	10	brain-only	5,8		
c.728G>C;p.(Arg243Pro)[61]	brain-only	8	brain-only	9		
c.732C>A;p.(Tyr244*)[61]	**brain-only**	**6**	**brain/thyroid**	**8**		
c.745C>T;p.(Q249*)[221]	**brain-only**	**3**	**brain-only**	**8**		
c.915del;p.(Ala306Argfs*75)[182]	**brain/lung**	**5**	**brain/lung**	**12, 13**		
c.978_1056del; p.(Ala330Serfs*25)[231]	**brain-only**	**3**	**brain-only**	**5**		
c.980_986del;p.(Ala327Glyfs*52)[232]	**brain/thyroid**	**5**	**brain/thyroid**	**7**		
c.1106C>T;p.(Ala369Val)[249]	thyroid-only	5	thyroid-only	2, 5		
	thyroid-only	5	thyroid-only	2		

*Symbols: *The provided scores are the overall scores for the respective patients. Families with at least a two-point increase in the disease severity are highlighted in bold letters and on gray backgrounds.*

4 Discussion

NKX2-1 haploinsufficiency is a rare and, so far, not well-understood genetic disorder with a predominantly neurological phenotype, mainly involving movement disorders resulting from disturbances in the basal ganglia circuitry. As analyzing the impairments caused by deficiency of *NKX2-1* on the neuroanatomical level in these patients appears to be difficult, animal models are necessary to study the role of Nkx2-1/*NKX2-1* and the effects of its absence. In the past few years, some progress has been made on both fronts: understanding Nkx2-1 in mice and gaining further insight into the brain-lung-thyroid syndrome in humans. However, the mechanisms underlying the syndrome are yet to be understood, and knowledge on the transcriptional targets of *NKX2-1* in the human brain is scarce. Moreover, the high number of different variants found in the patients (see Sections 3.5.4 and 3.5.6) and the fact that, to date, no genotype-phenotype correlation has been identified, pose additional challenges when aiming at a comparison between the results obtained in our mice and findings in humans. Therefore, in this study, I performed an extensive literature review, including the first systematic genotype-phenotype correlation analysis in addition to my mice studies. In the following section, I first address the results of the mouse studies, then present the conclusions from the literature review, and finally discuss the extent to which these results in mice are comparable to findings in humans and can therefore contribute to a better understanding of *NKX2-1* haploinsufficiency.

4.1 The role of Nkx2-1 in the mouse brain

In mice, our working group has previously already shown that Nkx2-1 is mainly expressed by cholinergic and PV-containing GABAergic neurons[1,2] of the BF, which are known to play a major role in movement and higher cognitive functions. Accordingly, the loss of these neuronal populations can result in severe impairments of both systems, in particular coordinated movement, learning, and memory impairments[3,36,261]. Using two different mutant mouse lines, our group demonstrated that (a) prenatal GAD67-driven deficiency leads to severe loss of both cholinergic and GABAergic PV-containing neurons resulting in disturbances in the basal ganglia circuitry and impairments in spatial memory, and (b) postnatal ChAT-driven inactivation leads to a significant loss of cholinergic BF neurons impairing the cholinergic centers. We have further demonstrated that ongoing Nkx2-1 expression is essential for the integrity of cholinergic BF neurons, which otherwise degenerate[1]. Moreover, *NKX2-1* was detected in the basal ganglia of the human brain (postmortem study of controls without known brain disease)[1], indicating a postnatal function of the transcription factor also in humans. The indispensable role of Nkx2-1 for the cholinergic system in mice prompted the idea that its lack might be accompanied by

impairments in the neurotrophin system that is required for proper development and function of cholinergic neurons.

Since these neurons cannot survive without access to *NGF*, we hypothesized that Nkx2-1 deficiency might initially impair the neurotrophin system, thereby causing loss of cholinergic neurons in both mutant lines. For this reason, I analyzed the expression levels of *NGF* and its receptors *trkA* and $p75^{NTR}$ in the pre- and postnatal mutant line in the present study. To determine the timepoint at which postnatal mutation occurs in the cholinergic population and to subsequently decide when to perform the expression study of *NGF* and its receptors, the postnatal mutation was monitored by studying the Cre-recombinase activity using ROSA LacZ reporter mice (see Section 3.1). The results are outlined in the following section.

4.1.1 The ChAT-Cre driven postnatal mutation affects already differentiated cholinergic neurons between P5 and P15

The postnatal deletion is performed under the ChAT promoter, (i.e. during ChAT-upregulation, which occurs between P5 and P15)[99,102,103]. Accordingly, X-gal, which is an indicator of the Cre activity and therefore measures the recombination extent, was not detected before P4/5 (see Figure 9A). In line with the upregulation of ChAT, at P8 strong X-gal labeling (or β-gal immunoreactivity which is an equivalent method) was detectable in the cholinergic centers (see Figure 9B–H). To further characterize the Cre-positive cells, Nkx2-1 co-labeling was performed at P8, resulting in the typical distribution pattern of Nkx2-1-positive cells in the BF[2] and strong co-localization of X-gal and Nkx2-1 in ChAT-Cre targeted regions (see Figure 10).

Since the number of cholinergic neurons reaches the level of adult mice after the second postnatal week[99,103] (cf. Section 1.4), labeling experiments were also performed at P15. As expected, at P15 a high density of cholinergic neurons positive for β-gal was found in all regions known to contain Nkx2-1-positive cholinergic neurons[2] (see Figures 11 and 12). These observations are also consistent with the description of the ChAT-Cre line by the Jackson Laboratories (Bar Harbor, ME, USA), which detected X-gal in all cholinergic BF neurons.

Interestingly, ChAT double-labeling did not always result in co-labeling with X-gal, and, cholinergic neurons were rarely found to be X-gal negative (see Figure 9C–E, indicated by black arrows). This can be explained by the heterogeneity of cholinergic BF neurons, which constitute three subpopulations (as described in Magno et al., 2011)[2]: (1) a Nkx2-1-positive population that is lost in postnatal mutants, thus accounting for a 50% reduction in cholinergic neurons; (2) a large subpopulation that remains unaffected in the ChAT-Cre line due to Nkx2-1 downregulation before the postnatal mutation occurs (i.e. prior to ChAT and Cre-upregulation); and (3) a small subpopulation deriving from Nkx2-1-negative domains[83,84]. The latter one most

likely accounts for the small percentage of cholinergic neurons that remains unaffected in the prenatal line despite Nkx2-1 deletion. More evidence for the heterogeneity of these neurons is provided by the fact that degenerating neurons can be observed near intact cholinergic BF neurons (results of our working group, not presented here)[2] that belong to the second subpopulation.

In addition, low co-localization of X-gal and Nkx2-1 was observed in the LGP (see Figure 10A and B) a region known for its high density of predominantly Nkx2-1-positive PV-containing GABAergic neurons[2] – which might be due to the cholinergic sparser neurons as no co-labeling with PV was found, confirming that PV-containing GABAergic neurons are not targeted by the postnatal mutation.

In line with the confirmation that the mutation occurs between P5 and P15 (this study), there was no loss of cholinergic neurons before P15 in the postnatal line[1,2] (results of our working group, not part of this study). Hence, P15 was chosen as the main investigation time point for the qRT-PCR study.

4.1.2 Degeneration of cholinergic BF neurons cannot be prevented by maintained *NGF* level due to *trkA* downregulation

The synthesis of *NGF* has been shown to not be impaired by pre- or postnatal deletion of Nkx2-1 (see Section 3.2), which can be explained by the presence of unaffected neocortical/allocortical GABAergic neurons that constitute the major sources of *NGF* in the mouse brain[4,5] (see Figure 28 for relative *NGF* expression in the mouse brain).

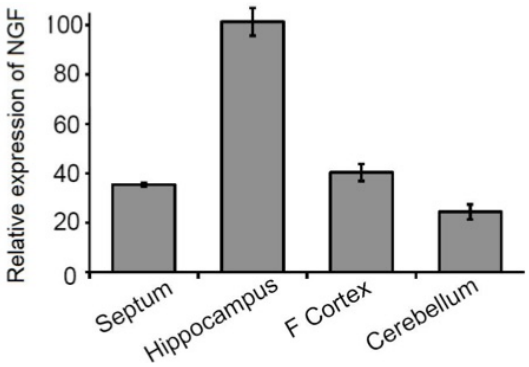

Figure 28. Relative NGF mRNA expression at P5 in selected brain regions of CD-1 mice*. For the analysis, RNA was isolated from four different brain regions at P5, and NGF expression was measured and normalized to Actb expression. Abbreviations: F Cortex = frontal cortex, NGF = nerve growth factor. The modified figure has been taken from Schnitzler et al. (2010)[154] and reprinted with permission from the Journal of Neuroscience: Society of Neuroscience. BMP9 (bone morphogenetic protein 9) induces NGF as an autocrine/paracrine cholinergic trophic factor in developing basal forebrain neurons.* ***Schnitzler et al. (2010).*** *Copyright © 2010.*

While this observation in the postnatal line is because the ChAT-Cre driven mutation only occurs in cholinergic cells, it can be explained as follows in the prenatal line: although most of the neocortical GABAergic interneurons originate from the MGE[43,93,94], which is well known for its strong expression of Nkx2-1[43,262], cortical neurons no longer express Nkx2-1 when they become postmitotic and start migrating[2,27]. Nobrega-Pereira et al. (2008)[27] provided an explanation for this by demonstrating that postmitotic GABAergic interneurons downregulate Nkx2-

1 as a prerequisite for tangential migration to the cortical mantle[83,85,91]. Fate-mapping studies have shown that two thirds of these cortical interneurons contain PV, and the remaining third contain SST[88,94-96]. Both neuronal populations primarily originate from the MGE[96] (cf. Section 1.3.2). Among allocortical hippocampal PV-containing neurons, only a subgroup seems to depend on Nkx2-1, and, accordingly, only a small proportion gets lost in the prenatal mutant line[1].

Interestingly, and in contrast to other studies, the endogenously available amount of *NGF* in our mutants (see Figure 13) could not prevent cholinergic neurons from degenerating. This is surprising, since postlesional *NGF* administration has been found to rescue axotomized cholinergic neurons, for example after fimbria-fornix transection (FFT)[263,264]. For instance, in a rat model for Alzheimer's disease Gu et al. (2009)[264] demonstrated that recombinant human *NGF* can promote survival of MSvDB neurons after FFT and improve the initially impaired spatial learning and memory in the Y-Maze-test. Another study with a mouse model for Huntington's disease, where the cholinergic deficit was accompanied by spatial memory deficits and a progressive, age-dependent reduction in the *NGF* level, indicated that *NGF* infusions result in an upregulation of cholinergic markers in the hippocampus and an improvement in spatial memory in the radial arm maze, thus suggesting plasticity of the SHS[265].

One reason that the maintained *NGF* level in our mutants could not rescue the cholinergic neurons might be the loss of *NGF* receptors (see Figure 13), which are necessary for mediation of *NGF* effects. The levels of receptor downregulation observed in both transgenic lines (Figure 13A prenatal line: *trkA* -82% and $p75^{NTR}$ -60%; Figure 13B postnatal line: only significant for *trkA* -56%) appear to be proportional to the respective extent of neuronal cell losses at P15 and 3 months[1,2] (53-98% loss of cholinergic neurons in the prenatal mutants depending on the region and about 50% overall loss of cholinergic neurons in postnatal mutants). The heavy impairment of cholinergic BF neurons despite relevant *NGF* synthesis in our mice is in line with findings in brain autopsies of Alzheimer's patients, where the *NGF* level was found to be maintained at the control level[266,267]. Similar results were obtained in the aging rat brain[269]. Furthermore, in an Alzheimer's model (TgMNAC13 transgenic mice) where anti-*trkA* antibodies were used to block *trkA* signaling without manipulation of *NGF*, a major loss of cholinergic neurons was observed that result in a significant visual memory deficit in the object recognition test[270].

In view of the research conducted on aging and Alzheimer's brains in humans, mice, and rats, it is presumable that the cognitive impairments in our postnatal mutants are primarily attributable to the loss of *trkA* receptors in the BF. The loss of striatal *trkA*-positive/$p75^{NTR}$-negative

cholinergic neurons[144] seems to play a subordinate role. Overall, *p75^{NTR}*, which was not significantly altered in our postnatal line, does not seem to be involved in the cognitive impairments. This is supported by the finding of an unaltered *p75^{NTR}* level in human Alzheimer's brains[268], and the study of Sanchez-Ortiz et al. (2012)[271], who demonstrated that additional deletion of *p75^{NTR}* in *trkA* knockout mice does not increase cognitive impairments. Their study provided further support for the prominent role of *trkA* by showing that a lack of *trkA* disturbs the formation of the septohippocampal projections, especially during the second postnatal week[271]. As an underlying mechanism, they proposed the interaction of *trkA* with extracellular ERKs[271]. ERK activation was shown to occur analogously to ChAT upregulation; that is, while no ERK activation is found at P5, the maximal activation is reached at P14 when the SHS is fully established (cf. Section 1.4.2) with dependency on *trkA*/*NGF* signaling. Accordingly, severely impaired septohippocampal projections were observed in *trkA*-deficient mice[271]. These were most likely caused by the loss of ERK activation resulting from *trkA* deletion[271]. On the level of cognition, the authors also found significant impairments in the novel object recognition test[271].

A *trkA*-mediated positive feedback loop involving ERKs and Lhx7 has recently been suggested for MSvDB cholinergic neurons[108], as described in detail in Section 1.3.4.1 (see also Figure 6). Tomioka et al. investigated the MSvDB neurons of rats *in vivo* and *in vitro*, finding that (a) Lhx7 directly regulates *trkA* transcription by binding to its AT-rich consensus sequence (i.e. its overexpression increases *trkA* activity, and its inactivation results in a reduced expression level); (b) *NGF* potentiates this Lhx7-stimulated *trkA* expression via the ERK pathway; and (c) Lhx-7 expression is stimulated by *NGF* via the PI3K pathway[108].

Consistent with these findings, in the prenatal line, the subtotal loss of cholinergic neurons in the striatum was accompanied by total loss of Lhx7 at 3 months of age[1], and SST-positive striatal neurons which are known to maintain Lhx6-expression[1], were not found to be affected (results of our working group). In another study[31], deletion of Ldb-1, which mediates the effects of both Lhx6 and Lhx7, resulted in substantial loss of both cholinergic and PV-containing GABAergic neurons, comparable to the loss in our prenatal mutants. It is, thus presumable that a reduction in Ldb-1 could also be found in our mutants. However, this was not investigated.

4.1.3 Neuronal loss leads to loss of axonal fibers and axonal sprouting in both lines

Our group has demonstrated that the loss of cholinergic neurons in both mutant lines is accompanied by the loss of cholinergic axonal fibers using acetylcholine esterase (AChE) staining in 3-month-old mice (see Figure 29, published in Magno et al., 2011, not part of this study)[1].

While there was a vast loss of axonal fibers in all of the analyzed regions (neocortex, hippocampal formation, and the striatum; see Figure 29B, F, and J) in prenatal mutants, this loss in the postnatal mutant line was only extensive in the neocortex and less pronounced in the hippocampal formation (see Figure 29H for "patch-like" AChE labeling pattern caused by the projection defects).

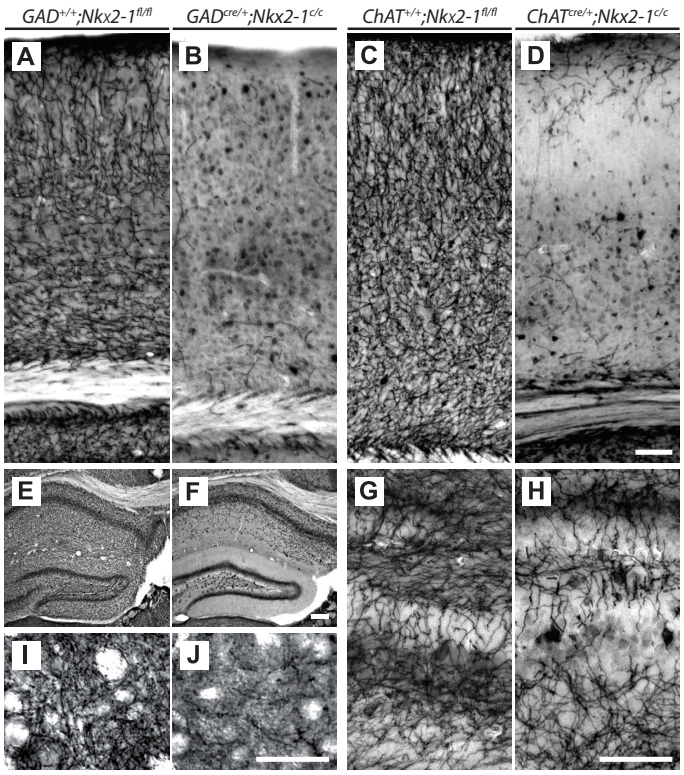

Figure 29. Loss of axonal fibers and target denervation in both mutant lines at 3 months. *Acetylcholine esterase (AChE) histochemistry is shown for Nkx2-1$^{fl/fl}$ controls (A, C, E, I, G), prenatal GAD$^{cre/+}$;Nkx2-1$^{c/c}$ mice (B, F, J), and postnatal ChAT$^{cre/+}$;Nkx2-1$^{c/c}$ (D, H). While intense labeling was observed in the somatosensory cortex (A, C), the hippocampal formation (E, G), and the CPu (I) in the controls, this pattern was severely reduced in the neocortex of both mutant lines (B, D). In the prenatal mutants, a severe impairment was also found in the hippocampal formation (F showing the gyrus dentatus, CA3, and CA1, and H showing the gyrus dentatus at higher magnification) and the CPu (J). In postnatal mutants, these observations were less severe (D,H). Scale bars: 100 µm (A–D) and 50 µm (E–J). Similar results were observed at P15 and 1 month of age (not shown here). Figure published in Magno et al. (2011)[1]. Reprinted with permission from the European Journal of Neuroscience: John Wiley and Sons, The integrity of cholinergic basal forebrain neurons depends on expression of Nkx2-1.* ***Magno et al. (2011).*** *Copyright © 2011.*

In general, after extensive neuronal loss or damage (e.g. following spinal cord injury or experimentally induced axotomy), compensatory axonal sprouting in terms of plasticity and reorganization can be observed. This is mediated by various factors, such as *GAP-43*[272]. *GAP-43* exerts a major role in the guidance of axonal outgrowth and the formation of synaptic contacts during development[273], and it is particularly important for the SHS as demonstrated in heterozygous knockout mice that develop neither a corpus callosum nor a hippocampal commissure[274]. Moreover, Haas et al. (2000)[275] have shown that *GAP-43* is upregulated in axotomized adult medial septal cholinergic and GABAergic neurons of the SHS after FFT, indicating their regenerative potential.

Naumann et al. (2002)[103] investigated the role of axonal sprouting in the SHS. Using the neurotoxin 192-IgG-saporin, which is coupled with a *p75^{NTR}* antibody causing its uptake by cholinergic neurons, they demonstrated that a unilateral entorhinal cortex lesion results in ipsilateral sprouting of cholinergic septohippocampal axons (shown by increased AChE staining), and

vice versa, a lesion of the septohippocampal projection leads to sprouting of the perforant pathway of the entorhinal cortex. Moreover, hippocampal GABAergic interneurons were found to sprout as well[276].

To investigate whether axonal sprouting is also present in our mutants, I analyzed *GAP-43* expression together with *NGF* and its receptors at P15 (cf. Section 3.3). Our expectation was to find *GAP-43* upregulation in the mutants as an indicator of axonal sprouting. However, the contrary finding (i.e. downregulation of *GAP-43*) would also have been plausible because of the reduced *trkA* and $p75^{NTR}$-level, as $p75^{NTR}$-sustained *trkA* autophosphorylation is a transcriptional activator of *GAP-43* (via the mitogen-activated protein-kinase isoforms ERK1 and 2) and thereby speeds up axonal outgrowth[277]. Although we did not find any major differences between the two mutant strains and their controls (see Figure 14), our results support the initial hypothesis for the following reason: the presence of a *GAP-43* expression level comparable to controls in the mutants, despite the extensive loss of BF cholinergic neurons and PV-containing GABAergic neurons and their fibers (see Figure 29), points to synaptic reorganization of surviving neuronal assembles rather than transient sprouting processes by the affected neurons. In line with our observations made by AChE staining in adult mice, the maintained *GAP-43* level is not accompanied by the maintenance or reestablishment of already affected connections. These data suggest that loss of both cholinergic and PV-containing GABAergic neurons, including their developing axonal projections, are compensated for by the sprouting processes of other, likely neighboring, neuronal systems – at least at the time point analyzed.

Another striking finding of the *GAP-43* expression analysis in this study was the overall higher expression level in wild-type mice compared to both mutants and controls of the pre- and postnatal mutant line. The analysis in wild-type mice of the same age was performed to have an additional control group for comparison in case of upregulation in the mutants. The reasons for the higher expression levels in wild-type mice are not known. However, it reasonable to assume that the overall lower expression levels in both mutants and controls are attributable to the genetic modifications in the transgenic lines.

4.1.4 Motor deficits and spatial memory impairment

As described in Magno et al. (2011)[1], our group revealed that the reduction of cholinergic and PV-containing GABAergic neurons in the prenatal line led to (1) severe impairments of the motor system due to disturbances in basal ganglia circuits and (2) deficits regarding spatial memory. In line with the cell counts, the impairments were attenuated in postnatal mutants. Motor deficits in the prenatal line were detected by a significantly reduced performance on the Rotarod test and spatial memory deficits in the Morris Water Maze, where the mutants failed

to learn the location of the platform. No motor impairments were found in the postnatal line using the Rotarod test. This finding is attributable to the fact that GABAergic neurons are not affected by the postnatal mutation. Interestingly, impairments of spatial memory in the postnatal line were only significant in female ChAT-Cre mice. The causes for this observation remain unclear, as no sex-specific effects of the mutations could be identified[1]. In a recent fate-mapping study, Magno et al. (2017) provided further evidence for the role of cholinergic BF neurons and the SHS in spatial learning and memory using Zic4-Cre mice with focal deletion of Nkx2-1[3]. The mutants presented with initial hyperactivity, and when exposed to cognitive tests, they exhibited significant impairments of short- and long-term memory in the novel object recognition test. They also had severe impairments of spatial learning and memory in the T-Maze test, while motor behavior, social recognition, and social memory were normal[3]. Magno et al. (2017) further investigated the theta rhythm of the hippocampus, which is an oscillatory brain activity that is associated with cognitive functions such as memory and spatial learning (cf.[3,115]), and found a shift in the oscillation frequency and running speed in these mice[3], which might cause impairments of hippocampal processing and memory formation[110].

The finding of the present thesis – namely, that there is a greater loss of *NGF* receptors in our prenatal mutants compared to the postnatal mutant line, which is in accordance with the cell counts (see above) – can be explained by the heterogeneity of cholinergic BF neurons[1]. While the remaining level of *trkA/p75^{NTR}* expression in both mutant lines is attributable to the Nkx2-1 negative subpopulation, the overall greater loss of cholinergic neurons and *trkA/p75^{NTR}* in the prenatal mutant line is owed to the fact that the GAD-Cre mutation also targets the cholinergic subpopulation that would otherwise later downregulate Nkx2-1 prior to migration (cf. Section 1.3.4.1). The less severe findings in postnatal mutants are also in line with other studies showing that manipulation of mature neurons leads to less pronounced impairments of higher cognitive functions[261]. However, the attenuated effects in postnatal mutants could also be explained by the mere fact that only approximately 50% of cholinergic neurons are lost and by the hypothesis that the remaining neurons can compensate for the loss. Either way, the results in both lines suggest that the neurological impairments in mice are proportional to the extent of neuronal loss.

A proportional relationship between the severity of cognitive impairments and the level of *trkA* reduction was also shown in another interesting neurodegeneration study[278]. Saragovi et al. demonstrated that the loss of *trkA* expression precedes the loss/ or degeneration of cholinergic BF neurons, and they therefore suggested that *trkA* could be a useful biomarker in assessing the

progression of neurodegenerative disorders such as Alzheimer's disease[278]. These findings further support the theory that *trkA* downregulation not only reflects the loss of the cholinergic BF population, but is also responsible for the populations' degeneration.

What remains to be elucidated is whether the GABAergic cell loss additionally contributes to spatial memory deficits, as their role in memory and learning is less investigated. In contrast to cholinergic neurons, which depend on *NGF*, no role of *NGF* or its receptors has been reported for GABAergic neurons. The downregulation of *NGF* receptors in prenatal mutants can thus be fully attributed to the cholinergic population. In GABAergic septal neurons and the hippocampal regions, *BDNF* and *trkB* expression was detected, and *trkB*-mediated *BDNF* signaling was found to play an important role during the development of these regions[164,279]. However, no such expression or dependency was found for cholinergic MSvDB neurons[165]. Our finding that there was no significant reduction in the expression of the *BDNF* receptor *trkB* following either of the mutations (cf. Section 3.4; see Figure 15) can be explained by two different reasons: either (1) *trkB* expression is not affected, which is highly likely in the postnatal line, or (2) the reduction cannot be detected in total brain preparations due to the overall high expression level of *trkB* in the brain compared to endogenous genes and other neurotrophin receptors, which might be true for both mutant lines. In hindsight, qRT-PCR of brain biopsies of the regions of interest would have been a more suitable technique to address this particular question. However, it could not be performed for technical reasons.

4.1.5 Cholinergic neurons die before they can gain access to *NGF*

The studies of our group have revealed a substantial loss of cholinergic neurons in both mutant lines[1,2] and an additional loss of Nkx2-1-dependent GABAergic neurons in the prenatal GAD-Cre line. We demonstrated for the first time, using electron microscopy (results of our group, not part of this thesis), that both neuronal groups in the prenatal line die rapidly (i.e. in the subventricular zone). Moreover, no evidence was found for a transformation of cholinergic neurons into other phenotypes or a replacement by other neurons[1]. Fine-structural analysis of the ventricular zone at E14.5 further revealed a high number of electron-dense cells with severe signs of degeneration (see Figure 30). In short with the findings on the expression of neurotrophic factors in the present study, it can be concluded that neurons in the prenatal line die before they can migrate and gain access to *NGF*.

In the postnatal line, however, cholinergic neurons were found to degenerate slowly after migrating to their final destination (results of our working group, not part of this thesis). Furthermore, severe signs of degeneration already appeared at the age of one month (see Figure 31)

and interestingly, the onset of the postnatal mutation was identified to be at a time point when the expression of the *NGF* receptors *trkA* and *p75*^{NTR} is at its peak (present study).

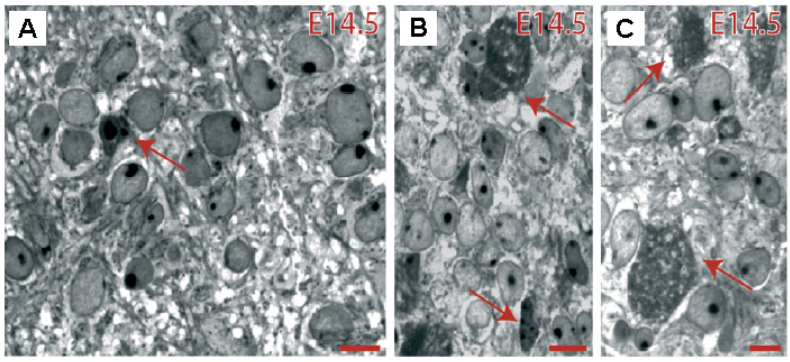

Figure 30. Electron microscopy of the subventricular zone of controls and prenatal mutants at E14.5*. In controls (A), apoptotic cell profiles (arrows) were rarely encountered. In prenatal mutants (B and C), many cells with severe signs of degeneration and cell debris were found (arrows). In these cell profiles, the nuclei and other cell organelles were rarely detectable. Scale bars: 10µm. Figure published in Magno et al. (2011)[1]. Reprinted with permission from the European Journal of Neuroscience: John Wiley and Sons, The integrity of cholinergic basal forebrain neurons depends on expression of Nkx2-1.* ***Magno et al. (2011).*** *Copyright © 2011.*

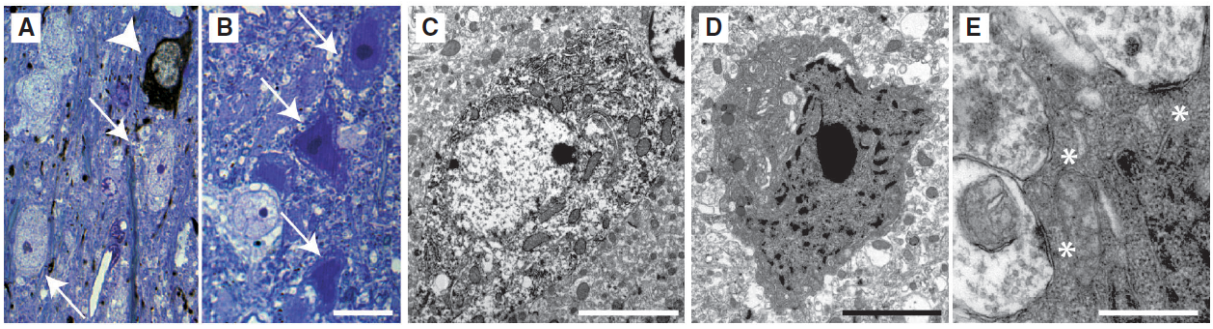

Figure 31. Semithin sections after immunohistochemical (IHC) analysis for ChAT and electron microscopy of the medial septum (MS) in postnatal mutants and controls at 1 month*. Controls (A, C) and postnatal mutants (B, D, E) are depicted. Scale bars: 20 um (A and B), 10 µm (C), 6 µm (D), and 2 µm (E). In controls (A), one neuron with ChAT labeling (arrowhead) and several unlabeled neurons are shown (arrows). In postnatal mutants (B), a high density of dark-colored neurons (arrows) was observed. Upon fine-structural analysis (D), these neurons showed severe signs of degeneration, including dark cytoplasm and condensed chromatin. However, they were found to still have intact input synapses (E; see asterisks). Figure published in Magno et al. (2011)[1]. Reprinted with permission from the European Journal of Neuroscience: John Wiley and Sons, The integrity of cholinergic basal forebrain neurons depends on expression of Nkx2-1.* ***Magno et al. (2011)*** *Copyright © 2011.*

Taken together, from these observations, it can be concluded that a) Nkx2-1 is critical for the development and survival of these neurons, and b) *NGF* is essential for the maintenance and differentiation of cholinergic neurons downstream of Nkx2-1, as already discussed in Section 4.1.2.

4.2 Clinical update on *NKX2-1* haploinsufficiency in humans and comparison with our findings in mice

4.2.1 The challenge of finding patients and identifying means to compare them

Twenty-one years after the first patient with a *NKX2-1* variant-associated brain-lung-thyroid syndrome was reported[190], this study provides what is so far the most comprehensive overview of the syndrome with current statistics on organ phenotypes, genotype-phenotype correlation,

and the first statistical analyses including survival. The main methodological difference from previous studies[50,54,56,60] is that, due to the rare nature of the syndrome, all cases with proven *NKX2-1* variants with at least one organ phenotype (brain, lung or thyroid) that could be found in the literature were included.

In the process of case finding and data extraction some major challenges were encountered. (1) Not all cases could be found in the PubMed database, despite the use of all common synonyms of *NKX2-1* and the disorder (see Table 7). While searching gene databases for information on genotypes and reading case reports from reference lists (including conference papers) additional cases were identified. (2) The level of information provided by the authors of case reports and series was highly heterogeneous (see also Section 3.5.2). For some patients, detailed descriptions of the patient history were available, whereas other reports lacked basic information on patient characteristics such as gender and age (see Table 8), and the respective cases hence had to be excluded from certain subanalyses. (3) Moreover, the descriptions of the variants differed depending on the type of publication, the journal requirements, and the year of publication, as well as the *NKX2-1* reference sequence used (sometimes not provided).

For better comparability, the nomenclature was standardized as best as possible (see Tables 25–28) according to current HGVS recommendations to be able to identify and group patients with the same variant for genotype-phenotype analysis. For this purpose, the most frequently used reference sequence isoform 1 (NM_001079668.2; see Figure 1 for gene structure) with the location of the HD at c.571-750[18] was chosen, and different genome databanks were subsequently searched (see Section 2.4.2.4). In several instances, patients were reported twice (in a primary report and with additional information in a follow-up study; see Sections 3.5.1 and 3.5.2). Identification of these cases was sometimes difficult, as the previous publication was often only mentioned as a side note (e.g. in a figure), and the information had to be merged. Moreover, in one journal, variants with coordinates not matching those of the *NKX2-1* gene were reported[203] as *NKX2-1* variants. The search of genome databases revealed that the authors characterized variants located on chromosome 9 and not 14, and those variants are also named *"NKX2.1"* as a synonym. The respective 25 patients[203] were therefore excluded from the present study.

4.2.2 Patient demographics

In total, 351 patients with 167 variants of the *NKX2-1* gene were identified from 91 case reports published between 1998 and May 2nd 2020. The patients were mainly Caucasians from industrialized countries (mostly Europe and North America; see Table 9), which is most likely due to the availability of genetic testing in these countries and does not account for a pattern of

geographical distribution. There were indeed slightly more females (56%), as previously reported for a smaller patient sample[61]. The vast majority of patients were children and teenagers, with a median age of 11.5 years (IQR 4–28 years) at last visit. This reflects the early onset of the disorder and rather short follow-up periods before publication. It is understandable that authors decide to publish early because of the rare nature of the disorder. In the last few years, there has fortunately been a rising number of follow-up studies offering insight into the course of the disease. The longest reported follow-up period of a patient series was 24.5 years[61].

4.2.3 Mode of inheritance and phenotypic variability

The mode of inheritance in this disorder is autosomal dominant[50,54,56,61], and the frequency of familial variants was shown to be higher than previously reported (73% in this study versus 62%[61] and 37%[62]). Earlier, Carré et al.[60] found more *de novo* variants (54%) in their smaller case study (versus 27% in the present study). Here, it should be noted that the rate of familial variants might even be higher due to possibly undiagnosed family members who are mildly symptomatic, are not yet symptomatic, or died before genetic testing could be performed. A subanalysis for the types of variants (see Figure 19b) revealed that leading subgroups were WGD in *de novo* variants (28%) and missense variants in familial cases (39%). Moreover, the highest frequency of *de novo* variants was detected in WGD, where they were found in equal abundance to familial variants (each 48%; see Tables 23, 25 and 30).

The present study (see Tables 25–28), and some previous ones, found a great phenotypic variability – that is, a wide range of manifestations occurring in different (related or unrelated) patients with the same variant – indicating variable expressivity[69,194], which is not unusual for autosomal dominant disorders[192]. There are various possible explanations for these observations. For instance, Maquet et al. propounded the intrafamilial variability in their study to be related to different individual genetic backgrounds that can influence the phenotypes. They suggested that the individual genetic background might be capable of compensating the *NKX2-1* deficiency to different degrees in various organ tissues[17], and they also proposed a possible impact of a variability in the wild-type allele[17]. Other authors have argued that intrafamilial variability cannot be fully explained by haploinsufficiency and that a possible role of other modifying genes must be considered[14,34,60,188,238]. In general, variable expressivity can be caused by known or unknown genetic factors (epistasis), environmental factors, coincidences, and other factors – including modifier genes and allelic variation[280]. As the concept of expressivity cannot be applied to phenotypic variability of patients with different variants, other causes must be identified. One possibility discussed by some authors is a partial dependency on the presence

of an intact HD[218]. In fact, the results of this study provide evidence for a role of not only the HD but also other domains in the disorder (see below).

4.2.4 Involvement of functional domains and impairments of gene functions

An analysis of the genotypes (cf. Section 3.5.4) revealed that the overwhelming majority of patients (95%) either had WGD (18%) or variants involving at least one functional domain (94% of patients with non-WGD variants; see Table 10 and Figure 18 for details), pointing to the critical role of the functional domains (see Section 1.1 and Figure 1 for details on the functional domains) in this disorder. However, the fact that there are also disease-causing variants without impairment of the functional domains or sites (see Table 28 for the respective variants) further demonstrates the vulnerability of the *NKX2-1* gene to even minor alterations. As knowledge on the *NKX2-1* gene is generally scarce, the possibility cannot be excluded that some of the variants categorized here as not being domain-involving (see Tables 10 and 28) might actually encompass some – *yet not identified* – functional domains or sites. Among all domains, the HD and the C-TAD were the ones affected in the majority of variants (both 67%), followed by the NK2-SD (58%) and the N-TAD (23%). In addition, three out of seven phosphorylation sites, located in either the NK2-SD (c.850_852;p.284) or the C-TAD (c.1069_1071;p.357 and c.1096_1098;p.366) were regularly involved. The degree to which the domain involvement varies among disease-causing variants highlights that the essential functions of the *NKX2-1* gene are distributed among different gene regions rather than being concentrated in one single domain.

In accordance with a previous observation from a smaller study[62], exon 3 is indeed the "hot spot" of variants. This is not surprising given that it is the largest exon (see Figure 1) and contains four functional domains (HD, Domain I, NK2-SD, and C-TAD) and three phosphorylation sites (c.850_852;p.284, c.1069_1071;p.357, and c.1096_1098;p.366). Overall, 66% of non-WGD variants are located on exon 3, and the others are found on exon 2 (24%) and intron 2 (10%). Interestingly, no variants of exon 1 or intron 1 and hence of the TN and the other four phosphorylation sites (c.100_102;p.34, c.124_126;p.42, c.142_144;p.48, and c.157_159;p.53) in the TN or adjacent to it were encountered (see Figure 18 for all non-WGD). Moreover, there were also no variants involving Domain I. While the TN is discussed as a potential promoter-specific transcriptional activator or inhibitor[23], the functional significance of Domain I is yet unknown[23]. This observation could either indicate that variants involving the TN and Domain I do not cause disease or simply that they are rare and/or not (yet) reported.

On the whole, 53 variants affecting 116 patients were studied for functional impairments (see Section 3.5.4.5.3). It should be noted here that the tests performed in these patients were highly

heterogenous; the findings are thus not representative of the patient population and were therefore not further statistically analyzed. In summary, the most important result of the subanalysis of functional impairments was the finding that all variants with impairments in gene function (see Table 11) did at least involve one functional domain, which was most often the HD (85%), followed by the C-TAD (70%). As can be seen from Table 11, HD-only, four-domain, and three-domain variants were nearly equally often found to cause functional impairments (28%, 20%, and 30% respectively). In contrast, two-domain, C-TAD-only, and N-TAD-only variants were rare. Interestingly, while no NK2-SD-only variants were studied, the NK2-SD was involved in the majority of variants (63%), namely, in those with the involvement of three or four domains.

Among all impairments, protein truncation was the most frequent one (48%, 21/53 variants). Its high frequency is in line with knowledge about disorders with haploinsufficiency, where the loss of one gene copy cannot be tolerated[283]. Variants causing protein truncation are generally expected to have major effects on gene function[281], as they impair transcription and result in a shortened or absent protein[282]. In line with this, protein truncation was found to primarily result from variants disrupting at least three functional domains (in 90% the HD combined with the NK2-SD and the C-TAD; see Figure 21). Moreover, half of those variants had an additional impairment of the N-TAD (53%), and interestingly, in a small subset (10%) C-TAD-only variants were also found to result in protein truncation. Taken together, the C-TAD was always involved in protein-truncating variants. Similarly, its impairment was also found in all variants with impaired protein synthesis (100%; see Figure 22). This observation is plausible, as the C-TAD is a transactivation domain and thus important for gene function. Transactivation domains generally promote transcription by recruiting other proteins (e.g. coactivator proteins and general transcription factors) without necessarily directly binding to DNA. However, the other transactivation domain, namely, the N-TAD, was less involved, suggesting that it is, in contrast to the C-TAD, not essential for proper protein synthesis. In conclusion, the findings concerning protein truncation highlight that the involvement of multiple domains causes more fundamental impairments of *NKX2-1* than single-domain variants and also points to a critical role of the C-TAD.

Evidence supporting this theory was provided by Moya et al. (2018)[20], who performed in vitro experiments to compare two *NKX2-1* variants with each other, namely, the c.313dup; p.(Val105Glyfs *334)[20] variant (Patient 212; Table 27, Part 1), which is located in the N-TAD and impairs four domains, and the c.915del;p.(Ala306Argfs*75)[230] variant (Patients 318-321; Table 28, Part 1), which is not located in any domain and solely impairs the C-TAD. First, they

tested the variants' binding activities on three specific promoters that are relevant for the brain-lung-thyroid syndrome: *LHX6* from the basal ganglia, SP-B from the lung, and TG from the thyroid. They discovered no binding activity in either of the variants compared to a wild-type protein with normal activity[20]. The authors further tested whether the transcriptional co-activator *TAZ,* which is co-expressed with Nkx2-1 in the mouse lung and interacts with it to activate SP-C[284], could rescue the transcriptional activity in these variants[168,216,229]. Indeed, *TAZ* co-expression was found to partially rescue the protein activation of the TG and *LHX6* promoters and to completely recover the SP-B promoter activation in the C-TAD-only variant c.915del; p.(Ala306Argfs*75)[230], while it had no effect on the four domain variant c.313dup; p.(Val105Glyfs *334)[20]. The latter finding is explained by the fact that *TAZ* binds to the transactivation domain 1[284] that contains the N-TAD[29] (see Figure 3 for details on the mutation in mice of the transactivation domain 1) that is, in the variant with an intact transactivation domain 1, the activity is rescued, while it remains unchanged in the one with its impairment.

It has been assumed that a certain amino acid sequence at position c.97-100 adjacent to or involving the phosphorylation site c.100_102;p.34 might be responsible for interaction between the N-TAD and *TAZ*[284]. However, there were contradicting studies that found this region to be inessential[20,285]. Moya et al. (2018) further suggested that *TAZ* could also be a genetic modifier of the lung phenotype in humans[20], as it is expressed in the mouse lung[284]. This assumption remains to be elucidated.

Furthermore, the finding that the NK2-SD is often involved in variants with functional impairments (63%) is in line with the knowledge that it plays a role in protein-protein interactions[22,23,25,28]. In addition, the observation that this involvement is always in conjunction with other domains (at least in the available data; see Table 11 and Figures 21–22) is in keeping with its characterization as an accessory DNA-binding domain without involvement in the specificity or affinity of DNA binding[25,28]. The finding that, apart from an intact transactivation domain 1, an intact transactivation domain 2, which encompasses the NK2-SD (see Figure 3 for details on the deletion of transactivation domain 2 in mice), is required for full functional interaction with *TAZ* in mice[284] is also in agreement with this.

In accordance with the functions of the HD in DNA recognition and nuclear localization of the protein[15,16], the HD had the highest overall involvement in variants with impaired DNA binding (79%), binding to promoters (71%), and migration to the nucleus (75%). Overall, in a subset of different types of variants, the following were described: a) loss of function (23%), which describes a state in which the complete protein is lost, and b) dominant negative effects (17%),

where the resulting protein interferes with the function of the wild-type protein. Both types of impairments can generally be found in variants of haploinsufficient genes.

Furthermore, some interesting findings were revealed regarding interactions with *PAX8* – another highly conserved transcription factor that is co-expressed with Nkx2-1 in the thyroid gland and is essential for its proper development in mice[286,287]. As in Nkx2-1 null mice, homozygous deletion of Pax8 also led to thyroid agenesis[288]. In some of the variants, which were tested for the interactions of the *NKX2-1* variants with *PAX8*, their synergy on the TG promoter was found to be impaired[16,62,182]. Among those variants, the presence of an intact *PAX8* gene could only partially rescue the *NKX2-1* activity on the TG promoter in the C-TAD-only variant c.915del; p.(Ala306Argfs*75)[230], but not in the c.313dup; p.(Val105Glyfs *334)[20] variant involving four domains. This could stem from more fundamental impairments caused by multiple gene involvements. However, this, along with many over observations of the functional impairments, warrants further investigation.

4.2.5 Is there a genotype-phenotype correlation?

Among the patients a wide range of *NKX2-1* variants, from WGD to point variants, were found. So far, it has been unclear whether a genotype-phenotype correlation exists. Some authors have postulated various potential correlations between deletion size, location and types of variants, and disease severity, while others have concluded that there are none[26,61,210]. The present study can, for the first time, shed some light on this question (see below).

4.2.5.1 Disease severity

4.2.5.1.1 A novel disease severity score for patients with *NKX2-1* variants and how it can help identify genotype-phenotype correlations

Given that examining possible differences in disease severity is difficult without clear definitions of mild and severe phenotypes, a clinical disease severity score (see Table 21) was developed based on the review findings for the purpose of this study, and it was used to compare various patient subgroups. Similar scores have already been used in other rare genetic diseases with multi-organ manifestations, such as Morbus Fabry, to assess disease severity, manage treatment, and study long-term outcomes of patients[289-291]. This score is a first attempt at establishing a disease severity classification that can a) be a useful tool to help create more uniformity and improve the comparability of case reports, thereby aiding in phenotypic characterization and genotype-phenotype correlation studies; and b) possibly be used as a checklist in the clinical setting, allowing clinicians to easily know what possible manifestations to look for during assessment of these patients. This checklist warrants further evaluation and optimization in the future as new cases are reported and specific studies are conducted.

4.2.5.1.2 Disease severity based on different genotypic groups, the involvement of functional domains, and the gene regions c.96–565 and c.753–881

A comparison of the disease severity scores between the genotypic groups based on the impaired gene region (WGD, HD, non-HD+ and non-HD; see Table 24) revealed a) significantly lower overall, brain and thyroid scores for the HD group compared to the WGD and non-HD+ groups as well as b) lower thyroid scores of the HD group compared to non-HD variants. This finding is partially surprising. While it is plausible that WGD and non-HD+ variants that impair larger regions of the *NKX2-1* gene have higher scores, the finding of higher scores in non-HD variants, which do not involve the HD, is surprising at first glance. However, a closer examination of the group compositions (see Table 23) and a comparison with the analysis of the genotypic groups based on the types of variants (WGD, missense, and AOV; see Table 30) might provide some explanations.

Similar to the HD group, missense variants were found to have lower scores compared to WGD and AOV for the overall, brain, and thyroid scores (see Table 31). For a better understanding of the present results, it is important to notice the link between missense variants and HD variants. As indicated in Table 30 and Figure 27a, missense variants either involve single domains (98%) or none (2%), and among all variants, the largest subgroup are HD variants (86%). In turn, 74% of the variants of the HD group are missense variants (see Table 23). These connections explain why the results of the HD and the missense group are analogous and also demonstrate the vulnerability of the HD towards missense variants in contrast to other domains that are rarely affected by them (see Table 30 and Figure 27).

Furthermore, a main finding from the subanalysis of disease severity scores in variants based on the number and combinations of involved functional domains (see Table 33) was that HD-only variants have significantly lower disease severity scores compared to a) variants with additionally impaired domains (significant for the overall score compared to three and four domains, and thyroid scores compared to two, three and four domains) and b) the isolated domain variants C-TAD-only and N-TAD-only (both significant for the thyroid score). The higher scores in multi-domain variants compared to HD-only variants seem plausible at first sight, as variants causing impairments of larger gene regions are presumably more disruptive and therefore result in more severe phenotypes. This is also supported by the finding that there were no differences for any of the scores between WGD and variants that impair four variants (see Table 33). However, this was not seen in the subanalysis of other single-domain variants, and the comparison of multi-domain variants with one another, as well as the comparison between any

variants with domain involvement and those without known domain involvement, led to inconclusive and partially conflicting results (see Table 33 for details).

Another striking finding pertained to the significantly higher scores in the C-TAD-only and N-TAD-only groups compared to HD-only variants. One possible explanation for this could be that HD-only variants cause less severe impairments of the *NKX2-1* gene. Yet, when considered with the above-mentioned findings, another underlying cause for the discrepancies seems more likely. As for the subanalysis of the genotypic groups based on the impaired gene region and type of variant, great differences in the share of missense variants appear to play a major role, as HD-only variants are exclusively caused by missense variants compared to other single-domain variants that are less (33% in C-TAD-only and 40% in N-TAD-only) or never (all multi-domain variants; see Table 30) found to be missense variants (see Figure 27a). The latter finding is plausible, as missense variants are point mutations that only affect the domain in which they are located.

In addition, there were no differences between the scores of the gene regions c.96–565 (transactivation domain 1) and c.753–881 (transactivation domain 2; see Table 35). This result can also be best explained by the share of missense variants, which is similar in both groups (96% for the gene region c.96-565 and 89% for the gene region c.753-881; see Table 34).

The overall finding of milder phenotypes in missense variants (see Table 31) using a disease severity score and statistical analysis in the present study is in line with previous smaller case series that described milder manifestations in patients with missense variants[52,54,63,68]. The data suggest that disease severity does not depend on the location of the variant on the *NKX2-1* gene, nor does it depend on the impaired functional domains or the number or combination thereof. They rather point to a major role of the type of variant.

It was further noticeable that there were no differences in the lung scores in any of the subgroup analyses of the genotypic groups based on the impaired gene region (Tables 24), types of variants (Table 31), and the gene regions c.96–565 and c.753–881 (Table 35). Interestingly, subgroup analysis for domain involvements revealed significantly lower scores in variants without domain involvement compared to WGD, multi-domain variants, and all single-domain variants, except for N-TAD-only and NK2-SD-only (Table 33). The N-TAD-only variants had higher median scores, and the isolated NK2-SD variant group had only one patient with a high lung score (see Table 33). However, the differences compared to no-domain variants, which had lung involvement in only 13% (see Section 3.5.6.4.1 for details on no-domain variants), were

not statistically significant. This might be due to the smaller sample sizes of both the N-TAD-only and NK2-SD-only groups (see Table 10). In addition, the overall scores of variants with no domain involvement were found to be significantly lower compared to variants with four- and three-domain involvement. These results might be influenced by the lower lung scores, which affect the overall scores. In conclusion, the results of the no domain group suggest that the lung is more vulnerable to impairments of the functional domains than those not involving the domains.

4.2.5.1.3 Disease severity in WGD with additional deletion of other genes

WGD with concomitant proximal (chromosomal coordinates hg19 chr14: 30631931 to 36.985.601) and distal (chromosomal coordinates hg19: chr14:36.990.3544 to 50031170) deletion of other genes (see Section 3.5.4.4. for details) were found to result in significantly higher overall disease severity scores compared to those with additional deletion of only the distal regions (see Table 36). This confirms some authors' previous observation of more severe phenotypes in patients with WGD with additional proximal deletions (previously defined as chromosomal coordinates hg19: 28.000.000–35.000.000), compared to WGD with additional distal deletion (defined as chromosomal coordinates hg19: 35.000.001–40.000.0000)[67,184]. The finding is also in line with studies that found severe phenotypes in patients with large WGD[40,41,46,51,224,271]. The more severe phenotypes in patients with contiguous gene deletion syndromes might be due to the loss of genes that possibly regulate *NKX2-1* or are otherwise able to at least partially compensate for *NKX2-1*-deficiency in smaller deletions.

4.2.5.1.4 Disease severity subanalysis for the mode of inheritance and gender

In line with the knowledge that *de novo* variants are often more disruptive and hence associated with severe early onset genetic disorders[292], *de novo* variants in this study also had significantly higher overall and organ-specific disease severity scores (see Table 22) and were more often found in variants with higher disease severity scores in other subanalysis (i.e. WGD and AOV [Tables 30 and 31] as well as non-HD+ and non-HD variants [Tables 23 and 24]).

In line with observations described in Section 4.2.5.1.3, the lower disease severity scores in familial variants are most likely explained by the higher frequency of missense variants (missense versus *de novo* variants: 39% versus 24%) and the smaller share of WGD (missense versus *de novo* variants: 10% versus 28%; see Figure 19b for details). In turn, the majority of missense variants (82%; see Table 30) and consequently also most of the HD variants (88%; see Table 23) were found to be familial variants. In addition, using ANOVA, familial variants were found to be associated with a) missense variants and AOV when compared to WGD (Table

32) and b) HD variants compared to WGD, non-HD+, and non-HD (Table 29). The one-way ANOVA subanalysis also showed an association between WGD and *de novo* variants when compared to missense and AOV (Table 32).

Interestingly, males had significantly higher overall and brain severity scores (see Table 22). As there are no known or plausible reasons why males should be more severely affected by the disorder, again the findings are most likely due to the higher frequency of AOV in males (males versus females: 52% versus 46%) and missense variants in females (females versus males: 35% versus 32%; see Figure 19a).

4.2.5.1.5 Disease severity in affected members of multigenerational families

Some authors have previously described a tendency of children of affected parents and grandparents to have more severe impairments with an earlier onset of symptoms[17,185,192]. A detailed analysis of multigenerational families with *NKX2-1* variants in this study revealed a disease progression (i.e. an increase in the overall disease severity score) among generations in half of the families (53%; see Table 38); therefore, no clear conclusions can be drawn from the analysis. While the increase in disease severity in these families could be incidental, it might also be due to an accumulation of the mutation load over generations. In addition, in two cases, parental mosaicism (i.e. the presence of two or more genetically different sets of cells that could be inherited by the parents) was suspected to play a role in this context[14].

4.2.5.2 Phenotypes

4.2.5.2.1 Combinations of organ manifestations

Comparable to Carré et al. (2009)[60] and Gras et al. (2012)[61], the triad was the leading organ combination (39% in this study versus 50% in Carré et al.[60] and 36% in Gras et al.[61]), followed by the brain-thyroid phenotype (28% in this study versus 30% in Carré et al.[60] and 32% in Gras et al.[61]), the brain-only phenotype (14% in this study versus 13% in Carré et al.[60] and 21% in Gras et al.[61]), a brain/lung combination (6% in this study versus 4% in Carré et al.[60] and 11% in Gras et al.[61]), and the thyroid-only phenotype (5% in this study versus 2% in Carré et al.[60] and 0% in Gras et al.[61]). The main difference between the present study and both Carré et al.[60] and Gras et al.[61] is the detection and reporting of isolated lung phenotypes (5% in this study), which contribute to some differences in the frequencies of organ combinations.

The triad was significantly more often found in patients with WGD and AOV compared to missense variants (see Table 32), and it was associated with WGD when compared to HD variants (see Table 29). In contrast, an association of single-organ phenotypes with missense variants was detectable (Table 32). In the subanalysis of genotypic groups based on the impaired

gene region, the isolated lung phenotype occurred more in HD variants and the isolated thyroid phenotype more in non-HD variants (Table 29), while no differences were found for the isolated brain phenotype (Tables 29 and 32). In contrast, no major differences were detectable for two-organ phenotypes (Tables 29 and 32). These findings confirm that the fundamental impairments of gene function found in WGD and AOV are more likely to cause dysfunctions of all three organs in contrast to point variants leading to single-organ impairments.

4.2.5.2.2 The brain phenotype – the most frequent manifestation

Overall, the brain phenotype remains the leading cause of morbidity in patients with *NKX2-1* variants, with an overall frequency of 88%. The frequency is slightly lower than previously reported by smaller studies (93% in Carré et al.[60] and 97% in Gras et al.[61]). This is most likely explained by the present study's inclusion of patients with isolated lung and thyroid phenotypes from before and after 2009 (see Figure 23 for the phenotypes found by Carré et al.[60] in comparison to this study). The median age at onset of the brain phenotype was found to be 1.1 (IQR 0.5 – 2.5) years in the present study, which is earlier than the previously reported mean age at onset of chorea of 2 years[61]. This finding is reasonable, since the time of diagnosis of any neurological symptom, including muscular hypotonia in neonates or developmental delay, was counted as the onset of the brain phenotype in this study. As a subset of patients without neurological symptoms were reported or died before the age of 4 years (41%), it is possible that the incidence of the neurological phenotype is even higher. At the same time, there might also be an overall recruitment bias of primarily neurological clinics that could contribute to the high incidence of the brain phenotype[63]. Conversely, it is also conceivable that the brain has a lower tolerance to *NKX2-1* deficiency, as suggested by Thorwarth et al.[26]. The present study confirmed that the leading neurological symptom is chorea (73% of the neurological manifestations) with or without athetosis. In line with the findings in mice[1,2] (see Figure 29), in one patient with a brain autopsy (Patient 263; see Table 27, Part 2) loss of ChAT-positive and PV-containing striatal interneurons and efferent fibers, as well as atrophy and abnormal morphology of ChAT-positive striatal interneurons, was found[65] (see Table 13 for details).

The term "benign hereditary chorea," which is often used as a synonym for the brain-lung-thyroid syndrome, might be misleading in view of the possibility and frequency of multiple-organ manifestations. Furthermore, the term "benign" should be reserved for cases of isolated chorea and not used for other patients who have a wide spectrum of manifestations, ranging from tremor to severe disabling disease. As already suggested by other authors[61], this term should be reconsidered, although the present study showed a tendency towards an improvement

in motor symptoms (see Section 3.5.5.2 for details), particularly regarding chorea in adolescence and adulthood[50,61,65]. This improvement could be due to compensational mechanisms of the brain and/or axonal sprouting and reestablishment of neuronal networks. Among all manifestations, the brain appears to be particularly vulnerable to WGD, since all patients, except one 1.3-year-old infant with a lung-thyroid phenotype[67] (Patient 35; see Table 25, Part 1), had neurological impairments (see Table 25). In view of the range of neurological manifestations with early or late onset (see Figure 17), it is possible that this infant could also develop a brain phenotype in the future.

The theory of a particular vulnerability of the brain towards WGD is supported by the results of the ANOVA analysis. It revealed an association between WGD and the overall brain phenotype, along with typical manifestations, such as choreoathetosis and muscular hypotonia when compared to missense (see Tables 30 and 32) and HD variants (see Tables 23 and 29). Moreover, overall cognitive impairments were found to be associated with WGD compared to AOV and missense variants (see Tables 30 and 32) as well as HD and non-HD variants (see Tables 23 and 29). The subanalysis of cognitive impairments further revealed a significantly higher frequency of WGD in patients suffering from intellectual disability and microcephaly compared to AOV and missense variants (see Tables 30 and 32) and HD, non-HD+, and non-HD variants (see Tables 23 and 29). Only chorea without athetosis was found to be associated with missense and HD variants compared to any other group (see Tables 23, 29, 30 and 32).

While there were generally only a few differences between WGD and non-HD+ variants, the comparison of non-HD+ variants with both HD and non-HD variants led to results similar to those for the comparison of WGD with HD variants, namely, higher frequencies in non-HD+ variants (see Tables 23 and 29). This most likely stems from the fact that non-HD+ variants are either three-domain (39%) or four domain variant (61%), thus impairing large parts of the *NKX2-1* gene. This finding is in line with the observation of a high frequency of the brain phenotype in WGD and the similarities in the brain severity scores of WGD and variants with three and four domains (see Table 33). In summary, these results indicate a particular vulnerability of the brain to variants with a complete lack of *NKX2-1* and multi-domain impairment. The importance of the functional domains for the brain phenotype, at least of the HD and NK2-SD, is further supported by the knowledge that the ventral nervous system defective gene, an essential factor of the embryonic central nervous system development, contains the HD and the NK2-SD[28].

A striking aspect of the brain phenotype to date has been that – although the findings in humans as well as in mice (cf. a recent study[3]) point to a major impairment of the cholinergic system and in particular the SHS, which is essential for memory and learning[261] – cognitive impairments were initially either not often reported in patients with *NKX2-1* variants, or not regarded as a main feature. In particular, memory function has not been well studied; it was only assessed in a few patients, and a deficit in spatial memory was found in one of them (Patient 285; Table 27, Part 3)[195]. The reporting on cognitive impairments has changed in recent years, especially as longer follow-ups were published describing patients with learning difficulties and low IQs, but the general perception of the brain phenotype is still that of a primary motor disorder. However, the present study suggests that cognitive impairments are in fact an integral part of the syndrome, with a frequency of 29% (see Table 12). It is furthermore possible that underreporting occured in the initial years because cognitive impairments might not always have been actively searched for, as suggested by earlier reports not providing information on this possible aspect. Another possibility is that the patients were not yet symptomatic at the time of reporting, which is supported by the differences between some initial reports and follow-up studies, as well as the analysis of the frequency of cognitive impairments in different age groups. While mental retardation and speech delay are mostly diagnosed in early childhood, learning difficulties, including ADHD, are more often found and/or recognized in school children and young adults.

However, some possible influential factors should be considered when assessing cognitive impairments in these patients. Motor impairments, such as difficulties with handwriting due to chorea, can lead to social embarrassment[247], lack of confidence, and poor school performance, as well as compromised IQ tests. Therefore, motor dysfunctions must be taken into account when screening for cognitive impairments (e.g. by comparing the results to patients with similar disorders without *NKX2-1* variants). Psychiatric disorders do not seem to play a major role in this disorder, as they were only found in 3% of patients. This finding is interesting in view of current research discussing *NKX2-1* as a candidate gene for schizophrenia[35] and other disorders such epilepsy[26]. While schizophrenia was only found in one patient, seizures were reported in seven patients (see Table 12).

4.2.5.2.3 Evidence for the comparability of the human brain phenotype with our data in mice – postmortem studies, neuroimaging, and pharmacological treatments

Although our mouse experiments cannot be regarded as a direct model for the brain phenotype in humans, they can help us understand the syndrome. Evidence for the comparability of our findings in mice with those in humans is provided by the identification of *NKX2-1* in brain

autopsies of two adult men[1]. Analysis of their basal ganglia revealed, just as in mice, strong *NKX2-1* expression in cholinergic and PV-containing GABAergic neurons of the external part of globus pallidus. *NKX2-1* was also found in cholinergic neurons, mainly in the putamen. Brain autopsies of patients with *NKX2-1* variants have so far only been performed in three patients with a brain phenotype, and pathologies were described in one of them (Table 13), namely, a 59-year-old patient (Patient 263; Table 27, Part 2) with choreoathetosis, clumsiness, and developmental delay, who died from leukemia. In this patient, a significant reduction in striatal cholinergic and PV-containing interneurons and their efferent fibers was detected[36,65]. A similar reduction in these neuronal profiles was also observed in postmortem studies of patients with Tourette's syndrome, which is another not well-understood neurodegenerative disorder involving the basal ganglia[293]. In Tourette's patients, *LHX6* was additionally found to be reduced[294].

To date, neuroimaging has generally not been considered a useful tool and is therefore only performed in a minority of patients (44%; see Figure 24). After most of the conventional MRIs in the first few years were normal, as reported by Carré et al.[60] in 2009 (imaging performed in 85% and pathologies found in only 18%), they were less often performed in subsequent years. While most of the scans were normal, some interesting findings were observed (see Figure 25) concerning the basal ganglia (39%) and the cerebral cortex (34%) among those with pathological results (27%, see Figure 25). In particular, the findings of Krude et al. (2002)[46], who reported hypoplastic pallida without proper division into the lateral and medial part and reduced striatal volume in a 15-year-old patient (missense variant, triad, severe choreoathetosis; Patient 139; Table 26, Part 2), are consistent with the observations made in Nkx2-1 null mutant mice that failed to develop pallida and entirely lacked *trkA*-positive BF cholinergic neurons[43]. Another family investigation with conventional MRI showed cortical atrophy and hyperintensities in the pallida of a 59-year-old woman (nonsense variant, mild chorea, learning difficulties; Patient 190; Table 26, Part 2) that were not present in her 30-year-old son (Patient 194; Table 26, Part 2) suffering from chorea and stuttering and who could not finish high school due to learning difficulties[221].

Moreover, using advanced MRI techniques, Maccabelli et al. (2010)[256] identified bilaterally reduced striatal volumes and spectroscopic abnormalities such as a reduced choline/creatine ratio (an imprecise indicator of hypomyelination) in two patients (whole gene deletion, brain-thyroid phenotype; Patients 44/45; Table 25, Part 2)[69,188,208,256] who previously (2002) had no abnormalities in conventional MRI scans[188]. This finding is in line with the significantly reduced weight and volume of the globus pallidus found in our prenatal mutants[1]. This example,

and the following ones, illustrate how advanced imaging techniques can help identify nonmacroscopic pathologies that are not visible on conventional scans. As these techniques were only used in the minority of cases, we do not know whether there could have been similar findings in the many other cases where patients with severe phenotypes had normal conventional scans. For instance, in a 19-year-old girl with the full triad and an intelligence quotient (IQ) of 58 (nonsense variant, normal MRI; Patient 208; Table 27, Part 1)[55], we would expect to find some pathologies in brain scans.

However, some promising imaging studies have been carried out in recent years. For instance, in 2010, Salvatore et al. studied a previously reported three-generation family[194] with a nonsense variant and ventricular dilation and empty sellae on conventional MRI scans (Patients 283–285; Table 27, Part 3). Using Fluordeoxyglucose positron emission tomography (FDG-PET), they detected hypometabolism in the basal ganglia that was most prominent in the nucleus caudatus and the temporo-parietal cortices of the adult patients (see Figure 32). While the 56-year-old grandfather (Patient 285; Table 27; Part 3) had delayed motor development in childhood, hyperkinesis, short-term verbal and spatial memory, as well as ADHD and hypothyroidism, his 26-year-old daughter (Patient 284; Table 27, Part 3) was suffering from progressive generalized chorea, long-term verbal memory deficit, and post-partum psychosis. Her 5-year-old son (Patient 283; Table 27, Part 3) displayed the full triad with psychomotor delay and a low IQ (76)[195].

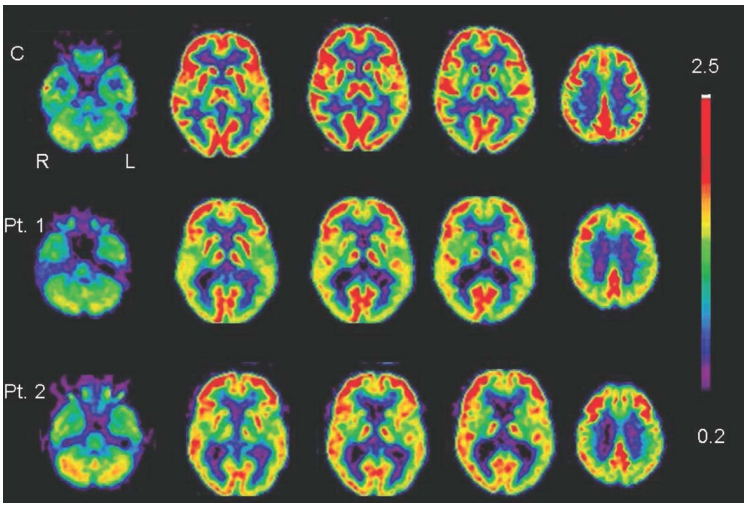

Figure 32. Fluorodeoxyglucose positron emission tomography (FDG-PET) showing hypometabolism in the basal ganglia of patients with the c.524C>A;p.(Ser175*) variant. *FDG-PET scans of a control (C, 39-year-old control) and two NKX2-1 haploinsufficient patients (Pt 1 and 2) are shown. In both patients, hypometabolism in the caudate nuclei and the medial frontal and temporo-parietal cortices was observed. High FDG uptake is indicated in red and low uptake in blue. R (right), L (left). Figure taken from Salvatore et al. (2010)[195]. Reprinted with permission from Movement Disorders: John Wiley and Sons, Benign hereditary chorea: Clinical and neuroimaging features in an Italian family,* ***Salvatore et al. (2010).*** *© Copyright 2010.*

Although this is a rare observation, it is plausible to find memory deficits in patients with *NKX2-1* variants in view of the data in mice. For instance, most recently, Magno et al. (2017)[18] have shown that Nkx2-1 is essential for spatial memory in mice[18]. Moreover, spatial memory deficits were also reported for Lhx7-deficient mice[30,104]. In view of the findings in mice, detecting similar changes, particularly involving *LHX7*, Isl1, and Ldb-1, in *NKX2-1* haploinsufficient patients is presumable (currently no data available).

In another study, Mahajnah et al. (2007) found a decreased uptake of technetium 99m ethyl cysteinate dimer (EDC) in the striatum of two siblings (11-year-old boy and 15-year-old girl; Patients 327 and 328 respectively; Table 28, Part 2) with a non-WGD deletion and chorea, who previously had normal conventional MRI scans[231]. These discrepancies suggest that normal conventional MRI scans on patients with *NKX2-1* variants do not necessarily prove that there are no pathologies, and they point to the utility of advanced MRI techniques and nuclear scans for detecting functional abnormalities. This is further supported by the findings of Uetmatsu et al. (2012), who investigated three children with choreoathetosis and a missense variant (two cousins aged 5 and 6 years, one with a brain-thyroid phenotype and one with a triad; Patients 137 and 138 respectively; Table 26, Part 2), and normal conventional MRIs. Using 99-technetium EDC single photon emission computed tomography (EDC-SPECT), the authors found a reduced cerebral perfusion in the CPu (see Figure 33) in both patients comparable to the reduction in Huntington's disease[238]. Based on their research, Uetmatsu et al. suggested that the

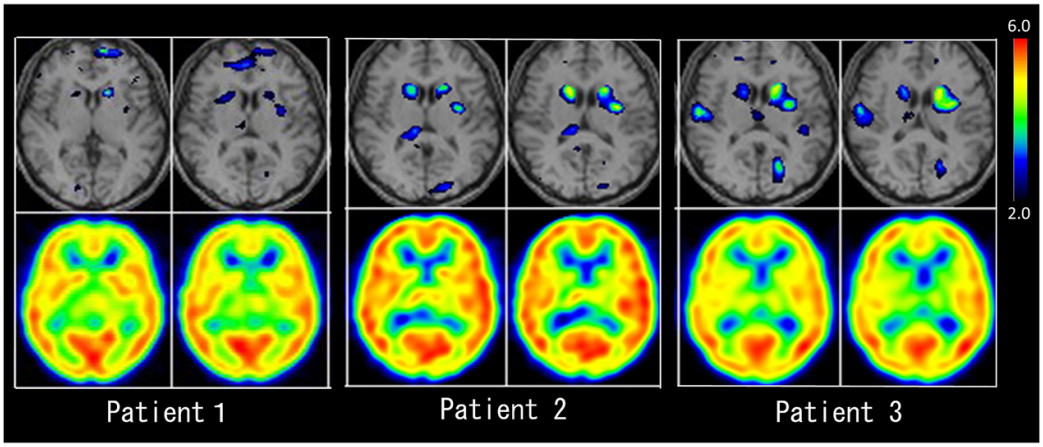

Figure 33. Ethyl cysteinate dimer single photon emission computed tomography (EDC-SPECT) showing hypoperfusion in the basal ganglia of patients with the c.978_1056del;p.(Ala330Serfs*25) variant*. The lower figures depict the 99-technetium EDC-SPECT for each patient, and the upper figures display the results of their statistical analysis using the easy Z-score Imaging System (eZIS). The blue color indicates the regions with reduced cerebral blood flow. Hypoperfusion in the caudate nuclei is found in all of the patients. Figure taken from Uematsu et al. (2012)[238]. Reprinted with permission from the Journal of Neurological Sciences: Elsevier, Hypoperfusion in caudate nuclei in patients with brain-lung-thyroid syndrome,* ***Uematsu et al. (2012).*** *© Copyright 2012.*

neurological symptoms of the brain-lung-thyroid syndrome are caused by a predominantly prenatal impairment of the differentiating striatum, and they proposed that deep-brain-stimulation might have a positive effect on these patients, as observed in patients with Huntington's disease. In accordance with the data highlighting an important role of the dopaminergic system, in the majority of patients who were treated with dopaminergic drugs, a rapid dose-dependent improvement in chorea[209,237] was observed (see Table 14). Konishi et al. (2013) provided a valuable explanation for the effects of dopaminergic drugs using dopaminergic PET[228] scans in two patients with a splicing site variant and chorea (father and 24-year-old son; Patients 269 and 286 respectively; Table 27, Part 2). In these patients, who previously had normal conventional MRIs, the researchers found heavily impaired relative dopamine binding to striatal D2 receptors (see Figure 34).

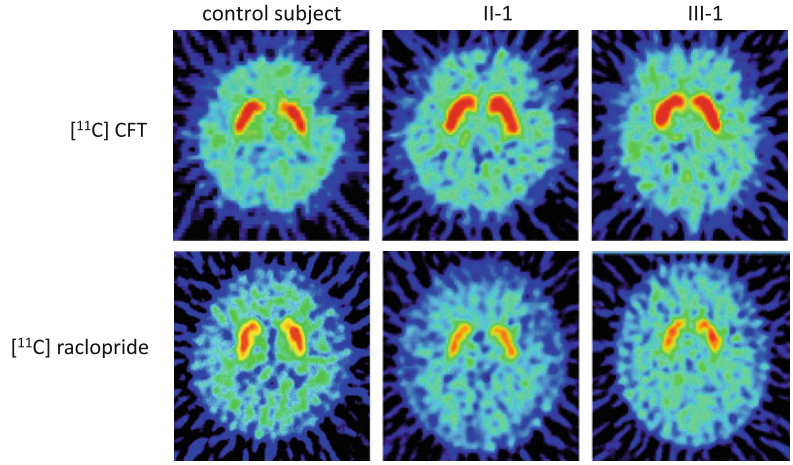

Figure 34: Dopaminergic positron emission tomography (PET) scans depicting impairments of the postsynaptic dopaminergic function in the basal ganglia of patients with the c.464-9C>A;p.(Ile155Thrfs*286) variant. *The upper figures show PET scans using the radiolabeled dopamine reuptake inhibitor (^{11}C)-CFT, which reveal no difference between controls and patients, thus indicating that the presynaptic dopaminergic function is intact. In contrast, in the lower figures, PET scans using the low-affinity dopamine D2 receptor probe (^{11}C)-raclopride show that D2 binding is reduced (by 70%) in the patients, compared to the control, indicating that the D2-receptors are less available and that the indirect pathway is therefore impaired. Figure taken from Konishi et al. (2013)[228] and reprinted with permission from the Journal of Neurology: Springer Nature. Benign hereditary chorea: dopaminergic brain imaging in patients with a novel intronic NKX2.1 gene mutation,* ***Konishi et al. (2013)*** *© Copyright 2013.*

These observations suggest an impairment of the indirect pathway in this disorder and possible compensation of impaired or underdeveloped neuronal networks[55,188,209,225]. The hypothesis of compensation is also supported by the tendency of improvement in motor symptoms in teenagers and adults[225] (see Section 3.5.5.2.2 for details). It also explains why antipsychotics, such as haloperidol[61,223,228], risperidone[236], olanzapine[55] and sulpiride[63], which are also used to treat Tourette's syndrome and Huntington's disease and act as D2-antagonists, mainly have no positive effects and can even worsen the chorea (see Table 14). Suppression of chorea by these

drugs in Huntington's disease is said to occur by blockage of the D2 receptors and/or sedating effects[295].

Among all dopaminergic drugs, mainly levodopa, which is primarily used to treat Parkinson's disease, was tried in patients with chorea. Levodopa is the precursor to dopamine that can penetrate the blood brain barrier and is converted to dopamine by the so-called dopa decarbocylase. Oral administration of levodopa is generally combined with an inhibitor of dopa decarboxylase (e.g. carbidopa) to prevent peripheral conversion of levodopa to dopamine, which would decrease the therapeutic effect and cause gastrointestinal side effects such as nausea and vomiting. In the majority of cases, it was not reported whether and, if yes, with which dopa decarboxylase it was combined. Two separate analyses were thus conducted in the present study (levodopa and levodopa/carbidopa; see Table 14). In the group without information on a combination with a decarboxylase, the improvement rate was low (38%), whereas it was moderate in the levodopa/carbidopa group (67%). Overall, the effectivity of levodopa in these patients was found to be limited by gastrointestinal side effects (see Table 14) that are associated with the necessity to use higher dosages than in disorders such as Parkinson's disease to achieve a clinical improvement[56,63,209]. Furthermore, in one patient, the dopamine agonist ropinirole, which triggers D2 receptors and is another drug used to treat Parkinson's disease and restless legs syndrome, was found to be successful[55]. Methylphenidate, which is a dopamine reuptake blocker used to treat ADHD in some patients, recently had a high success rate (100%) in improving chorea[61,68,215,235,241]. ADHD has furthermore been linked to striatal pathologies[61]. In addition, tetrabenazine[63,65,195,202,237,245,247], an orphan drug used to treat Huntington's disease, exhibited promising results (improvement in 73%), while often being limited by side effects (40%). Tetrabenazine is an oral monoamine depleter acting as a reversible high-affinity blocker of monoamine uptake by binding to the vesicular monoamine transporter 2, which transports dopamine in addition to serotonin, norepinephrine, and histamine. As a result, the monoamine degradation, particularly of dopamine, in the neurons is augmented[296]. To better understand the effects of *NKX2-1* deficiency in the brain and the treatment options, future studies investigating the dopaminergic deficit and the role of dopaminergic drugs are necessary[209].

Moreover, the oral anticholinergic drugs trihexyphenidyl (used to treat Parkinson's disease, essential tremor, and akathisia side effects of antipsychotics)[61,63,201,209], biperiden (used to treat all forms of parkinsonism and extrapyramidal symptoms)[209], and benztropine[62] (used for the treatment of Parkinson's disease and akathisia side effects of antipsychotics) were all found to be helpful (improvement in 100%; see Table 14). This is plausible because of the suspected

impairment of GABAergic and cholinergic neurons in humans[209], similarly to mice. In summary, all of these studies point to a major role of the striatum in *NKX2-1* haploinsufficiency and demonstrate that our data in mice can indeed be related to the disorder in humans. In addition, two treatments used in essential tremor were tried in the patients: propranolol and ethanol consumption. The mechanism of how propranolol, which is a nonselective β-adrenergic receptor antagonist, reduces tremor is not fully understood; however, it is widely accepted that blockage of peripheral non-cardiac β-2 receptors in the muscle spindle plays a role.[298] Ethanol intake was found to significantly reduce symptoms of essential tremor, myoclonus-dystonia, and some other hyperkinesis. One of the suspected underlying mechanisms for the responsiveness to ethanol is that ethanol enhances the tonic inhibition of dysfunctional GABAergic pathways[299]. Ethanol consumption in *NKX2-1*-deficient patients was found to be effective in only half of them, and it thus remains unclear whether chorea in this disorder is actually ethanol-responsive. In contrast, anticonvulsive drugs, except for topiramate, did not improve the symptoms and often had to be stopped because of side effects. The mechanisms underlying the effects of topiramate, which is used to treat epilepsy, cluster headaches, and migraines, are not fully understood; however, apart from other effects, suppression of mesocorticolimbic dopaminergic activity by enhancement of the inhibitory action of GABA has been found. Interestingly, this effect was shown to be useful in the treatment of cocaine dependency[297].

Using neuroimaging, other abnormalities that could also be linked to *NKX2-1* deficiency were occasionally found in the patients. For instance, there were pathologies of the corpus callosum, such as thinning[181], hypoplasm or agenesis[13,60,181,184], and truncation[250] (see Figure 25). These could be related to interneuronal loss, as accumulations of Nkx2-1-positive, reelin-secreting interneurons were found in the corpus callosum of rats[93]. Moreover, pituitary pathologies were detected, ranging from cystic masses[46,62] to an empty sella turcica[194,195] (3%, see Figure 20), suggesting that *NKX2-1* is also expressed in the human pituitary gland, similarly to mice[2,300], where it was shown to be essential, as Nkx2-1 null mice have pituitary agenesis[32].

Furthermore, mice with a homozygous ablation of the transactivation domain 1 (i.e. gene region c.96–565), lack a pituitary (see Figure 3), while the ablation of the transactivation domain 2 (i.e. gene region c.753–881) and a variant with selective deletion of the seven phosphorylation sites (see Figure 3)[29,293] did not impair pituitary development. In a subanalysis of patients with variants falling into the region of the gene region c.96–565 (see Figure 20a), an empty sella syndrome was detected in 2% (2/99 patients), namely, a father (Patient 235; see Table 27, Part 2) and daughter (Patient 236; see Table 27, Part 2) with a the same nonsense variant[243], while

in patients with variants of the gene region c.753–881 (see Figure 20b), no pituitary abnormalities were found. The difference for pituitary manifestations between the two groups was not found to be statistically significant (see Table 35). Moreover, the majority of pituitary abnormalities were rather found in WGD or variants of other regions (cf. 3.5.5.2.1). Therefore, no conclusions can be drawn from the finding in the c.96–565 group.

4.2.5.2.4 The lung phenotype – the leading cause of mortality

Agreeing with previous studies, approximately half of the patients were found to have a lung phenotype (46% in this study and 46% in Gras et al.[61] versus 54% in Carré et al.[60]). In mice, Nkx2-1 expression was identified in type II pneumocytes and Clara cells[33] and was found to be essential for pulmonary branching, alveolarization and surfactant production[9,33,35,38]. In accordance with the findings in mice, histological studies in humans with a *NKX2-1*-deficiency-associated lung phenotype (see Tables 16 and 17) revealed 1) weak or absent *NKX2-1* staining in type II pneumocytes[13,196,250]; 2) the location of a disease-causing *NKX2-1* variant in the lung tissue (Patient 13; Table 25, Part 1)[250]; 3) impaired pulmonary architecture[17]; and 4) reduced and hyperplastic type II pneumocytes[13,14,17,47,196,234] and consequently reduced or absent surfactant staining[13,234].

Overall, the pulmonary phenotype in *NKX2-1* variants (see Tables 15–17) appears to be mainly attributable to the deficiency in surfactant[60]. Surfactant is critical for the regulation of the alveolar surface tension, suppression of bacterial growth and immunomodulation[35] and its deficiency is known to cause IRDS and is also linked to inflammatory lung diseases such as asthma[35]. To date, the lung phenotype in patients with *NKX2-1* deficiency has been mainly characterized by IRDS (51% in this study versus 76% in Carré et al.[60]), recurrent pulmonary infections (40% in this study versus 24% in Carré et al.[60]) and ILD (15% in this study versus 11% in Carré et al.[60]). In support of the clinical observations, a) lung CTs (see Table 16) often revealed signs of obstructive or interstitial lung disease with atelectasis[14,213,250], emphysema[47,250,255], ground-glass opacities[14,61,191,192,196], and fibrosis[14,47,247,255], and b) histologic studies futher showed increased macrophages[14,192,196] and increased alveolar muscularization[250] in patients with ILD and/or IRDS. Additionally, alveolar proteinosis[47] (a surfactant-associated rare disorder) and NEHI[165,166] (a rare form of ILD) were found.

Furthermore, in line with the link of surfactant deficiency to inflammatory diseases, the present study showed that the lung phenotype extends these typical manifestations (see above) by bronchial hyperreactivity/asthma (16% in this study versus no previous statistics), which is increasingly reported as the age of the reported patients rises[14,26,46,55,61,64,193,210,227,241].

In the present analysis, the lung phenotype was found to be associated with a high morbidity, particularly when IRDS was involved. The analysis revealed that one fifth of patients (21%) with a lung phenotype require mechanical ventilation and one quarter (27%) short or long term supplemental oxygen with a median duration of 1.5 years (IQR 0.2–4.8).

Most importantly, the study showed an early lung-associated mortality (i.e. 86% of lethal cases are due to the lung phenotype). The overall mortality rate for the whole patient population was 6%, which is comparable to previous data (7% in Carré et al.[60]). Subanalysis of lethal cases showed that a) the primary pulmonary cause of demise was respiratory failure in infancy (71%) followed by lung cancer (10%); and b) the median age at death was 1.5 years (IQR 0.4–7) for patients with pulmonary causes of death compared to 28 years (8, 23, and 58) for patients without respiratory cause. The difference in survival between patients with and without a lung phenotype were shown to be statistically significant in a Kaplan-Meier analysis (see Figure 26). The mean estimated survival of patients with a lung phenotype was found to be only 51.1 years (45.5–56.6 years, 95% confidence interval) compared to 71 years (65.7–76.3 years, 95% confidence interval) for those without pulmonary manifestations and the overall estimated patient survival was 64 years (59.7–68.6 years, 95% confidence interval). In summary, in accordance with previous studies[60], it can be concluded that all lethal cases attributable to *NKX2-1* deficiency are lung-associated. Interestingly, lung cancer was rarely seen in these patients (2% of the patients with a lung phenotype) although the role of *NKX2-1* in lung cancer is well-established[50]. This observation might be explained by the young age of the patients (see Figure 17).

In contrast to the findings for the brain phenotype, a subanalysis of the different genotypic groups mostly did not reveal any significant differences (see Tables 23–24 and Tables 29–31). Interestingly, both the HD group (compared to WGD, non-HD+ and non-HD; see Table 29) and the missense group (compared to WGD and AOV; see Table 32) showed an association with isolated lung phenotypes. As already discussed in Section 4.2.5.1.2 the analogies between the HD group and the missense group are most likely owed to the high frequency of missense variants in the HD group (see Table 24). While the association of lung-only phenotypes with the HD group and missense variants that are predominantly located in the HD (86%; see Table 30) suggests a higher vulnerability of the lungs to HD variants, the subanalysis for functional domains (see Table 34) does not support the assumption that the HD is more important than other domains for the lungs. On the contrary, the finding of rare lung manifestations (see Section 4.2.5.1.2) and significantly lower lung scores in no-domain variants (see Table 33) rather point to the importance of functional domains in general for the lungs and less vulnerability to

variants without domain impairment. Moreover, observations provided in mice with three different variants that do not involve the HD (see Figure 3) support the theory that the HD is not more important than other domains for the lungs. As illustrated in Figure 3, homozygous mutation of the gene region c.96-565 or the gene region c.753-881 as well as the mutation of the seven phosphorylation sites were all found to result in lung dysgenesis or agenesis, which is incompatible with life. Subanalysis of the patients with variants of the gene region c.96-565 and gene region c.753-881 showed an association of the gene region c.753-881 with IRDS (see Table 34). This finding might be explained by the mere fact, that only the variants of the c.753-881 gene region were identified to involve any phosphorylation sites.

As only three out of seven phosphorylation sites were found to be impaired by the variants in the presented patient series (see Figure 18), namely, the ones located in either the NK2-SD (c.850_852;p.284) or the C-TAD (c.1069_1071;p.357 and c.1096;p.366), it was not possible to conduct a subanalysis of the phosphorylation sites independent of the functional domains. However, there was one patient with a missense variant located in one of the phosphorylation sites, namely c.1096;p.366, without impairment of any other gene regions (Patient 338; Table 28, Part 2)[13]. This patient with a lung-thyroid phenotype suffered from severe lung manifestations (IRDS and pulmonary hypertension) that led to end-stage respiratory failure and death at the age of 6 months[13]. While this phenotype supports the theory of a particular role of phosphorylation sites for the lungs, this singular observation could also be incidental. To further evaluate the potential role of phosphorylation sites in humans, more case reports with predominant involvement of the phosphorylation sites would have to be studied.

4.2.5.2.5 The thyroid phenotype – the manifestation that might affect the brain phenotype and can be detected by screening

The thyroid is the organ with the second most frequent manifestations (see Table 18; 64% in this study versus 67% in Gras et al.[61] and 87% in Carré et al.[60]) that predominantly involve hypothyroidism (93% in this study versus 87% in Carré et al.[60] and 67% in Gras et al.[61]), which is often congenital (45% in this study versus 55% in Carré et al.[60]) and associated with thyroid dysgenesis (i.e. agenesis, hypoplasia, and ectopy; overall 23% in this study versus 45% in Carré et al.[60]). Interestingly, autopsies in two patients with congenital hypothyroidism and missense variants of the HD revealed different extents of impairments of *NKX2-1* staining. While in one patient both the *NKX2-1* staining in the epithelium of follicles and the TG labeling were reduced (Patient 71; Table 26, Part 1)[196], in the other patient *NKX2-1* staining was intact and colloid depleted[234] (Patient 141; Table 26, Part 2). These findings are in line with the knowledge that Nkx2-1 that exerts its role by binding to the promoters of the TG and thyroperoxidase genes[24]

is required for the differentiation of thyrocytes and the establishment and functional maintenance of their follicular architecture[39,40] in mice.

Congenital hypothyroidism, which is the typical manifestation of the thyroid phenotype, is often diagnosed through neonatal screenings allowing for early identification and treatment. In the general population its incidence is 1:2,000–1:4,000[203,227] and it involves thyroid dysgenesis in 80–85% that includes ectopy (50–60%)[203], athyrosis (20–30%), and hypoplasia (5%). However, newborn screenings are not performed everywhere, and the laboratory thresholds used for TSH are also different among laboratories and centers in some countries, thus leading to inconsistent diagnosis[227]. As already discussed for the lung, also carcinomas of the thyroid gland were rarely found. Considering the role of *NKX2-1* in thyroid malignancies[50], this finding might be best explained by the young age of the patients (see Figure 17).

Similar to the findings in the brain phenotype, a subanalysis of the genotypic groups revealed that 1) the frequency of the overall thyroid phenotype (see Tables 23 and 29) and the overall thyroid score (see Table 24) were significantly lower in HD variants compared to all other groups and 2) overall thyroid manifestations, hypothyroidism and congenital hypothyroidism were significantly less frequent (see Tables 30 and 32), and the overall thyroid score was lower (see Table 31) in missense variants, compared to both WGD and AOV. Again the analogy of the findings of the HD group and the missense group can be explained by the high frequency of missense variants in the HD group (cf. Section 4.2.5.1.2).

In analogy to the finding of lung agenesis in Nkx2-1 null mice[32] and homozygous mutants with deletion of the gene regions c.96–565 and c.753–881 (see Figure 3), these mice also lack a thyroid[29]. The fact that the isolated deletion of either of the transactivation domains causes thyroid agenesis suggests that both regions are essential for proper thyroid development, while the intact HD cannot save the thyroid organogenesis[29]. The similarities of the two regions in mice (see above) are reflected in the similarities and the lack of statistical differences in humans with variants of the gene regions c.96–565 and c.753–881 (see Tables 34 and 35).

In contrast to the lung phenotype, there were mostly no differences between variants involving domains and those without domain impairment for the thyroid scores (Table 33) suggesting that the thyroid is also vulnerable to alterations of gene regions outside its known functional domains and sites.

Interestingly, the non-HD group exhibited an association with the isolated thyroid phenotype compared to all other groups (WGD, HD, and non-HD+; see Tables 23 and 29). This might be because non-HD variants have a particularly high phosphorylation site involvement (85% of variants), as they mostly involve the NK2-SD and/or C-TAD (see Section 3.5.6.2.1 for details). The finding that HD-only variants in humans have significantly lower thyroid scores compared to multi-domain variants and C-TAD-only and N-TAD-only variants is in line with this (see Table 33) and provides further evidence for the finding that the HD is less important for the thyroid development and function than the transactivation domains N-TAD and C-TAD.

Moreover, evidence for the importance of the phosphorylation sites for the thyroid is provided by mice with homozygous mutation of the seven phosphorylation sites that display smaller but correctly located thyroid glands (see Figure 3)[29] and the findings in heterozygous mutants of the same mutant line that developed thyroids but had impairments of thyroid architecture, while DNA binding, protein expression, and nuclear localization of Nkx2-1 remained intact[29]. Furthermore, as for the lung phenotype, a comparison of the single point variant of a phosphorylation site (c.1096;p.366) with the phosphorylation site mutation in mice revealed some similarities. The patient was suffering from a severe lung-thyroid phenotype with congenital hypothyroidism but had no thyroid dysgenesis (Patient 338; Table 28, Part 2)[13]. Together with the data obtained from mice, these findings, although they might be incidental as they were only reported in one patient, might indicate that phosphorylation is not required for early thyroid development, but for proper cellular organization thereafter also in humans as already reported in mice[32,291].

In summary, these analyses indicate that the thyroid is generally vulnerable to any type of variants independent of domain involvement. However, variants involving larger regions of the *NKX2-1* gene and/or more phosphorylation sites (i.e. WGD, AOV, non-HD+, and non-HD variants) appear to result more frequently in a thyroid phenotype and more severe manifestations.

An important aspect that we did not investigate in our study, but that should be considered, is the potential influence of hypothyroidism on the neurological phenotype. One case report exemplifies the importance of thyroid hormones for brain function: the case of an infant with congenital hypothyroidism and choreoathetosis, who displayed significant improvements in the latter after substitution of levothyroxine (Patient 253; Table 27, Part 2).[66] Interestingly, Nkx2-1 is expressed by follicular cells in the thyroid that secrete T4, but also by tanycytes of the periventricular ependyma of the third ventricle[301] that produce the enzyme deiodinase 2, which locally converts T4 into the biologically active T3[302]. During brain development, even the

slightest T3 deficiency can cause severe impairments of the motor system, memory, and language[35]. In particular, PV-containing GABAergic neurons highly depend on a stable T3–level, as has been demonstrated by Gilbert et al.[303] (2007), who found that deficiency results in a severe loss of cortical and hippocampal GABAergic neurons. The loss was most pronounced in the gyrus dentatus (70%). A similarly intense loss, particularly of the fibers, was also observed in our prenatal Nkx2-1 mutants[1] (see Figure 29). Hypothyroidism was moreover found to harm cholinergic BF neurons, as the deficiency in thyroid hormones prevented proper dendritic branching in another study[304]. Thus, *NKX2-1* variants have the potential to impair the effects of thyroid hormones in the brain without the actual presence of hypothyroidism.

4.2.5.3 Additional clinical features and possible links to *NKX2-1* or other deleted genes

An analysis of additional clinical features (see Sections 3.5.5.6 and 3.5.6.5), which were encountered in 20% of patients (see Table 20), revealed that they were more frequent in non-WGD (58%) compared to WGD (42%) variants. A closer examination of the respective variants showed that nearly all patients with WGD had additional deletion of other genes (97%), in contrast to patients with non-WGD variants who had an involvement of only 5% (see Table 37). Figure 35 illustrates the location of the *NKX2-1* gene on chromosome 14 and the genes in close proximity that are often found to be additionally deleted (i.e. *PAX9, NKX2-8, SLC25A21, RALGAPA1, SFTA3, MBIP,* and *MIPOL1*).

Several authors have discussed the possible roles of additional variants of other genes in patients with *NKX2-1* vairants[69,184]. However, the overall degree to which the WGD and non-WGD variants with additional clinical features differ in terms of involvement of other genes (see Table 37 for details) suggests that the additional clinical features might stem from *NKX2-1* deficiency or be incidental rather than caused by deficiency in other genes. In the following section, the possible role of additionally deleted genes is discussed for the most frequent additional clinical features.

4.2.5.3.1 Hypo- and oligodontia

Patients with hypo-/oligodontia (4% of all patients) nearly exclusively had WGD (93%, see Table 37). This association was found to be statistically significant in comparison to missense, AOV, HD, non-HD+, and non-HD variants (Tables 23, 29, 30, and 32) and the vast majority of patients had an additional deletion of *PAX9* (93%) – a gene involved in craniofacial development and odontogenesis. *PAX9* is known to be associated with hypo- and oligodontia[69], and hypo-/oligodontia in these patients is thus most likely linked to *PAX9* deletion.

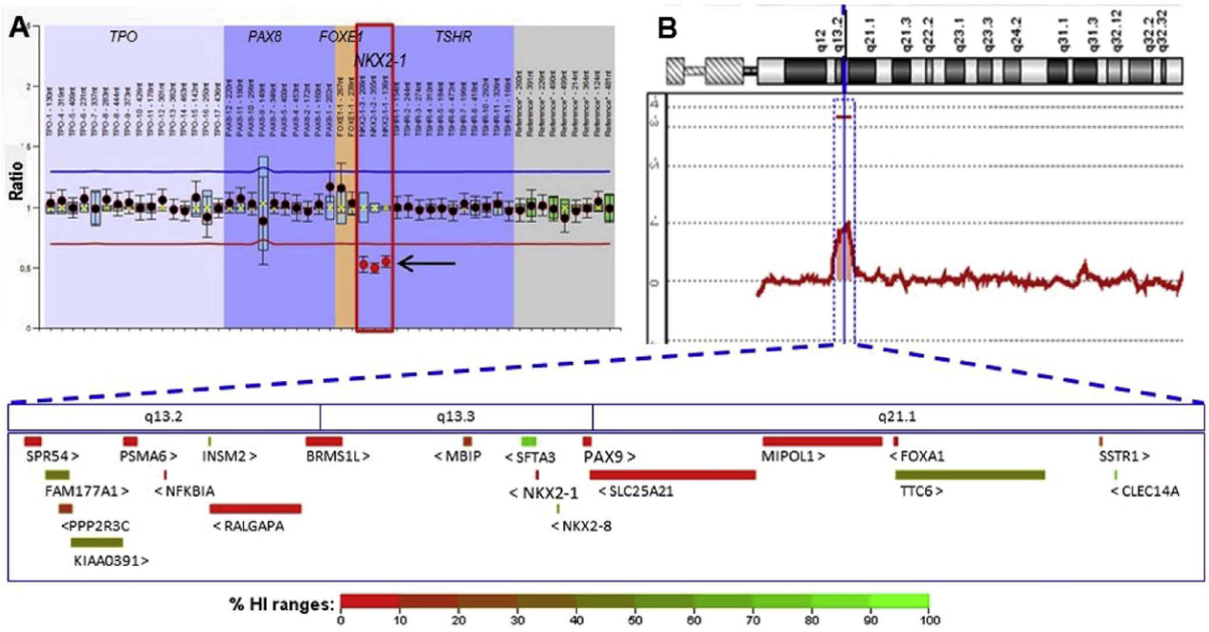

Figure 35: Location of frequently additionally deleted genes in WGD. *The figure shows other frequently deleted genes in proximity to NKX2-1. At the bottom of the figure, the legend indicates a color-coded range of haploinsufficiency. For genes with known haploinsufficiency, the grade of haploinsufficiency is displayed below the gene name. Figure taken from Villafuerte et al. (2018)[239] and reprinted with permission from the European Journal of Medical Genetics: Elsevier, The Brain-Lung-Thyroid syndrome (BLTS): A novel deletion in chromosome 14q13.2-q21.1 expands the phenotype to humoral immunodeficiency,* ***Villafuerte et al. (2018).*** *© Copyright 2018.*

In addition, deletion of *NKX2-8* (64%; a gene involved in neural tube development, foregut development, and lung cancer) and *SLC25A21* (57%; no known human phenotype[239]; associated with one case of a brain-lung-thyroid syndrome without *NKX2-1* variant[186]) was frequently found in these patients. While these additional genes' deletions in variants with *PAX9* deletion are explained by their close proximity to both *NKX2-1* and *PAX9* (see Figure 36), no role in tooth development of humans is known. However, *SLC25A21* was found to regulate *Pax9* in mice[305], and there may thus be an additional impact of *SLC25A21* deletion in some of the patients with *PAX9* deletion and hypo-/oligodontia.

4.2.5.3.2 Facial dysmorphism

Facial dysmorphism (reported in 4% of all patients; see Table 20) was found to be associated with WGD (67% of cases involved WGD; see Table 37) when compared to missense, AOV, HD, non-HD+, and HD variants (Tables 23, 29, 30 and 32). In the majority of cases, concomitant deletion of *PAX9* (70%) and *RALGAPA1* (70%) was found. While there is no known function of *NKX2-1* that, if missing, could lead to facial dysmorphism, both *PAX9* and *RALGAPA1* deficiency can result in impairments of craniofacial development. Therefore, it is presumable that facial dysmorphism is rather associated with the deletion of these genes than *NKX2-1* deficiency.

Interestingly, *RALGAPA1* haploinsufficiency is furthermore associated with a human phenotype involving facial dysmorphism in addition to severe intellectual disability, microcephaly, developmental and speech delay, muscular hypotonia, and seizures[260]. In fact, the majority of patients with a deletion of *RALGAPA1* in this study also had most of these symptoms (63% developmental delay, 53% muscular hypotonia, 37% severe intellectual disability, and 21% microcephaly; cf. Section 3.5.6.5). Seizures and speech delay were not found in our patients with concomitant deletion of *RALGAPA1*.

4.2.5.3.3 Short stature and retarded skeletal age

Short stature (7% of all patients; see Table 20) and retarded skeletal age (5% of all patients; see Table 20) were more frequently seen in patients without deletion of other genes (see Table 37). However, it remains unclear whether these features can be linked to *NKX2-1* deficiency. While Nkx2-1 expression was found in both the calcitonin-producing C-cells of the thyroid and the parathyroid of mice (see Malt et al. for review)[35], Nkx2-1 null mice had normal parathyroids[32], suggesting that Nkx2-1 expression is not essential for parathyroid development. Detailed studies in heterozygous mice discovered that there was no fusion of the Nkx2-1-positive ultimobranchial body cyst – an embryological structure (out-pocketing of the forth pharyngeal pouch) normally fusing with the thyroid diverticulum to give rise to the C-cells[175] – and thus there was also no dissemination of C-cells into the thyroid[175]. As hypoparathyroidism was rare in this syndrome (three cases[62,68]) and since not much is known about the role of Nkx2-1/*NKX2-1* in the parathyroid and C-cells, further investigation is needed to evaluate a possible link between the findings in mice and humans. In conclusion, it remains to be elucidated whether these features are related to the *NKX2-1* variants.

4.2.5.3.4 Hyper- and hypophagia

Disturbed food intake, namely hypophagia (2% of all patients; see Table 20) and hyperphagia (3% of all patients; see Table 20), were mostly found in patients without deletion of other genes (see Table 37). As Nkx2-1 is strongly expressed in the hypothalamic arcuate nucleus, which is the center of feeding behavior[301] (see Malt et al. for a review[35]), its deficiency might cause these disturbances. Detailed studies further detected Nkx2-1 to stimulate both the food-intake stimulating NPY[301] and agouti-related peptide, as well as to inhibit the anorexia-inducing proopiomelanocortin[306]. In line with this, blockage of Nkx2-1 by leptin decreases food intake in mice, and its upregulation increases food intake[306]. Therefore, it is reasonable to assume that hypo- and hyperphagia in the patients are related to *NKX2-1* deficiency.

4.2.5.3.5 Fever of unknown origin

Fever of unknown origin without infection (6% of all patients; see Table 20) was mainly found in patients without deletion of other genes (see Table 37). This phenomenon might be attributable to *NKX2-1* deficiency, as Nkx2-1 expression was detected in cells involved in core temperature regulation and circadian rhythm. Its expression was identified in melanopsin-containing, adenylate cyclase-activating, polypeptide-producing, intrinsically photosensitive retinal ganglion cells, as well as in the suprachiasmatic and dorsomedial hypothalamic nucleus (see Malt et al. for a review[35]).

4.2.5.3.6 Other additional clinical features

Other additional clinical features and disorders (Tables 20 and 37) which are not detailed above, are listed in Table 39. They have been found in frequencies comparable to the general population and can therefore be regarded as incidental.

Table 39: Frequencies of rare additional clinical features compared to the general population

Additional clinical feature/disorder	Incidence among patients with *NKX2-1* variants	Incidence in the general population
Congenital heart disease	9/351 patients (3%)[26,46,194,254]	8:1,000[307]
Hyperextendable joints	4/351 patients (1%)[239]	5–18%[308]
Hydronephrosis	2/351 patients (<1%)[67,183]	1–2%[309]
Megabladder	2/351 patients (<1%)[67]	<1%[310]
Maldescensus testis	2/351 patients (1%)[67,183]	1–8%[311]
Hypogammaglobulinemia	1/351 patients (<1%)[239]	1:10,000-30,000[312]
Developmental dysplasia of the hip	1/351 patients (<1%)[239]	1–7%[313]
Polydactyly	1/351 patients (<1%)[206]	1:1,500-3,000[314]
Vesico-ureteral reflux	1/351 patients (<1%)[181]	1–25%[315]

4.2.6 Limitations of the genotype-phenotype correlation analysis

Two important limitations of the study must be kept in mind when interpreting the results. First, the significance of the *p*-values obtained in this study remain exploratory, as no complete correction for multiple testing was possible. Second, some of the data depend on the newly developed disease severity score, and the results might have been different, if the score would have been designed differently. Nevertheless, on the whole, the data mostly support previous knowledge on the syndrome and findings in mice. Further evaluation is, however necessary. The present study is intended as an overview that allows for a better understanding of the current knowledge about the syndrome and assists future research on the disorder.

4.2.7 Conclusions for clinical practice

4.2.7.1 Identification of patients with possible *NKX2-1* deficiency

Genetic testing and thyroid screening should be performed on a routine basis in patients with suspicious symptoms, particularly those with early onset chorea[242,62] and suspected ataxic and

dyskinetic cerebral palsy, to prevent misdiagnosis[212,222]. It is important to check thyroid values in children with these types of neurological impairments to allow for an early start of levothyroxine substitution, if necessary. *NKX2-1* screening can also be helpful in patients with congenital hypothyroidism who develop neurological impairments despite sufficient levothyroxine substitution[208]. In addition, testing should be considered in patients with diffuse lung disease of unknown etiology and additional neurological or endocrinological peculiarities[13]. Due to the variable combination of symptoms, screening should also be performed when a thyroid phenotype is missing[60,226]. Moreover, the variant analysis should not be restricted to *NKX2-1,* as there might also be deletions of regulatory elements and closely related genes such as *MBIP*[26], *SLC25A21*[186], and *SFTA3*[185], which have been found to mimic the brain-lung-thyroid syndrome in a few patients.

For this purpose, it is preferable to also search for larger deletions using multiplex ligation-dependent probe amplification. Furthermore, parental testing, family screening, and anamnesis are of paramount importance to study the expressivity of the various variants[13,63,66,224], including the relevance of environmental factors[230,252]; to identify genotype-phenotype correlations; and to calculate the risk of recurrence in families that wish to have more children[13].

4.2.7.2 Diagnostics and treatments

Patients with *NKX2-1* variants should be systematically screened and followed up for thyroid and lung diseases in terms of prevention and early onset of treatment[67]. Knowledge of the pulmonary manifestations is particularly important here, since patients with lung manifestations have a significantly higher mortality rate at a young age (this study). As cognitive impairments are in fact relevant to this disorder, screening for them must be actively pursued in all cases, so that patients can receive special assistance and treatments early on, if necessary.

Furthermore, since there are no specific treatments for these patients, each symptom should be treated according to standard guidelines, if available. Whenever possible, brain MRIs and nuclear brain scans using advanced imaging techniques should be performed in these patients. In terms of clinical research, one promising method to investigate pathophysiological mechanisms underlying *NKX2-1* deficiency could be to utilize molecular brain scans with dopaminergic radioligands to explore the direct effects of dopaminergic drugs on these patients. In the future, as *NKX2-1* haploinsufficient patients age, the possible oncogenic role of *NKX2-1* should considered. In fact, a few cases of lung[61,47,210] and thyroid[249] cancer have already been reported.

4.2.7.3 Reporting of patients in the literature and genotype-phenotype analysis

The case reports need to be more comparable in terms of patient information and diagnostics providing information on all three organs. This will eventually allow for a better analysis of

genotype-phenotype relations. To this end, the severity score developed in this study might be helpful. Additional patient follow-up studies are ultimately required to observe the development in adulthood, and more information obtained from brain autopsies is needed[65] to enable the actual identification of the effects of *NKX2-1* variants in the human brain. As patients are currently predominantly children or young adults with a more or less normal life expectancy (see this study), the amount of available data will most likely remain low in the next few decades. Therefore, neuroimaging is crucial. Nonetheless, essential information can be drawn from every single case, and postmortem studies should hence be encouraged in all lethal cases.

4.3 Summary

Cholinergic septohippocampal neurons express the neurotrophin receptors *trkA* and *p75^{NTR}* and undergo degeneration when they are disconnected (e.g. by axotomy from their target the hippocampus). An analysis of *NGF* and its receptors in the present study revealed that the loss of BF neurons and their axons in both mutant mouse lines is indeed accompanied by a substantial reduction in the expression of the *NGF* receptors *trkA* and *p75^{NTR}*, despite a maintained *NGF* level. Therefore, *NGF* deficiency, which appears to be proportional to the amount of the loss of cholinergic BF neurons in the pre- and postnatal mutant line, could be ruled out as the cause of the receptor downregulation. Nkx2-1 is known to be expressed from E10 onwards in the MGE, which is the origin of cholinergic and PV-containing GABAergic neurons of the BF. Our working group has previously demonstrated that cholinergic and PV-containing GABAergic neurons of the BF degenerate and die already at the subventricular zone when they become postmitotic by using a GAD67-Cre mutant line with prenatal deletion of Nkx2-1. Therefore, at this point, neurotrophins cannot influence them. In the ChAT-Cre postnatal mutant line, the present study showed that the Nkx2-1 deletion exerts its effect exactly at the moment when *NGF* and its receptors become upregulated. We thus conclude that the presence of *NGF* is neither capable of overwriting the effects of the pre- nor of the postnatal Nkx2-1 deletion. The analysis of *GAP-43* furhter suggests that the loss of both cholinergic and PV-containing GABAergic neurons, including their developing axonal projections, are compensated for by the sprouting processes of other, likely neighboring, neuronal systems – at least at the time point analyzed.

Evidence for the comparability of our mice studies with the human brain phenotype have been provided by brain autopsy studies of a) two healthy controls, where the same neuronal populations of the globus pallidus and putamen were shown to express *NKX2-1*[1] (data of our working group), and b) the brain of a 59-year-old women with an isolated chorea, where striatal cholinergic and PV-containing interneurons and their axonal fibers were significantly reduced[65].

More than 21 years after the first case report of a patient with a *NKX2-1* variant[190], this study provides the most comprehensive review of the brain-lung-thyroid syndrome – reporting 351 patients with 167 different variants from the literature – with the first statistical analysis, including survival, genotype-phenotype correlation, and disease severity analysis, using an original and newly developed disease severity score. Despite the variable expressivity of symptoms and the great variety of variants certain genotype-phenotype correlations could be identified.

For the first time, the present study analyzed the role of the different functional domains of the *NKX2-1* gene and found that nearly all patients had variants that involve at least one of four functional domains (i.e. the N-TAD, the HD, the NK2-SD and the C-TAD), indicating their vulnerability to genetic alterations and highlighting their significance for the brain, lungs and thyroid. Moreover, the wide variety of domain involvements in these variants, ranging from no domain impairment to WGD, demonstrates that different parts of the *NKX2-1* gene exert important roles in organogenesis and function. In contrast to previous studies that assumed the HD to play a major role in this disorder, the genotype-phenotype correlation and disease severity analysis revealed that the organ manifestations and the disease severity are significantly influenced by the types of variants (i.e. WGD, missense and AOV), with missense variants being associated with milder phenotypes and single-organ manifestations and WGD being associated with severe organ manifestations and multi-organ phenotypes.

Among the three organs, the brain was found to be the leading manifestation and main cause of morbidity in *NKX2-1* deficiency. In contrast to previous studies, cognitive impairments were found to be an integral part of the syndrome. The brain was demonstrated to be particularly vulnerable to WGD, which nearly always result in a brain phenotype and are associated with cognitive impairments. Moreover, the present data contradict previous studies that reported brain imaging not to be useful in this disorder. In fact, advanced brain imaging techniques (i.e. MRI and nuclear imaging) were shown to be able to identify morphological and functional pathologies of the basal ganglia. In the thyroid, the second most affected organ, WGD were found to cause more severe manifestations compared to any other group of variants. Both the brain and the thyroid phenotype were demonstrated to be vulnerable to any type of variant irrespective of domain involvement and to have lower disease severity scores and less multi-organ manifestations in missense variants. The lung-phenotype was shown extend the classical manifestations of IRDS and recurrent pulmonary infections by asthma/brochial hyperreactivity. And the lungs, however, showed no correlation for the types of variants. Instead, lower lung scores were found in variants without domain involvement, pointing to a stronger role of the functional domains in general and less vulnerability to alterations outside them. Kaplan-Meier analysis for survival revealed an early mortality from pulmonary disorders.

The data obtained in mice and the analysis of the brain phenotype in the literature review (present study) indicate that the symptoms found in *NKX2-1* haploinsufficiency are caused not only by disturbances in the basal ganglia circuitry, but also, on a more fundamental level, namely, by the loss of cholinergic and GABAergic neurons pre- and/or postnatally. Overall, the present study points to a pre- and postnatal lack of *NKX2-1* in the syndrome. As cholinergic neurons

cannot be held entirely accountable for the motor dysfunctions in these patients, our study further emphasizes the major role of GABAergic neurons in this disorder. However, the role of GABAergic neurons for higher cognitive functions such as learning and memory is generally less studied than cholinergic neurons; therefore, their role in this regard is not clear[316].

The postnatal expression and essential character of Nkx2-1 for both neuronal populations[1,2] is a new aspect in Nkx2-1 research (see also commentary of Marin, 2011[317]), and the present study is the first to suggest a possible mechanism underlying the loss of cholinergic neurons in Nkx2-1 deficiency, namely, the loss of *trkA* receptors. In summary, Nkx2-1 is shown to be essential for the development of the cell lines, while *NGF* is required for proper differentiation and maintenance of the neurons downstream of Nkx2-1 in mice. The role of the neurotrophin system in *NKX2-1* haploinsufficiency in humans remains to be elucidated. However, data from research in the aging brain and neurodegenerative disorders provide evidence for the role of *NGF* and its receptors as being critical also in humans.

5. References

1. Magno L, Kretz O, Bert B, Ersozlu S, Vogt J, Fink H, Kimura S, Vogt A, Monyer H, Nitsch R, Naumann T. The integrity of cholinergic basal forebrain neurons depends on expression of Nkx2-1. Eur J Neurosci 2011;34:1767-82.
2. Magno L, Catanzariti V, Nitsch R, Krude H, Naumann T. Ongoing expression of Nkx2.1 in the postnatal mouse forebrain: potential for understanding NKX2.1 haploinsufficiency in humans? Brain Res 2009;1304:164-86.
3. Magno L, Barry C, Schmidt-Hieber C, Theodotou P, Hausser M, Kessaris N. NKX2-1 Is Required in the Embryonic Septum for Cholinergic System Development, Learning, and Memory. Cell Rep 2017;20:1572-84.
4. Rocamora N, Pascual M, Acsady L, de Lecea L, Freund TF, Soriano E. Expression of NGF and NT3 mRNAs in hippocampal interneurons innervated by the GABAergic septohippocampal pathway. J Neurosci 1996;16:3991-4004.
5. McAllister AK, Katz LC, Lo DC. Neurotrophins and synaptic plasticity. Annu Rev Neurosci 1999;22:295-318.
6. Sofroniew MV, Howe CL, Mobley WC. Nerve growth factor signaling, neuroprotection, and neural repair. Annu Rev Neurosci 2001;24:1217-81.
7. Guazzi S, Price M, De Felice M, Damante G, Mattei MG, Di Lauro R. Thyroid nuclear factor 1 (TTF-1) contains a homeodomain and displays a novel DNA binding specificity. EMBO J 1990;9:3631-9.
8. Kim Y, Nirenberg M. Drosophila NK-homeobox genes. Proc Natl Acad Sci U S A 1989;86:7716-20.
9. Ikeda K, Clark JC, Shaw-White JR, Stahlman MT, Boutell CJ, Whitsett JA. Gene structure and expression of human thyroid transcription factor-1 in respiratory epithelial cells. J Biol Chem 1995;270:8108-14.
10. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, Billis K, Cummins C, Gall A, Giron CG, Gil L, Gordon L, Haggerty L, Haskell E, Hourlier T, Izuogu OG, Janacek SH, Juettemann T, To JK, Laird MR, Lavidas I, Liu Z, Loveland JE, Maurel T, McLaren W, Moore B, Mudge J, Murphy DN, Newman V, Nuhn M, Ogeh D, Ong CK, Parker A, Patricio M, Riat HS, Schuilenburg H, Sheppard D, Sparrow H, Taylor K, Thormann A, Vullo A, Walts B, Zadissa A, Frankish A, Hunt SE, Kostadima M, Langridge N, Martin FJ, Muffato M, Perry E, Ruffier M, Staines DM, Trevanion SJ, Aken BL, Cunningham F, Yates A, Flicek P. Ensembl 2018. Nucleic Acids Res. 2017/11/21 ed2018:D754-D61.
11. Oguchi H, Pan YT, Kimura S. The complete nucleotide sequence of the mouse thyroid-specific enhancer-binding protein (T/EBP) gene: extensive identity of the deduced amino acid sequence with the human protein. Biochim Biophys Acta 1995;1261:304-6.
12. Hamdan H, Liu H, Li C, Jones C, Lee M, deLemos R, Minoo P. Structure of the human Nkx2.1 gene. Biochim Biophys Acta 1998;1396:336-48.
13. Hamvas A, Deterding RR, Wert SE, White FV, Dishop MK, Alfano DN, Halbower AC, Planer B, Stephan MJ, Uchida DA, Williames LD, Rosenfeld JA, Lebel RR, Young LR, Cole FS, Nogee LM. Heterogeneous pulmonary phenotypes associated with mutations in the thyroid transcription factor gene NKX2-1. Chest 2013;144:794-804.
14. Nattes E, Lejeune S, Carsin A, Borie R, Gibertini I, Balinotti J, Nathan N, Marchand-Adam S, Thumerelle C, Fauroux B, Bosdure E, Houdouin V, Delestrain C, Louha M, Couderc R, De Becdelievre A, Fanen P, Funalot B, Crestani B, Deschildre A, Dubus JC, Epaud R. Heterogeneity of lung disease associated with NK2 homeobox 1 mutations. Respir Med 2017;129:16-23.

15. Christophe-Hobertus C, Duquesne V, Pichon B, Roger PP, Christophe D. Critical residues of the homeodomain involved in contacting DNA bases also specify the nuclear accumulation of thyroid transcription factor-1. Eur J Biochem 1999;265:491-7.
16. Ghaffari M, Zeng X, Whitsett JA, Yan C. Nuclear localization domain of thyroid transcription factor-1 in respiratory epithelial cells. Biochem J 1997;328 (Pt 3):757-61.
17. Maquet E, Costagliola S, Parma J, Christophe-Hobertus C, Oligny LL, Fournet JC, Robitaille Y, Vuissoz JM, Payot A, Laberge S, Vassart G, Van Vliet G, Deladoey J. Lethal respiratory failure and mild primary hypothyroidism in a term girl with a de novo heterozygous mutation in the TITF1/NKX2.1 gene. J Clin Endocrinol Metab 2009;94:197-203.
18. https://www.uniprot.org/uniprot/P43699 TUCIUtupkNARD-DcAf.
19. Mitchell AL, Attwood TK, Babbitt PC, Blum M, Bork P, Bridge A, Brown SD, Chang HY, El-Gebali S, Fraser MI, Gough J, Haft DR, Huang H, Letunic I, Lopez R, Luciani A, Madeira F, Marchler-Bauer A, Mi H, Natale DA, Necci M, Nuka G, Orengo C, Pandurangan AP, Paysan-Lafosse T, Pesseat S, Potter SC, Qureshi MA, Rawlings ND, Redaschi N, Richardson LJ, Rivoire C, Salazar GA, Sangrador-Vegas A, Sigrist CJA, Sillitoe I, Sutton GG, Thanki N, Thomas PD, Tosatto SCE, Yong SY, Finn RD. InterPro in 2019: improving coverage, classification and access to protein sequence annotations. Nucleic Acids Res 2018.
20. Moya CM, Zaballos MA, Garzon L, Luna C, Simon R, Yaffe MB, Gallego E, Santisteban P, Moreno JC. TAZ/WWTR1 Mediates the Pulmonary Effects of NKX2-1 Mutations in Brain-Lung-Thyroid Syndrome. J Clin Endocrinol Metab 2018;103:839-52.
21. Geer LY M-BA, Geer RC, Han L, He J, He S, Liu C, Shi W, Bryant SH. The NCBI BioSystems database. Nucleic Acids Res. 2010 Jan; 38(Database issue):D492-6. Datablase cited for the entry "NKX2-1" [cited 20.11.2018], Available from: https://www.ncbi.nlm.nih.gov/gene/7080.
22. Harvey RP. NK-2 homeobox genes and heart development. Dev Biol 1996;178:203-16.
23. Zhang LP, Ma BY, Han FX, Wan HL, Wu JP, Yu LH, Wang XR, Zhu JY. Molecular characterization and functional analysis of sheep thyroid transcription factor-1. Genet Mol Res 2012;11:2585-97.
24. Zannini M, Acebron A, De Felice M, Arnone MI, Martin-Perez J, Santisteban P, Di Lauro R. Mapping and functional role of phosphorylation sites in the thyroid transcription factor-1 (TTF-1). J Biol Chem 1996;271:2249-54.
25. Watada H, Mirmira RG, Kalamaras J, German MS. Intramolecular control of transcriptional activity by the NK2-specific domain in NK-2 homeodomain proteins. Proc Natl Acad Sci U S A 2000;97:9443-8.
26. Thorwarth A, Schnittert-Hubener S, Schrumpf P, Muller I, Jyrch S, Dame C, Biebermann H, Kleinau G, Katchanov J, Schuelke M, Ebert G, Steininger A, Bonnemann C, Brockmann K, Christen HJ, Crock P, deZegher F, Griese M, Hewitt J, Ivarsson S, Hubner C, Kapelari K, Plecko B, Rating D, Stoeva I, Ropers HH, Gruters A, Ullmann R, Krude H. Comprehensive genotyping and clinical characterisation reveal 27 novel NKX2-1 mutations and expand the phenotypic spectrum. J Med Genet 2014;51:375-87.
27. Nobrega-Pereira S, Kessaris N, Du T, Kimura S, Anderson SA, Marin O. Postmitotic Nkx2-1 controls the migration of telencephalic interneurons by direct repression of guidance receptors. Neuron 2008;59:733-45.
28. Yoo S. Kinetic analysis of Drosophila Vnd protein containing homeodomain with its target sequence. BMB Rep 2010;43:407-12.
29. Silberschmidt D, Rodriguez-Mallon A, Mithboakar P, Cali G, Amendola E, Sanges R, Zannini M, Scarfo M, De Luca P, Nitsch L, Di Lauro R, De Felice M. In vivo role of

different domains and of phosphorylation in the transcription factor Nkx2-1. BMC Dev Biol 2011;11:9.
30. Fragkouli A, Hearn C, Errington M, Cooke S, Grigoriou M, Bliss T, Stylianopoulou F, Pachnis V. Loss of forebrain cholinergic neurons and impairment in spatial learning and memory in LHX7-deficient mice. Eur J Neurosci 2005;21:2923-38.
31. Zhao Y, Flandin P, Vogt D, Blood A, Hermesz E, Westphal H, Rubenstein JL. Ldb1 is essential for development of Nkx2.1 lineage derived GABAergic and cholinergic neurons in the telencephalon. Dev Biol 2014;385:94-106.
32. Kimura S, Hara Y, Pineau T, Fernandez-Salguero P, Fox CH, Ward JM, Gonzalez FJ. The T/ebp null mouse: thyroid-specific enhancer-binding protein is essential for the organogenesis of the thyroid, lung, ventral forebrain, and pituitary. Genes & Development 1996;10:60-9.
33. Boggaram V. Thyroid transcription factor-1 (TTF-1/Nkx2.1/TITF1) gene regulation in the lung. Clin Sci (Lond) 2009;116:27-35.
34. Montanelli L, Tonacchera M. Genetics and phenomics of hypothyroidism and thyroid dys- and agenesis due to PAX8 and TTF1 mutations. Mol Cell Endocrinol 2010;322:64-71.
35. Malt EA, Juhasz K, Malt UF, Naumann T. A Role for the Transcription Factor Nk2 Homeobox 1 in Schizophrenia: Convergent Evidence from Animal and Human Studies. Front Behav Neurosci 2016;10:59.
36. Kleiner-Fisman G, Calingasan NY, Putt M, Chen J, Beal MF, Lang AE. Alterations of striatal neurons in benign hereditary chorea. Mov Disord 2005;20:1353-7.
37. Goulburn AL, Alden D, Davis RP, Micallef SJ, Ng ES, Yu QC, Lim SM, Soh CL, Elliott DA, Hatzistavrou T, Bourke J, Watmuff B, Lang RJ, Haynes JM, Pouton CW, Giudice A, Trounson AO, Anderson SA, Stanley EG, Elefanty AG. A targeted NKX2.1 human embryonic stem cell reporter line enables identification of human basal forebrain derivatives. Stem Cells 2011;29:462-73.
38. Maeda Y, Hunter TC, Loudy DE, Dave V, Schreiber V, Whitsett JA. PARP-2 interacts with TTF-1 and regulates expression of surfactant protein-B. J Biol Chem 2006;281:9600-6.
39. Dentice M, Luongo C, Elefante A, Ambrosio R, Salzano S, Zannini M, Nitsch R, Di Lauro R, Rossi G, Fenzi G, Salvatore D. Pendrin is a novel in vivo downstream target gene of the TTF-1/Nkx-2.1 homeodomain transcription factor in differentiated thyroid cells. Mol Cell Biol 2005;25:10171-82.
40. Fernandez LP, Lopez-Marquez A, Santisteban P. Thyroid transcription factors in development, differentiation and disease. Nat Rev Endocrinol 2015;11:29-42.
41. Civitareale D, Lonigro R, Sinclair AJ, Di Lauro R. A thyroid-specific nuclear protein essential for tissue-specific expression of the thyroglobulin promoter. EMBO J 1989;8:2537-42.
42. Mizuno K, Gonzalez FJ, Kimura S. Thyroid-specific enhancer-binding protein (T/EBP): cDNA cloning, functional characterization, and structural identity with thyroid transcription factor TTF-1. Mol Cell Biol 1991;11:4927-33.
43. Sussel L, Marin O, Kimura S, Rubenstein JL. Loss of Nkx2.1 homeobox gene function results in a ventral to dorsal molecular respecification within the basal telencephalon: evidence for a transformation of the pallidum into the striatum. Development 1999;126:3359-70.
44. DeFelice M, Silberschmidt D, DiLauro R, Xu Y, Wert SE, Weaver TE, Bachurski CJ, Clark JC, Whitsett JA. TTF-1 phosphorylation is required for peripheral lung morphogenesis, perinatal survival, and tissue-specific gene expression. J Biol Chem 2003;278:35574-83.

45. Pohlenz J, Dumitrescu A, Zundel D, Martine U, Schonberger W, Koo E, Weiss RE, Cohen RN, Kimura S, Refetoff S. Partial deficiency of thyroid transcription factor 1 produces predominantly neurological defects in humans and mice. J Clin Invest 2002;109:469-73.
46. Krude H, Schutz B, Biebermann H, von Moers A, Schnabel D, Neitzel H, Tonnies H, Weise D, Lafferty A, Schwarz S, DeFelice M, von Deimling A, van Landeghem F, DiLauro R, Gruters A. Choreoathetosis, hypothyroidism, and pulmonary alterations due to human NKX2-1 haploinsufficiency. J Clin Invest 2002;109:475-80.
47. Willemsen MA, Breedveld GJ, Wouda S, Otten BJ, Yntema JL, Lammens M, de Vries BB. Brain-Thyroid-Lung syndrome: a patient with a severe multi-system disorder due to a de novo mutation in the thyroid transcription factor 1 gene. Eur J Pediatr 2005;164:28-30.
48. Online Mendelian Inheritance in Man OIM-NIoGM, Johns Hopkins University (Baltimore, MD). [cited 10.10.2018] Available from: https://omim.org/.
49. Orphanet: an online database of rare diseases and orphan drugs. Copyright IAcAfhwon.
50. Patel NJ, Jankovic J. NKX2-1-Related Disorders. In: In: Adam MP AH, Pagon RA, et al., editors. , ed. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-20182014 Feb 20 [Updated 2016 Jul 29].
51. Harper PS. Benign hereditary chorea. Clinical and genetic aspects. Clin Genet 1978;13:85-95.
52. Asmus F, Hjermind LE, Dupont E, Wagenstaller J, Haberlandt E, Munz M, Strom TM, Gasser T. Genomic deletion size at the epsilon-sarcoglycan locus determines the clinical phenotype. Brain 2007;130:2736-45.
53. de Vries BB, Arts WF, Breedveld GJ, Hoogeboom JJ, Niermeijer MF, Heutink P. Benign hereditary chorea of early onset maps to chromosome 14q. Am J Hum Genet 2000;66:136-42.
54. Inzelberg R, Weinberger M, Gak E. Benign hereditary chorea: an update. Parkinsonism Relat Disord 2011;17:301-7.
55. Nakamura K, Sekijima Y, Nagamatsu K, Yoshida K, Ikeda S. A novel nonsense mutation in the TITF-1 gene in a Japanese family with benign hereditary chorea. J Neurol Sci 2012;313:189-92.
56. Peall KJ, Kurian MA. Benign Hereditary Chorea: An Update. Tremor Other Hyperkinet Mov (N Y) 2015;5:314.
57. Seidman JG, Seidman C. Transcription factor haploinsufficiency: when half a loaf is not enough. Journal of Clinical Investigation 2002;109:451-5.
58. Dang VT, Kassahn KS, Marcos AE, Ragan MA. Identification of human haploinsufficient genes and their genomic proximity to segmental duplications. Eur J Hum Genet 2008;16:1350-7.
59. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources [Internet]. Firth HVeaAJHG, 524-533. [cited 14.09.2018]. Available from: https://decipher.sanger.ac.uk/.
60. Carre A, Szinnai G, Castanet M, Sura-Trueba S, Tron E, Broutin-L'Hermite I, Barat P, Goizet C, Lacombe D, Moutard ML, Raybaud C, Raynaud-Ravni C, Romana S, Ythier H, Leger J, Polak M. Five new TTF1/NKX2.1 mutations in brain-lung-thyroid syndrome: rescue by PAX8 synergism in one case. Hum Mol Genet 2009;18:2266-76.
61. Gras D, Jonard L, Roze E, Chantot-Bastaraud S, Koht J, Motte J, Rodriguez D, Louha M, Caubel I, Kemlin I, Lion-Francois L, Goizet C, Guillot L, Moutard ML, Epaud R, Heron B, Charles P, Tallot M, Camuzat A, Durr A, Polak M, Devos D, Sanlaville D, Vuillaume I, Billette de Villemeur T, Vidailhet M, Doummar D. Benign hereditary chorea: phenotype, prognosis, therapeutic outcome and long term follow-up in a large

series with new mutations in the TITF1/NKX2-1 gene. J Neurol Neurosurg Psychiatry 2012;83:956-62.
62. Veneziano L, Parkinson MH, Mantuano E, Frontali M, Bhatia KP, Giunti P. A novel de novo mutation of the TITF1/NKX2-1 gene causing ataxia, benign hereditary chorea, hypothyroidism and a pituitary mass in a UK family and review of the literature. Cerebellum 2014;13:588-95.
63. Peall KJ, Lumsden D, Kneen R, Madhu R, Peake D, Gibbon F, Lewis H, Hedderly T, Meyer E, Robb SA, Lynch B, King MD, Lin JP, Morris HR, Jungbluth H, Kurian MA. Benign hereditary chorea related to NKX2.1: expansion of the genotypic and phenotypic spectrum. Dev Med Child Neurol 2014;56:642-8.
64. Nettore IC, Mirra P, Ferrara AM, Sibilio A, Pagliara V, Kay CS, Lorenzoni PJ, Werneck LC, Bruck I, Dos Santos LH, Beguinot F, Salvatore D, Ungaro P, Fenzi G, Scola RH, Macchia PE. Identification and functional characterization of a novel mutation in the NKX2-1 gene: comparison with the data in the literature. Thyroid 2013;23:675-82.
65. Kleiner-Fisman G, Rogaeva E, Halliday W, Houle S, Kawarai T, Sato C, Medeiros H, St George-Hyslop PH, Lang AE. Benign hereditary chorea: clinical, genetic, and pathological findings. Ann Neurol 2003;54:244-7.
66. Monti S, Nicoletti A, Cantasano A, Krude H, Cassio A. NKX2.1-Related Disorders: a novel mutation with mild clinical presentation. Ital J Pediatr 2015;41:45.
67. Gentile M, De Mattia D, Pansini A, Schettini F, Buonadonna AL, Capozza M, Ficarella R, Laforgia N. 14q13 distal microdeletion encompassing NKX2-1 and PAX9: Patient report and refinement of the associated phenotype. Am J Med Genet A 2016;170:1884-8.
68. Devos D, Vuillaume I, de Becdelievre A, de Martinville B, Dhaenens CM, Cuvellier JC, Cuisset JM, Vallee L, Lemaitre MP, Bourteel H, Hachulla E, Wallaert B, Destee A, Defebvre L, Sablonniere B. New syndromic form of benign hereditary chorea is associated with a deletion of TITF-1 and PAX-9 contiguous genes. Mov Disord 2006;21:2237-40.
69. Guala A, Falco V, Breedveld G, De Filippi P, Danesino C. Deletion of PAX9 and oligodontia: a third family and review of the literature. Int J Paediatr Dent 2008;18:441-5.
70. Zaborszky L vdPA, Gyengesi E. Chapter 28 - The Basal Forebrain Cholinergic Projection System in Mice, In The Mouse Nervous System. San Diego: edited by Charles Watson George Paxinos Luis Puelles, Academic Press; 2012.
71. Reil J. Untersuchungen über den Bau des großen Gehirns im Menschen. Archiv für die Physiologie, Halle: Curtsche Buchhandlung 1809;Vol. 9:136-208.
72. de Olmos JS, Heimer L. The concepts of the ventral striatopallidal system and extended amygdala. Ann N Y Acad Sci 1999;877:1-32.
73. Auld DS, Kornecook TJ, Bastianetto S, Quirion R. Alzheimer's disease and the basal forebrain cholinergic system: relations to beta-amyloid peptides, cognition, and treatment strategies. Prog Neurobiol 2002;68:209-45.
74. Hedreen JC, Struble RG, Whitehouse PJ, Price DL. Topography of the magnocellular basal forebrain system in human brain. J Neuropathol Exp Neurol 1984;43:1-21.
75. Meynert T. Vom Gehirne der Säugethiere. In: Stricker S e, ed. Handbuch der Lehre von den Geweben des Menschen und der Thiere. Leipzig: Engelmann; 1872:694-808.
76. Mufson EJ, Ginsberg SD, Ikonomovic MD, DeKosky ST. Human cholinergic basal forebrain: chemoanatomy and neurologic dysfunction. J Chem Neuroanat 2003;26:233-42.
77. Gritti I, Mainville L, Mancia M, Jones BE. GABAergic and other noncholinergic basal forebrain neurons, together with cholinergic neurons, project to the mesocortex and isocortex in the rat. J Comp Neurol 1997;383:163-77.

78. Zaborszky L, Pang K, Somogyi J, Nadasdy Z, Kallo I. The basal forebrain corticopetal system revisited. Ann N Y Acad Sci 1999;877:339-67.
79. Gritti I, Manns ID, Mainville L, Jones BE. Parvalbumin, calbindin, or calretinin in cortically projecting and GABAergic, cholinergic, or glutamatergic basal forebrain neurons of the rat. J Comp Neurol 2003;458:11-31.
80. Henderson Z, Lu CB, Janzso G, Matto N, McKinley CE, Yanagawa Y, Halasy K. Distribution and role of Kv3.1b in neurons in the medial septum diagonal band complex. Neuroscience 2010;166:952-69.
81. Hur EE, Zaborszky L. Vglut2 afferents to the medial prefrontal and primary somatosensory cortices: a combined retrograde tracing in situ hybridization study [corrected]. J Comp Neurol 2005;483:351-73.
82. Shimamura K, Hartigan DJ, Martinez S, Puelles L, Rubenstein JL. Longitudinal organization of the anterior neural plate and neural tube. Development 1995;121:3923-33.
83. Flames N, Pla R, Gelman DM, Rubenstein JL, Puelles L, Marin O. Delineation of multiple subpallial progenitor domains by the combinatorial expression of transcriptional codes. J Neurosci 2007;27:9682-95.
84. Garcia-Lopez M, Abellan A, Legaz I, Rubenstein JL, Puelles L, Medina L. Histogenetic compartments of the mouse centromedial and extended amygdala based on gene expression patterns during development. J Comp Neurol 2008;506:46-74.
85. Xu Q, Tam M, Anderson SA. Fate mapping Nkx2.1-lineage cells in the mouse telencephalon. J Comp Neurol 2008;506:16-29.
86. Flandin P, Kimura S, Rubenstein JL. The progenitor zone of the ventral medial ganglionic eminence requires Nkx2-1 to generate most of the globus pallidus but few neocortical interneurons. J Neurosci 2010;30:2812-23.
87. Rubin AN, Alfonsi F, Humphreys MP, Choi CK, Rocha SF, Kessaris N. The germinal zones of the basal ganglia but not the septum generate GABAergic interneurons for the cortex. J Neurosci 2010;30:12050-62.
88. Butt SJ, Sousa VH, Fuccillo MV, Hjerling-Leffler J, Miyoshi G, Kimura S, Fishell G. The requirement of Nkx2-1 in the temporal specification of cortical interneuron subtypes. Neuron 2008;59:722-32.
89. Sweeney JE, Hohmann CF, Oster-Granite ML, Coyle JT. Neurogenesis of the basal forebrain in euploid and trisomy 16 mice: an animal model for developmental disorders in Down syndrome. Neuroscience 1989;31:413-25.
90. Furusho M, Ono K, Takebayashi H, Masahira N, Kagawa T, Ikeda K, Ikenaka K. Involvement of the Olig2 transcription factor in cholinergic neuron development of the basal forebrain. Dev Biol 2006;293:348-57.
91. Marin O, Anderson SA, Rubenstein JL. Origin and molecular specification of striatal interneurons. J Neurosci 2000;20:6063-76.
92. Pleasure SJ, Anderson S, Hevner R, Bagri A, Marin O, Lowenstein DH, Rubenstein JL. Cell migration from the ganglionic eminences is required for the development of hippocampal GABAergic interneurons. Neuron 2000;28:727-40.
93. Lavdas AA, Grigoriou M, Pachnis V, Parnavelas JG. The medial ganglionic eminence gives rise to a population of early neurons in the developing cerebral cortex. J Neurosci 1999;19:7881-8.
94. Wichterle H, Turnbull DH, Nery S, Fishell G, Alvarez-Buylla A. In utero fate mapping reveals distinct migratory pathways and fates of neurons born in the mammalian basal forebrain. Development 2001;128:3759-71.
95. Wonders CP, Taylor L, Welagen J, Mbata IC, Xiang JZ, Anderson SA. A spatial bias for the origins of interneuron subgroups within the medial ganglionic eminence. Dev Biol 2008;314:127-36.

96. Xu Q, Cobos I, De La Cruz E, Rubenstein JL, Anderson SA. Origins of cortical interneuron subtypes. J Neurosci 2004;24:2612-22.
97. Magno L. The integrity of basal forebrain neurons depends on permanent expression of Nkx2-1: potential for understanding haploinsufficiency in humans. Berlin, Germany: Medizinische Fakultät der Charité- Universitätsmedizin Berlin; 2010.
98. Young KM, Fogarty M, Kessaris N, Richardson WD. Subventricular zone stem cells are heterogeneous with respect to their embryonic origins and neurogenic fates in the adult olfactory bulb. J Neurosci 2007;27:8286-96.
99. Bender R, Plaschke M, Naumann T, Wahle P, Frotscher M. Development of cholinergic and GABAergic neurons in the rat medial septum: different onset of choline acetyltransferase and glutamate decarboxylase mRNA expression. J Comp Neurol 1996;372:204-14.
100. Katarova Z, Sekerkova G, Prodan S, Mugnaini E, Szabo G. Domain-restricted expression of two glutamic acid decarboxylase genes in midgestation mouse embryos. J Comp Neurol 2000;424:607-27.
101. Puelles L, Kuwana E, Puelles E, Bulfone A, Shimamura K, Keleher J, Smiga S, Rubenstein JL. Pallial and subpallial derivatives in the embryonic chick and mouse telencephalon, traced by the expression of the genes Dlx-2, Emx-1, Nkx-2.1, Pax-6, and Tbr-1. J Comp Neurol 2000;424:409-38.
102. Korsching S. The neurotrophic factor concept: a reexamination. J Neurosci 1993;13:2739-48.
103. Naumann T, Casademunt E, Hollerbach E, Hofmann J, Dechant G, Frotscher M, Barde YA. Complete deletion of the neurotrophin receptor p75NTR leads to long-lasting increases in the number of basal forebrain cholinergic neurons. J Neurosci 2002;22:2409-18.
104. Fragkouli A, van Wijk NV, Lopes R, Kessaris N, Pachnis V. LIM homeodomain transcription factor-dependent specification of bipotential MGE progenitors into cholinergic and GABAergic striatal interneurons. Development 2009;136:3841-51.
105. Flames N, Marin O. Developmental mechanisms underlying the generation of cortical interneuron diversity. Neuron 2005;46:377-81.
106. Liodis P, Denaxa M, Grigoriou M, Akufo-Addo C, Yanagawa Y, Pachnis V. Lhx6 activity is required for the normal migration and specification of cortical interneuron subtypes. J Neurosci 2007;27:3078-89.
107. Cho HH, Cargnin F, Kim Y, Lee B, Kwon RJ, Nam H, Shen R, Barnes AP, Lee JW, Lee S, Lee SK. Isl1 directly controls a cholinergic neuronal identity in the developing forebrain and spinal cord by forming cell type-specific complexes. PLoS Genet 2014;10:e1004280.
108. Tomioka T, Shimazaki T, Yamauchi T, Oki T, Ohgoh M, Okano H. LIM homeobox 8 (Lhx8) is a key regulator of the cholinergic neuronal function via a tropomyosin receptor kinase A (TrkA)-mediated positive feedback loop. J Biol Chem 2014;289:1000-10.
109. Li H, Qin J, Jin G, Zou L, Shi J, Han X, Cheng X, Zhang X. Overexpression of Lhx8 inhibits cell proliferation and induces cell cycle arrest in PC12 cell line. In Vitro Cell Dev Biol Anim 2015;51:329-35.
110. Zhao Y, Marin O, Hermesz E, Powell A, Flames N, Palkovits M, Rubenstein JL, Westphal H. The LIM-homeobox gene Lhx8 is required for the development of many cholinergic neurons in the mouse forebrain. Proc Natl Acad Sci U S A 2003;100:9005-10.
111. Mori T, Yuxing Z, Takaki H, Takeuchi M, Iseki K, Hagino S, Kitanaka J, Takemura M, Misawa H, Ikawa M, Okabe M, Wanaka A. The LIM homeobox gene, L3/Lhx8, is necessary for proper development of basal forebrain cholinergic neurons. Eur J Neurosci 2004;19:3129-41.

112. Elshatory Y, Gan L. The LIM-homeobox gene Islet-1 is required for the development of restricted forebrain cholinergic neurons. J Neurosci 2008;28:3291-7.
113. Manabe T, Tatsumi K, Inoue M, Matsuyoshi H, Makinodan M, Yokoyama S, Wanaka A. L3/Lhx8 is involved in the determination of cholinergic or GABAergic cell fate. J Neurochem 2005;94:723-30.
114. Lopes R, Verhey van Wijk N, Neves G, Pachnis V. Transcription factor LIM homeobox 7 (Lhx7) maintains subtype identity of cholinergic interneurons in the mammalian striatum. Proc Natl Acad Sci U S A 2012;109:3119-24.
115. Colom LV. Septal networks: relevance to theta rhythm, epilepsy and Alzheimer's disease. J Neurochem 2006;96:609-23.
116. Park KJ, Grosso CA, Aubert I, Kaplan DR, Miller FD. p75NTR-dependent, myelin-mediated axonal degeneration regulates neural connectivity in the adult brain. Nat Neurosci 2010;13:559-66.
117. Swanson LW KC, Björklund A. The limbic region I: The septohippocampal system. In: A. Björklund TH, and L.W. Swanson, eds, ed. Handbook of Chemical Neuroanatomy ntegrated Systems of the CNS, Part1. Amsterdam: Elsevier; 1987:125-277.
118. Stephan H. Allocortex. Handbuch der mikroskopischen Anatomie des Menschen. Berlin, Heidelberg, New York: Springer-Verlag; 1975.
119. Swanson LW, Cowan WM. The connections of the septal region in the rat. J Comp Neurol 1979;186:621-55.
120. Amaral DG, Kurz J. An analysis of the origins of the cholinergic and noncholinergic septal projections to the hippocampal formation of the rat. J Comp Neurol 1985;240:37-59.
121. Freund TF, Antal M. GABA-containing neurons in the septum control inhibitory interneurons in the hippocampus. Nature 1988;336:170-3.
122. Freund TF. GABAergic septohippocampal neurons contain parvalbumin. Brain Res 1989;478:375-81.
123. Linke R, Frotscher M. Development of the rat septohippocampal projection: tracing with DiI and electron microscopy of identified growth cones. J Comp Neurol 1993;332:69-88.
124. Naumann T, Linke R, Frotscher M. Fine structure of rat septohippocampal neurons: I. Identification of septohippocampal projection neurons by retrograde tracing combined with electron microscopic immunocytochemistry and intracellular staining. J Comp Neurol 1992;325:207-18.
125. Naumann T, Schnell O, Zhi Q, Kirsch M, Schubert KO, Sendtner M, Hofmann HD. Endogenous ciliary neurotrophic factor protects GABAergic, but not cholinergic, septohippocampal neurons following fimbria-fornix transection. Brain Pathol 2003;13:309-21.
126. Sofroniew MV, Galletly NP, Isacson O, Svendsen CN. Survival of adult basal forebrain cholinergic neurons after loss of target neurons. Science 1990;247:338-42.
127. Jakab RL LC. Septum. In: G.Paxinos e, ed. The rat nervous system. Sydney, New York, London: Academic Press; 1995:405-42.
128. Takei N, Kuramoto H, Endo Y, Hatanaka H. NGF and BDNF increase the immunoreactivity of vesicular acetylcholine transporter in cultured neurons from the embryonic rat septum. Neurosci Lett 1997;226:207-9.
129. Oosawa H, Fujii T, Kawashima K. Nerve growth factor increases the synthesis and release of acetylcholine and the expression of vesicular acetylcholine transporter in primary cultured rat embryonic septal cells. J Neurosci Res 1999;57:381-7.
130. Auld DS, Mennicken F, Day JC, Quirion R. Neurotrophins differentially enhance acetylcholine release, acetylcholine content and choline acetyltransferase activity in basal forebrain neurons. J Neurochem 2001;77:253-62.

131. Cohen S, Levi-Montalcini R, Hamburger V. A Nerve Growth-Stimulating Factor Isolated from Sarcom as 37 and 180. Proc Natl Acad Sci U S A 1954;40:1014-8.
132. Skaper SD. The biology of neurotrophins, signalling pathways, and functional peptide mimetics of neurotrophins and their receptors. CNS Neurol Disord Drug Targets 2008;7:46-62.
133. Bartkowska K, Turlejski K, Djavadian RL. Neurotrophins and their receptors in early development of the mammalian nervous system. Acta Neurobiol Exp (Wars) 2010;70:454-67.
134. Thoenen H. Neurotrophins and neuronal plasticity. Science 1995;270:593-8.
135. Thoenen H, Barde YA. Physiology of nerve growth factor. Physiol Rev 1980;60:1284-335.
136. Altar CA, Burton LE, Bennett GL, Dugich-Djordjevic M. Recombinant human nerve growth factor is biologically active and labels novel high-affinity binding sites in rat brain. Proc Natl Acad Sci U S A 1991;88:281-5.
137. Hempstead BL, Martin-Zanca D, Kaplan DR, Parada LF, Chao MV. High-affinity NGF binding requires coexpression of the trk proto-oncogene and the low-affinity NGF receptor. Nature 1991;350:678-83.
138. Kaplan DR, Hempstead BL, Martin-Zanca D, Chao MV, Parada LF. The trk proto-oncogene product: a signal transducing receptor for nerve growth factor. Science 1991;252:554-8.
139. Johnson D, Lanahan A, Buck CR, Sehgal A, Morgan C, Mercer E, Bothwell M, Chao M. Expression and structure of the human NGF receptor. Cell 1986;47:545-54.
140. Radeke MJ, Misko TP, Hsu C, Herzenberg LA, Shooter EM. Gene transfer and molecular cloning of the rat nerve growth factor receptor. Nature 1987;325:593-7.
141. Chao MV, Hempstead BL. p75 and Trk: a two-receptor system. Trends Neurosci 1995;18:321-6.
142. Bothwell M. p75NTR: a receptor after all. Science 1996;272:506-7.
143. Dechant G, Barde YA. The neurotrophin receptor p75(NTR): novel functions and implications for diseases of the nervous system. Nat Neurosci 2002;5:1131-6.
144. Sobreviela T, Clary DO, Reichardt LF, Brandabur MM, Kordower JH, Mufson EJ. TrkA-immunoreactive profiles in the central nervous system: colocalization with neurons containing p75 nerve growth factor receptor, choline acetyltransferase, and serotonin. J Comp Neurol 1994;350:587-611.
145. Batchelor PE, Armstrong DM, Blaker SN, Gage FH. Nerve growth factor receptor and choline acetyltransferase colocalization in neurons within the rat forebrain: response to fimbria-fornix transection. J Comp Neurol 1989;284:187-204.
146. Gibbs RB, Pfaff DW. In situ hybridization detection of trkA mRNA in brain: distribution, colocalization with p75NGFR and up-regulation by nerve growth factor. J Comp Neurol 1994;341:324-39.
147. Esposito D, Patel P, Stephens RM, Perez P, Chao MV, Kaplan DR, Hempstead BL. The cytoplasmic and transmembrane domains of the p75 and Trk A receptors regulate high affinity binding to nerve growth factor. J Biol Chem 2001;276:32687-95.
148. Benedetti M, Levi A, Chao MV. Differential expression of nerve growth factor receptors leads to altered binding affinity and neurotrophin responsiveness. Proc Natl Acad Sci U S A 1993;90:7859-63.
149. Bibel M, Barde YA. Neurotrophins: key regulators of cell fate and cell shape in the vertebrate nervous system. Genes Dev 2000;14:2919-37.
150. Davies AM, Lee KF, Jaenisch R. p75-deficient trigeminal sensory neurons have an altered response to NGF but not to other neurotrophins. Neuron 1993;11:565-74.

151. Mahadeo D, Kaplan L, Chao MV, Hempstead BL. High affinity nerve growth factor binding displays a faster rate of association than p140trk binding. Implications for multi-subunit polypeptide receptors. J Biol Chem 1994;269:6884-91.
152. Bentley CA, Lee KF. p75 is important for axon growth and schwann cell migration during development. J Neurosci 2000;20:7706-15.
153. Harrison SM, Jones ME, Uecker S, Albers KM, Kudrycki KE, Davis BM. Levels of nerve growth factor and neurotrophin-3 are affected differentially by the presence of p75 in sympathetic neurons in vivo. J Comp Neurol 2000;424:99-110.
154. Schnitzler AC, Mellott TJ, Lopez-Coviella I, Tallini YN, Kotlikoff MI, Follettie MT, Blusztajn JK. BMP9 (bone morphogenetic protein 9) induces NGF as an autocrine/paracrine cholinergic trophic factor in developing basal forebrain neurons. J Neurosci 2010;30:8221-8.
155. Gnahn H, Hefti F, Heumann R, Schwab ME, Thoenen H. NGF-mediated increase of choline acetyltransferase (ChAT) in the neonatal rat forebrain: evidence for a physiological role of NGF in the brain? Brain Res 1983;285:45-52.
156. Mobley WC, Rutkowski JL, Tennekoon GI, Gemski J, Buchanan K, Johnston MV. Nerve growth factor increases choline acetyltransferase activity in developing basal forebrain neurons. Brain Res 1986;387:53-62.
157. Sofroniew MV, Pearson RC, Powell TP. The cholinergic nuclei of the basal forebrain of the rat: normal structure, development and experimentally induced degeneration. Brain Res 1987;411:310-31.
158. Levi-Montalcini R, Hamburger V. Selective growth stimulating effects of mouse sarcoma on the sensory and sympathetic nervous system of the chick embryo. J Exp Zool 1951;116:321-61.
159. Whittemore SR, Seiger A. The expression, localization and functional significance of beta-nerve growth factor in the central nervous system. Brain Res 1987;434:439-64.
160. Shooter EM. Early days of the nerve growth factor proteins. Annu Rev Neurosci 2001;24:601-29.
161. Huang EJ, Reichardt LF. Neurotrophins: roles in neuronal development and function. Annu Rev Neurosci 2001;24:677-736.
162. Levi-Montalcini R. The nerve growth factor 35 years later. Science 1987;237:1154-62.
163. Hempstead BL. Brain-Derived Neurotrophic Factor: Three Ligands, Many Actions. Trans Am Clin Climatol Assoc 2015;126:9-19.
164. Reichardt LF. Neurotrophin-regulated signalling pathways. Philos Trans R Soc Lond B Biol Sci 2006;361:1545-64.
165. Miranda RC, Sohrabji F, Toran-Allerand CD. Neuronal colocalization of mRNAs for neurotrophins and their receptors in the developing central nervous system suggests a potential for autocrine interactions. Proc Natl Acad Sci U S A 1993;90:6439-43.
166. Klein R, Smeyne RJ, Wurst W, Long LK, Auerbach BA, Joyner AL, Barbacid M. Targeted disruption of the trkB neurotrophin receptor gene results in nervous system lesions and neonatal death. Cell 1993;75:113-22.
167. Ernfors P, Kucera J, Lee KF, Loring J, Jaenisch R. Studies on the physiological role of brain-derived neurotrophic factor and neurotrophin-3 in knockout mice. Int J Dev Biol 1995;39:799-807.
168. Smeyne RJ, Klein R, Schnapp A, Long LK, Bryant S, Lewin A, Lira SA, Barbacid M. Severe sensory and sympathetic neuropathies in mice carrying a disrupted Trk/NGF receptor gene. Nature 1994;368:246-9.
169. Fagan AM, Garber M, Barbacid M, Silos-Santiago I, Holtzman DM. A role for TrkA during maturation of striatal and basal forebrain cholinergic neurons in vivo. J Neurosci 1997;17:7644-54.

170. Van der Zee CE, Ross GM, Riopelle RJ, Hagg T. Survival of cholinergic forebrain neurons in developing p75NGFR-deficient mice. Science 1996;274:1729-32.
171. Peterson DA, Leppert JT, Lee KF, Gage FH. Basal forebrain neuronal loss in mice lacking neurotrophin receptor p75. Science 1997;277:837-9.
172. Yeo TT, Chua-Couzens J, Butcher LL, Bredesen DE, Cooper JD, Valletta JS, Mobley WC, Longo FM. Absence of p75NTR causes increased basal forebrain cholinergic neuron size, choline acetyltransferase activity, and target innervation. J Neurosci 1997;17:7594-605.
173. Hagg T. Neuronal cell death: retraction. Science 1999;285:340.
174. Voigt A. Generation and analysis of transgenic mice expressing CRE recom-binase in defined interneurons: University Heidelberg; 2007.
175. Kusakabe T, Kawaguchi A, Hoshi N, Kawaguchi R, Hoshi S, Kimura S. Thyroid-specific enhancer-binding protein/NKX2.1 is required for the maintenance of ordered architecture and function of the differentiated thyroid. Mol Endocrinol 2006;20:1796-809.
176. Soriano P. Generalized lacZ expression with the ROSA26 Cre reporter strain. Nat Genet 1999;21:70-1.
177. Paulick K. Die Funktionen der Neureguline im Kortex und Muskel: Fachbereich Biologie, Chemie, Pharmazie der Freien Universität Berlin; 2013.
178. den Dunnen JT, Dalgleish R, Maglott DR, Hart RK, Greenblatt MS, McGowan-Jordan J, Roux AF, Smith T, Antonarakis SE, Taschner PE. HGVS Recommendations for the Description of Sequence Variants: 2016 Update. Hum Mutat 2016;37:564-9.
179. Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, Thomas NS, Abeysinghe S, Krawczak M, Cooper DN. Human Gene Mutation Database (HGMD): 2003 update. Hum Mutat 2003;21:577-81.
180. Wildeman M, van Ophuizen E, den Dunnen JT, Taschner PE. Improving sequence variant descriptions in mutation databases and literature using the Mutalyzer sequence variation nomenclature checker. Hum Mutat 2008;29:6-13.
181. Kamnasaran D, O'Brien PC, Schuffenhauer S, Quarrell O, Lupski JR, Grammatico P, Ferguson-Smith MA, Cox DW. Defining the breakpoints of proximal chromosome 14q rearrangements in nine patients using flow-sorted chromosomes. Am J Med Genet 2001;102:173-82.
182. Salvado M, Boronat-Guerrero S, Hernandez-Vara J, Alvarez-Sabin J. [Chorea due to TITF1/NKX2-1 mutation: phenotypical description and therapeutic response in a family]. Rev Neurol 2013;56:515-20.
183. Petek E, Plecko-Startinig B, Windpassinger C, Egger H, Wagner K, Kroisel PM. Molecular characterisation of a 3.5 Mb interstitial 14q deletion in a child with several phenotypic anomalies. J Med Genet 2003;40:e47.
184. Santen GW, Sun Y, Gijsbers AC, Carre A, Holvoet M, Haeringen A, Lesnik Oberstein SA, Tomoda A, Mabe H, Polak M, Devriendt K, Ruivenkamp CA, Bijlsma EK. Further delineation of the phenotype of chromosome 14q13 deletions: (positional) involvement of FOXG1 appears the main determinant of phenotype severity, with no evidence for a holoprosencephaly locus. J Med Genet 2012;49:366-72.
185. Barnett CP, Mencel JJ, Gecz J, Waters W, Kirwin SM, Vinette KM, Uppill M, Nicholl J. Choreoathetosis, congenital hypothyroidism and neonatal respiratory distress syndrome with intact NKX2-1. Am J Med Genet A 2012;158A:3168-73.
186. Kharbanda M, Hermanns P, Jones J, Pohlenz J, Horrocks I, Donaldson M. A further case of brain-lung-thyroid syndrome with deletion proximal to NKX2-1. Eur J Med Genet 2017;60:257-60.
187. Wang C, Hata Y, Hirono K, Takasaki A, Ozawa SW, Nakaoka H, Saito K, Miyao N, Okabe M, Ibuki K, Nishida N, Origasa H, Yu X, Bowles NE, Ichida F, for LSC. A Wide

and Specific Spectrum of Genetic Variants and Genotype-Phenotype Correlations Revealed by Next-Generation Sequencing in Patients with Left Ventricular Noncompaction. J Am Heart Assoc 2017;6.
188. Breedveld GJ, Percy AK, MacDonald ME, de Vries BB, Yapijakis C, Dure LS, Ippel EF, Sandkuijl LA, Heutink P, Arts WF. Clinical and genetic heterogeneity in benign hereditary chorea. Neurology 2002;59:579-84.
189. Iwatani N, Mabe H, Devriendt K, Kodama M, Miike T. Deletion of NKX2.1 gene encoding thyroid transcription factor-1 in two siblings with hypothyroidism and respiratory failure. J Pediatr 2000;137:272-6.
190. Devriendt K, Vanhole C, Matthijs G, de Zegher F. Deletion of thyroid transcription factor-1 gene in an infant with neonatal thyroid dysfunction and respiratory failure. N Engl J Med 1998;338:1317-8.
191. Nevel RJ, Garnett ET, Worrell JA, Morton RL, Nogee LM, Blackwell TS, Young LR. Persistent Lung Disease in Adults with NKX2.1 Mutation and Familial Neuroendocrine Cell Hyperplasia of Infancy. Ann Am Thorac Soc 2016;13:1299-304.
192. Young LR, Deutsch GH, Bokulic RE, Brody AS, Nogee LM. A mutation in TTF1/NKX2.1 is associated with familial neuroendocrine cell hyperplasia of infancy. Chest 2013;144:1199-206.
193. Koht J, Lostegaard SO, Wedding I, Vidailhet M, Louha M, Tallaksen CM. Benign hereditary chorea, not only chorea: a family case presentation. Cerebellum Ataxias 2016;3:3.
194. Ferrara AM, De Michele G, Salvatore E, Di Maio L, Zampella E, Capuano S, Del Prete G, Rossi G, Fenzi G, Filla A, Macchia PE. A novel NKX2.1 mutation in a family with hypothyroidism and benign hereditary chorea. Thyroid 2008;18:1005-9.
195. Salvatore E, Di Maio L, Filla A, Ferrara AM, Rinaldi C, Sacca F, Peluso S, Macchia PE, Pappata S, De Michele G. Benign hereditary chorea: clinical and neuroimaging features in an Italian family. Mov Disord 2010;25:1491-6.
196. Guillot L, Carre A, Szinnai G, Castanet M, Tron E, Jaubert F, Broutin I, Counil F, Feldmann D, Clement A, Polak M, Epaud R. NKX2-1 mutations leading to surfactant protein promoter dysregulation cause interstitial lung disease in "Brain-Lung-Thyroid Syndrome". Hum Mutat 2010;31:E1146-62.
197. Attarian SJ, Leibel SL, Yang P, Alfano DN, Hackett BP, Cole FS, Hamvas A. Mutations in the thyroid transcription factor gene NKX2-1 result in decreased expression of SFTPB and SFTPC. Pediatr Res 2018.
198. Segel R, Ben-Pazi H, Zeligson S, Fatal-Valevski A, Aran A, Gross-Tsur V, Schneebaum-Sender N, Shmueli D, Lev D, Perlberg S, Blumkin L, Deutsch L, Levy-Lahad E. Copy number variations in cryptogenic cerebral palsy. Neurology 2015;84:1660-8.
199. Dale RC, Grattan-Smith P, Nicholson M, Peters GB. Microdeletions detected using chromosome microarray in children with suspected genetic movement disorders: a single-centre study. Dev Med Child Neurol 2012;54:618-23.
200. Bodian DL, Klein E, Iyer RK, Wong WS, Kothiyal P, Stauffer D, Huddleston KC, Gaither AD, Remsburg I, Khromykh A, Baker RL, Maxwell GL, Vockley JG, Niederhuber JE, Solomon BD. Utility of whole-genome sequencing for detection of newborn screening disorders in a population cohort of 1,696 neonates. Genet Med 2016;18:221-30.
201. Iodice A, Pisani F. Status dystonicus: management and prevention in children at high risk. Acta Biomed 2019;90:207-12.
202. de Filippis T, Marelli F, Vigone MC, Di Frenna M, Weber G, Persani L. Novel NKX2-1 Frameshift Mutations in Patients with Atypical Phenotypes of the Brain-Lung-Thyroid Syndrome. Eur Thyroid J 2014;3:227-33.

203. Wang F, Liu C, Jia X, Liu X, Xu Y, Yan S, Jia X, Huang Z, Liu S, Gu M. Next-generation sequencing of NKX2.1, FOXE1, PAX8, NKX2.5, and TSHR in 100 Chinese patients with congenital hypothyroidism and athyreosis. Clin Chim Acta 2017;470:36-41.
204. Narumi S, Muroya K, Asakura Y, Adachi M, Hasegawa T. Transcription factor mutations and congenital hypothyroidism: systematic genetic screening of a population-based cohort of Japanese patients. J Clin Endocrinol Metab 2010;95:1981-5.
205. Zorzi G, Invernizzi F, Zibordi F, Costa C, Ciano C, Garavaglia B, Nardocci N. Clinical features of a new family with benign hereditary chorea carrying a novel TITF-1 mutation. Movement Disorders 2008;23:1-6.
206. Teissier R, Guillot L, Carre A, Morandini M, Stuckens C, Ythier H, Munnich A, Szinnai G, de Blic J, Clement A, Leger J, Castanet M, Epaud R, Polak M. Multiplex Ligation-dependent Probe Amplification improves the detection rate of NKX2.1 mutations in patients affected by brain-lung-thyroid syndrome. Horm Res Paediatr 2012;77:146-51.
207. Accornero S, Danesino C, Bastianello S, D'Errico I, Guala A, Chiovato L. Duplication of the pituitary stalk in a patient with a heterozygous deletion of chromosome 14 harboring the thyroid transcription factor-1 gene. J Clin Endocrinol Metab 2010;95:3595-6.
208. Hayashi S, Yagi M, Morisaki I, Inazawa J. Identical deletion at 14q13.3 including PAX9 and NKX2-1 in siblings from mosaicism of unaffected parent. J Hum Genet 2015;60:203-6.
209. Asmus F, Horber V, Pohlenz J, Schwabe D, Zimprich A, Munz M, Schoning M, Gasser T. A novel TITF-1 mutation causes benign hereditary chorea with response to levodopa. Neurology 2005;64:1952-4.
210. Glik A, Vuillaume I, Devos D, Inzelberg R. Psychosis, short stature in benign hereditary chorea: a novel thyroid transcription factor-1 mutation. Mov Disord 2008;23:1744-7.
211. Asmus F, Devlin A, Munz M, Zimprich A, Gasser T, Chinnery PF. Clinical differentiation of genetically proven benign hereditary chorea and myoclonus-dystonia. Mov Disord 2007;22:2104-9.
212. Doyle DA, Gonzalez I, Thomas B, Scavina M. Autosomal dominant transmission of congenital hypothyroidism, neonatal respiratory distress, and ataxia caused by a mutation of NKX2-1. J Pediatr 2004;145:190-3.
213. Safi KH, Bernat JA, Keegan CE, Ahmad A, Hershenson MB, Arteta M. Interstitial lung disease of infancy caused by a new NKX2-1 mutation. Clin Case Rep 2017;5:739-43.
214. Milone R, Masson R, Di Cosmo C, Tonacchera M, Bertini V, Guzzetta A, Battini R. A Not So Benign Family Pedigree With Hereditary Chorea: A Broader Phenotypic Expression or Additional Picture? Child Neurol Open 2019;6:2329048X19828881.
215. Blumkin L, Lerman-Sagie T, Westenberger A, Ben-Pazi H, Zerem A, Yosovich K, Lev D. Multiple Causes of Pediatric Early Onset Chorea-Clinical and Genetic Approach. Neuropediatrics 2018;49:246-55.
216. Lof C, Patyra K, Kuulasmaa T, Vangipurapu J, Undeutsch H, Jaeschke H, Pajunen T, Kero A, Krude H, Biebermann H, Kleinau G, Kuhnen P, Rantakari K, Miettinen P, Kirjavainen T, Pursiheimo JP, Mustila T, Jaaskelainen J, Ojaniemi M, Toppari J, Ignatius J, Laakso M, Kero J. Detection of Novel Gene Variants Associated with Congenital Hypothyroidism in a Finnish Patient Cohort. Thyroid 2016;26:1215-24.
217. Ferrara JM, Adam OR, Kirwin SM, Houghton DJ, Shepherd C, Vinette KM, Litvan I. Brain-lung-thyroid disease: clinical features of a kindred with a novel thyroid transcription factor 1 mutation. J Child Neurol 2012;27:68-73.
218. Provenzano C, Veneziano L, Appleton R, Frontali M, Civitareale D. Functional characterization of a novel mutation in TITF-1 in a patient with benign hereditary chorea. J Neurol Sci 2008;264:56-62.

219. Shinohara H, Takagi M, Ito K, Shimizu E, Fukuzawa R, Hasegawa T. A Novel Mutation in NKX2-1 Shows Dominant-Negative Effects Only in the Presence of PAX8. Thyroid 2018.
220. Coon EA, Ahlskog JE, Patterson MC, Niu Z, Milone M. Expanding Phenotypic Spectrum of NKX2-1-Related Disorders-Mitochondrial and Immunologic Dysfunction. JAMA Neurol 2016;73:237-8.
221. Costa MC, Costa C, Silva AP, Evangelista P, Santos L, Ferro A, Sequeiros J, Maciel P. Nonsense mutation in TITF1 in a Portuguese family with benign hereditary chorea. Neurogenetics 2005;6:209-15.
222. McMichael G, Haan E, Gardner A, Yap TY, Thompson S, Ouvrier R, Dale RC, Gecz J, Maclennan AH. NKX2-1 mutation in a family diagnosed with ataxic dyskinetic cerebral palsy. Eur J Med Genet 2013;56:506-9.
223. Sempere AP, Aparicio S, Mola S, Perez-Tur J. Benign hereditary chorea: clinical features and long-term follow-up in a Spanish family. Parkinsonism Relat Disord 2013;19:394-6.
224. Asmus F, Langseth A, Doherty E, Nestor T, Munz M, Gasser T, Lynch T, King MD. "Jerky" dystonia in children: spectrum of phenotypes and genetic testing. Mov Disord 2009;24:702-9.
225. Fons C, Rizzu P, Garcia-Cazorla A, Martorell L, Ormazabal A, Artuch R, Campistol J, Fernandez-Alvarez E. TITF-1 gene mutation in a case of sporadic non-progressive chorea. Response to levodopa treatment. Brain Dev 2012;34:255-7.
226. Jovien S, Borie R, Doummar D, Clement A, Nathan N. Respiratory Distress, Congenital Hypothyroidism and Hypotonia in a Newborn. Respiration 2016;92:188-91.
227. Barreiro J, Alonso-Fernandez JR, Castro-Feijoo L, Colon C, Cabanas P, Heredia C, Castano LA, Gomez-Lado C, Couce ML, Pombo M. Congenital hypothyroidism with neurological and respiratory alterations: a case detected using a variable diagnostic threshold for TSH. J Clin Res Pediatr Endocrinol 2011;3:208-11.
228. Konishi T, Kono S, Fujimoto M, Terada T, Matsushita K, Ouchi Y, Miyajima H. Benign hereditary chorea: dopaminergic brain imaging in patients with a novel intronic NKX2.1 gene mutation. J Neurol 2013;260:207-13.
229. Nagasaki K, Narumi S, Asami T, Kikuchi T, Hasegawa T, Uchiyama M. Mutation of a gene for thyroid transcription factor-1 (TITF1) in a patient with clinical features of resistance to thyrotropin. Endocr J 2008;55:875-8.
230. Moya CM, Perez de Nanclares G, Castano L, Potau N, Bilbao JR, Carrascosa A, Bargada M, Coya R, Martul P, Vicens-Calvet E, Santisteban P. Functional study of a novel single deletion in the TITF1/NKX2.1 homeobox gene that produces congenital hypothyroidism and benign chorea but not pulmonary distress. J Clin Endocrinol Metab 2006;91:1832-41.
231. Mahajnah M, Inbar D, Steinmetz A, Heutink P, Breedveld GJ, Straussberg R. Benign hereditary chorea: clinical, neuroimaging, and genetic findings. J Child Neurol 2007;22:1231-4.
232. Shetty VB, Kiraly-Borri C, Lamont P, Bikker H, Choong CS. NKX2-1 mutations in brain-lung-thyroid syndrome: a case series of four patients. J Pediatr Endocrinol Metab 2014;27:373-8.
233. Provenzano C, Zamboni M, Veneziano L, Mantuano E, Garavaglia B, Zorzi G, Pagonabarraga J, Giunti P, Civitareale D. Functional characterization of two novel mutations in TTF-1/NKX2.1 homeodomain in patients with benign hereditary chorea. J Neurol Sci 2016;360:78-83.
234. Gillett ES, Deutsch GH, Bamshad MJ, McAdams RM, Mann PC. Novel NKX2.1 mutation associated with hypothyroidism and lethal respiratory failure in a full-term neonate. J Perinatol 2013;33:157-60.

235. Tubing J, Bohnenpoll J, Spiegler J, Gillessen-Kaesbach G, Baumer T, Max C, Sperner J, Klein C, Munchau A. Methylphenidate Can Improve Chorea in NKX2.1 and ADCY5 Mutation-positive Patients-A Report of Two Children. Mov Disord Clin Pract 2018;5:343-5.
236. Parnes M, Bashir H, Jankovic J. Is Benign Hereditary Chorea Really Benign? Brain-Lung-Thyroid Syndrome Caused by NKX2-1 Mutations. Mov Disord Clin Pract 2019;6:34-9.
237. Rosati A, Berti B, Melani F, Cellini E, Procopio E, Guerrini R. Recurrent drop attacks in early childhood as presenting symptom of benign hereditary chorea caused by TITF1 gene mutations. Dev Med Child Neurol 2015;57:777-9.
238. Uematsu M, Haginoya K, Kikuchi A, Nakayama T, Kakisaka Y, Numata Y, Kobayashi T, Hino-Fukuyo N, Fujiwara I, Kure S. Hypoperfusion in caudate nuclei in patients with brain-lung-thyroid syndrome. J Neurol Sci 2012;315:77-81.
239. Villafuerte B, Natera-de-Benito D, Gonzalez A, Mori MA, Palomares M, Nevado J, Garcia-Minaur S, Lapunzina P, Gonzalez-Granado LI, Allende LM, Moreno JC. The Brain-Lung-Thyroid syndrome (BLTS): A novel deletion in chromosome 14q13.2-q21.1 expands the phenotype to humoral immunodeficiency. Eur J Med Genet 2018;61:393-8.
240. Graf S, Bösch N, Bachmann S, Zumsteg U, Heinimann K, Szinnai G. Familial brain-lung-thyroid syndrome due to a new NKX2-1 mutation p.Q172L causing disabling benign hereditary chorea. Horm Res Paediatr 2012;78:47-150.
241. Gauquelin L, Tran LT, Chouinard S, Bernard G. The Movement Disorder of Brain-Lung-Thyroid Syndrome Can be Responsive to Methylphenidate. Tremor Other Hyperkinet Mov (N Y) 2017;7:508.
242. Williamson S, Kirkpatrick M, Greene S, Goudie D. A novel mutation of NKX2-1 affecting 2 generations with hypothyroidism and choreoathetosis: part of the spectrum of brain-thyroid-lung syndrome. J Child Neurol 2014;29:666-9.
243. Balicza P, Grosz Z, Molnar V, Illes A, Csaban D, Gezsi A, Dezsi L, Zadori D, Vecsei L, Molnar MJ. NKX2-1 New Mutation Associated With Myoclonus, Dystonia, and Pituitary Involvement. Front Genet 2018;9:335.
244. de Gusmao CM, Kok F, Casella EB, Waugh JL. Benign hereditary chorea related to NKX2-1 with ataxia and dystonia. Neurol Genet 2016;2:e40.
245. Zou M, Alzahrani AS, Al-Odaib A, Alqahtani MA, Babiker O, Al-Rijjal RA, BinEssa HA, Kattan WE, Al-Enezi AF, Al Qarni A, Al-Faham MSA, Baitei EY, Alsagheir A, Meyer BF, Shi Y. Molecular Analysis of Congenital Hypothyroidism in Saudi Arabia: SLC26A7 Mutation Is a Novel Defect in Thyroid Dyshormonogenesis. J Clin Endocrinol Metab 2018;103:1889-98.
246. Santos-Silva R, Rosario M, Grangeia A, Costa C, Castro-Correia C, Alonso I, Leao M, Fontoura M. Genetic analyses in a cohort of Portuguese pediatric patients with congenital hypothyroidism. J Pediatr Endocrinol Metab 2019;32:1265-73.
247. Tozawa T, Yokochi K, Kono S, Konishi T, Yamamoto T, Nishimura A, Chiyonobu T, Morimoto M, Hosoi H. A Video Report of Brain-Lung-Thyroid Syndrome in a Japanese Female With a Novel Frameshift Mutation of the NKX2-1 Gene. Child Neurol Open 2016;3:2329048X16665012.
248. Sun F, Zhang JX, Yang CY, Gao GQ, Zhu WB, Han B, Zhang LL, Wan YY, Ye XP, Ma YR, Zhang MM, Yang L, Zhang QY, Liu W, Guo CC, Chen G, Zhao SX, Song KY, Song HD. The genetic characteristics of congenital hypothyroidism in China by comprehensive screening of 21 candidate genes. Eur J Endocrinol 2018;178:623-33.
249. Ngan ES, Lang BH, Liu T, Shum CK, So MT, Lau DK, Leon TY, Cherny SS, Tsai SY, Lo CY, Khoo US, Tam PK, Garcia-Barcelo MM. A germline mutation (A339V) in

thyroid transcription factor-1 (TITF-1/NKX2.1) in patients with multinodular goiter and papillary thyroid carcinoma. J Natl Cancer Inst 2009;101:162-75.
250. Galambos C, Levy H, Cannon CL, Vargas SO, Reid LM, Cleveland R, Lindeman R, deMello DE, Wert SE, Whitsett JA, Perez-Atayde AR, Kozakewich H. Pulmonary pathology in thyroid transcription factor-1 deficiency syndrome. Am J Respir Crit Care Med 2010;182:549-54.
251. Kleinlein B, Griese M, Liebisch G, Krude H, Lohse P, Aslanidis C, Schmitz G, Peters J, Holzinger A. Fatal neonatal respiratory failure in an infant with congenital hypothyroidism due to haploinsufficiency of the NKX2-1 gene: alteration of pulmonary surfactant homeostasis. Arch Dis Child Fetal Neonatal Ed 2011;96:F453-6.
252. Salerno T, Peca D, Menchini L, Schiavino A, Petreschi F, Occasi F, Cogo P, Danhaive O, Cutrera R. Respiratory insufficiency in a newborn with congenital hypothyroidism due to a new mutation of TTF-1/NKX2.1 gene. Pediatr Pulmonol 2014;49:E42-4.
253. Hermanns P, Kumorowicz-Czoch M, Grasberger H, Refetoff S, Pohlenz J. Novel Mutations in the NKX2.1 gene and the PAX8 gene in a Boy with Brain-Lung-Thyroid Syndrome. Exp Clin Endocrinol Diabetes 2017.
254. Gu R, Ye G, Zhou Y, Jiang Z. Combined mutations of NKX2-1 and surfactant protein C genes for refractory low oxyhemoglobin saturation and interstitial pneumonia: A case report. Medicine (Baltimore) 2020;99:e19650.
255. Moya CM, Garzón L, Luna C, R. S, Zaballos MA, Santisteban P, Gallego E, Moreno JC. TAZ does not rescue the lung promoter activity of a novel NKX2-1 mutation in a boy with severe lung emphysema. Horm Res 2012;78.
256. Maccabelli G, Pichiecchio A, Guala A, Ponzio M, Palesi F, Maranzana Rt D, Poloni GU, Bastianello S, Danesino C. Advanced magnetic resonance imaging in benign hereditary chorea: study of two familial cases. Mov Disord 2010;25:2670-4.
257. Costa RM, Martul EV, Reboredo JM, Cigarran S. Curvilinear bodies in hydroxychloroquine-induced renal phospholipidosis resembling Fabry disease. Clin Kidney J 2013;6:533-6.
258. Peca D, Petrini S, Tzialla C, Boldrini R, Morini F, Stronati M, Carnielli VP, Cogo PE, Danhaive O. Altered surfactant homeostasis and recurrent respiratory failure secondary to TTF-1 nuclear targeting defect. Respir Res 2011;12:115.
259. Gauquelin N, van den Bos KHW, Beche A, Krause FF, Lobato I, Lazar S, Rosenauer A, Van Aert S, Verbeeck J. Determining oxygen relaxations at an interface: A comparative study between transmission electron microscopy techniques. Ultramicroscopy 2017;181:178-90.
260. Liu Y, Michelson D, Clark R, Gold JA. Child Neurology: Siblings with infantile epilepsy and developmental delay: A circuitous path to genomic diagnosis. Neurology 2018;91:143-7.
261. Berger-Sweeney J. The cholinergic basal forebrain system during development and its influence on cognitive processes: important questions and potential answers. Neurosci Biobehav Rev 2003;27:401-11.
262. Stoykova A, Treichel D, Hallonet M, Gruss P. Pax6 modulates the dorsoventral patterning of the mammalian telencephalon. J Neurosci 2000;20:8042-50.
263. Hefti F. Nerve growth factor promotes survival of septal cholinergic neurons after fimbrial transections. J Neurosci 1986;6:2155-62.
264. Gu H, Long D, Song C, Li X. Recombinant human NGF-loaded microspheres promote survival of basal forebrain cholinergic neurons and improve memory impairments of spatial learning in the rat model of Alzheimer's disease with fimbria-fornix lesion. Neurosci Lett 2009;453:204-9.
265. Zhang H, Petit GH, Gaughwin PM, Hansen C, Ranganathan S, Zuo X, Smith R, Roybon L, Brundin P, Mobley WC, Li JY. NGF rescues hippocampal cholinergic neuronal

markers, restores neurogenesis, and improves the spatial working memory in a mouse model of Huntington's Disease. J Huntingtons Dis 2013;2:69-82.
266. Goedert M, Fine A, Hunt SP, Ullrich A. Nerve growth factor mRNA in peripheral and central rat tissues and in the human central nervous system: lesion effects in the rat brain and levels in Alzheimer's disease. Brain Res 1986;387:85-92.
267. Boissiere F, Faucheux B, Ruberg M, Agid Y, Hirsch EC. Decreased TrkA gene expression in cholinergic neurons of the striatum and basal forebrain of patients with Alzheimer's disease. Exp Neurol 1997;145:245-52.
268. Mufson EJ, Counts SE, Fahnestock M, Ginsberg SD. Cholinotrophic molecular substrates of mild cognitive impairment in the elderly. Curr Alzheimer Res 2007;4:340-50.
269. Cooper JD, Lindholm D, Sofroniew MV. Reduced transport of [125I]nerve growth factor by cholinergic neurons and down-regulated TrkA expression in the medial septum of aged rats. Neuroscience 1994;62:625-9.
270. Capsoni S, Tiveron C, Vignone D, Amato G, Cattaneo A. Dissecting the involvement of tropomyosin-related kinase A and p75 neurotrophin receptor signaling in NGF deficit-induced neurodegeneration. Proc Natl Acad Sci U S A 2010;107:12299-304.
271. Sanchez-Ortiz E, Yui D, Song D, Li Y, Rubenstein JL, Reichardt LF, Parada LF. TrkA gene ablation in basal forebrain results in dysfunction of the cholinergic circuitry. J Neurosci 2012;32:4065-79.
272. Spejo AB, Oliveira AL. Synaptic rearrangement following axonal injury: Old and new players. Neuropharmacology 2015;96:113-23.
273. Benowitz LI, Routtenberg A. GAP-43: an intrinsic determinant of neuronal development and plasticity. Trends Neurosci 1997;20:84-91.
274. Shen Y, Mani S, Donovan SL, Schwob JE, Meiri KF. Growth-associated protein-43 is required for commissural axon guidance in the developing vertebrate nervous system. J Neurosci 2002;22:239-47.
275. Haas CA, Hollerbach E, Deller T, Naumann T, Frotscher M. Up-regulation of growth-associated protein 43 mRNA in rat medial septum neurons axotomized by fimbria-fornix transection. Eur J Neurosci 2000;12:4233-42.
276. Aubert I, Poirier J, Gauthier S, Quirion R. Multiple cholinergic markers are unexpectedly not altered in the rat dentate gyrus following entorhinal cortex lesions. J Neurosci 1994;14:2476-84.
277. Diolaiti D, Bernardoni R, Trazzi S, Papa A, Porro A, Bono F, Herbert JM, Perini G, Della Valle G. Functional cooperation between TrkA and p75(NTR) accelerates neuronal differentiation by increased transcription of GAP-43 and p21(CIP/WAF) genes via ERK1/2 and AP-1 activities. Exp Cell Res 2007;313:2980-92.
278. Saragovi HU. Progression of age-associated cognitive impairment correlates with quantitative and qualitative loss of TrkA receptor protein in nucleus basalis and cortex. J Neurochem 2005;95:1472-80.
279. Wetmore C, Ernfors P, Persson H, Olson L. Localization of brain-derived neurotrophic factor mRNA to neurons in the brain by in situ hybridization. Exp Neurol 1990;109:141-52.
280. Lobo I. Same genetic mutation, different genetic disease phenotype. Nature Education 2008;1:64.
281. Rivas MA, Pirinen M, Conrad DF, Lek M, Tsang EK, Karczewski KJ, Maller JB, Kukurba KR, DeLuca DS, Fromer M, Ferreira PG, Smith KS, Zhang R, Zhao F, Banks E, Poplin R, Ruderfer DM, Purcell SM, Tukiainen T, Minikel EV, Stenson PD, Cooper DN, Huang KH, Sullivan TJ, Nedzel J, Consortium GT, Geuvadis C, Bustamante CD, Li JB, Daly MJ, Guigo R, Donnelly P, Ardlie K, Sammeth M, Dermitzakis ET, McCarthy MI, Montgomery SB, Lappalainen T, MacArthur DG. Human genomics.

Effect of predicted protein-truncating genetic variants on the human transcriptome. Science 2015;348:666-9.
282. DeBoever C, Tanigawa Y, Lindholm ME, McInnes G, Lavertu A, Ingelsson E, Chang C, Ashley EA, Bustamante CD, Daly MJ, Rivas MA. Medical relevance of protein-truncating variants across 337,205 individuals in the UK Biobank study. Nat Commun 2018;9:1612.
283. Bartha I, Rausell A, McLaren PJ, Mohammadi P, Tardaguila M, Chaturvedi N, Fellay J, Telenti A. The Characteristics of Heterozygous Protein Truncating Variants in the Human Genome. PLoS Comput Biol 2015;11:e1004647.
284. Park KS, Whitsett JA, Di Palma T, Hong JH, Yaffe MB, Zannini M. TAZ interacts with TTF-1 and regulates expression of surfactant protein-C. J Biol Chem 2004;279:17384-90.
285. Di Palma T, Conti A, de Cristofaro T, Scala S, Nitsch L, Zannini M. Identification of novel Pax8 targets in FRTL-5 thyroid cells by gene silencing and expression microarray analysis. PLoS One 2011;6:e25162.
286. Ohno M, Zannini M, Levy O, Carrasco N, di Lauro R. The paired-domain transcription factor Pax8 binds to the upstream enhancer of the rat sodium/iodide symporter gene and participates in both thyroid-specific and cyclic-AMP-dependent transcription. Mol Cell Biol 1999;19:2051-60.
287. Zannini M, Francis-Lang H, Plachov D, Di Lauro R. Pax-8, a paired domain-containing protein, binds to a sequence overlapping the recognition site of a homeodomain and activates transcription from two thyroid-specific promoters. Mol Cell Biol 1992;12:4230-41.
288. Mansouri A, Chowdhury K, Gruss P. Follicular cells of the thyroid gland require Pax8 gene function. Nat Genet 1998;19:87-90.
289. Whybra C, Kampmann C, Krummenauer F, Ries M, Mengel E, Miebach E, Baehner F, Kim K, Bajbouj M, Schwarting A, Gal A, Beck M. The Mainz Severity Score Index: a new instrument for quantifying the Anderson-Fabry disease phenotype, and the response of patients to enzyme replacement therapy. Clin Genet 2004;65:299-307.
290. Mehta A, Ricci R, Widmer U, Dehout F, Garcia de Lorenzo A, Kampmann C, Linhart A, Sunder-Plassmann G, Ries M, Beck M. Fabry disease defined: baseline clinical manifestations of 366 patients in the Fabry Outcome Survey. Eur J Clin Invest 2004;34:236-42.
291. Mehta A, Clarke JT, Giugliani R, Elliott P, Linhart A, Beck M, Sunder-Plassmann G, Investigators FOS. Natural course of Fabry disease: changing pattern of causes of death in FOS - Fabry Outcome Survey. J Med Genet 2009;46:548-52.
292. Acuna-Hidalgo R, Veltman JA, Hoischen A. New insights into the generation and role of de novo mutations in health and disease. Genome Biol 2016;17:241.
293. Kataoka Y, Kalanithi PS, Grantz H, Schwartz ML, Saper C, Leckman JF, Vaccarino FM. Decreased number of parvalbumin and cholinergic interneurons in the striatum of individuals with Tourette syndrome. J Comp Neurol 2010;518:277-91.
294. Paschou P, Stylianopoulou E, Karagiannidis I, Rizzo R, Tarnok Z, Wolanczyk T, Hebebrand J, Nothen MM, Lehmkuhl G, Farkas L, Nagy P, Szymanska U, Lykidis D, Androutsos C, Tsironi V, Koumoula A, Barta C, Klidonas S, Ypsilantis P, Simopoulos C, Consortium TS, Skavdis G, Grigoriou M. Evaluation of the LIM homeobox genes LHX6 and LHX8 as candidates for Tourette syndrome. Genes Brain Behav 2012;11:444-51.
295. Coppen EM, Roos RA. Current Pharmacological Approaches to Reduce Chorea in Huntington's Disease. Drugs 2017;77:29-46.
296. Yero T, Rey JA. Tetrabenazine (Xenazine), An FDA-Approved Treatment Option For Huntington's Disease-Related Chorea. P T 2008;33:690-4.

297. Johnson BA, Ait-Daoud N, Wang XQ, Penberthy JK, Javors MA, Seneviratne C, Liu L. Topiramate for the treatment of cocaine addiction: a randomized clinical trial. JAMA Psychiatry 2013;70:1338-46.
298. Hedera P, Cibulcik F, Davis TL. Pharmacotherapy of essential tremor. J Cent Nerv Syst Dis 2013;5:43-55.
299. Wu J, Tang H, Chen S, Cao L. Mechanisms and Pharmacotherapy for Ethanol-Responsive Movement Disorders. Front Neurol 2020;11:892.
300. Kim MS, Hur MK, Son YJ, Park JI, Chun SY, D'Elia AV, Damante G, Cho S, Kim K, Lee BJ. Regulation of pituitary adenylate cyclase-activating polypeptide gene transcription by TTF-1, a homeodomain-containing transcription factor. J Biol Chem 2002;277:36863-71.
301. Yee CL, Wang Y, Anderson S, Ekker M, Rubenstein JL. Arcuate nucleus expression of NKX2.1 and DLX and lineages expressing these transcription factors in neuropeptide Y(+), proopiomelanocortin(+), and tyrosine hydroxylase(+) neurons in neonatal and adult mice. J Comp Neurol 2009;517:37-50.
302. Gereben B, Salvatore D, Harney JW, Tu HM, Larsen PR. The human, but not rat, dio2 gene is stimulated by thyroid transcription factor-1 (TTF-1). Mol Endocrinol 2001;15:112-24.
303. Gilbert ME, Sui L, Walker MJ, Anderson W, Thomas S, Smoller SN, Schon JP, Phani S, Goodman JH. Thyroid hormone insufficiency during brain development reduces parvalbumin immunoreactivity and inhibitory function in the hippocampus. Endocrinology 2007;148:92-102.
304. Gould E, Butcher LL. Developing cholinergic basal forebrain neurons are sensitive to thyroid hormone. J Neurosci 1989;9:3347-58.
305. Santagati F, Abe K, Schmidt V, Schmitt-John T, Suzuki M, Yamamura K, Imai K. Identification of Cis-regulatory elements in the mouse Pax9/Nkx2-9 genomic region: implication for evolutionary conserved synteny. Genetics 2003;165:235-42.
306. Kim JG, Park BS, Yun CH, Kim HJ, Kang SS, D'Elia AV, Damante G, Lee KU, Park JW, Kim ES, Namgoong IS, Kim YI, Lee BJ. Thyroid transcription factor-1 regulates feeding behavior via melanocortin pathway in the hypothalamus. Diabetes 2011;60:710-9.
307. Prabhakaran D, Anand S, Watkins D, Gaziano T, Wu Y, Mbanya JC, Nugent R, Disease Control Priorities-3 Cardiovascular R, Related Disorders Author G. Cardiovascular, respiratory, and related disorders: key messages from Disease Control Priorities, 3rd edition. Lancet 2018;391:1224-36.
308. Peterson B, Coda A, Pacey V, Hawke F. Physical and mechanical therapies for lower limb symptoms in children with Hypermobility Spectrum Disorder and Hypermobile Ehlers-Danlos Syndrome: a systematic review. J Foot Ankle Res 2018;11:59.
309. Krajewski W, Wojciechowska J, Dembowski J, Zdrojowy R, Szydelko T. Hydronephrosis in the course of ureteropelvic junction obstruction: An underestimated problem? Current opinions on the pathogenesis, diagnosis and treatment. Adv Clin Exp Med 2017;26:857-64.
310. Chen L, Guan J, Gu H, Zhang M. Outcomes in fetuses diagnosed with megacystis: Systematic review and meta-analysis. Eur J Obstet Gynecol Reprod Biol 2018;233:120-6.
311. Kalfa N, Gaspari L, Ollivier M, Philibert P, Bergougnoux A, Paris F, Sultan C. Molecular genetics of hypospadias and cryptorchidism recent developments. Clin Genet 2019;95:122-31.
312. Glocker E, Ehl S, Grimbacher B. Common variable immunodeficiency in children. Curr Opin Pediatr 2007;19:685-92.

313. Swarup I, Penny CL, Dodwell ER. Developmental dysplasia of the hip: an update on diagnosis and management from birth to 6 months. Curr Opin Pediatr 2018;30:84-92.
314. Comer GC, Potter M, Ladd AL. Polydactyly of the Hand. J Am Acad Orthop Surg 2018;26:75-82.
315. Diamond DA, Chan IHY, Holland AJA, Kurtz MP, Nelson C, Estrada CR, Jr., Bauer S, Tam PKH. Advances in paediatric urology. Lancet 2017;390:1061-71.
316. Schwegler H, Boldyreva M, Linke R, Wu J, Zilles K, Crusio WE. Genetic variation in the morphology of the septo-hippocampal cholinergic and GABAergic systems in mice: II. Morpho-behavioral correlations. Hippocampus 1996;6:535-45.
317. Marin O. A postnatal function for Nkx2-1 in basal forebrain integrity (Commentary on Magno et al.). Eur J Neurosci 2011;34:1766.

Eidesstattliche Versicherung

„Ich, Sara Ersözlü, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema: „Effects of pre- and postnatal deletion of the transcription factor Nkx2-1 on the expression of *NGF, trkA, trkB* and $p75^{NTR}$ in mice and a clinical update on *NKX2-1* haploinsufficiency in humans“ selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe.

Alle Stellen, die wörtlich oder dem Sinne nach auf Publikationen oder Vorträgen anderer Autoren beruhen, sind als solche in korrekter Zitierung kenntlich gemacht. Die Abschnitte zu Methodik (insbesondere praktische Arbeiten, Laborbestimmungen, statistische Aufarbeitung) und Resultaten (insbesondere Abbildungen, Graphiken und Tabellen werden von mir verantwortet.

Meine Anteile an etwaigen Publikationen zu dieser Dissertation entsprechen denen, die in der untenstehenden gemeinsamen Erklärung mit dem Betreuer, angegeben sind. Für sämtliche im Rahmen der Dissertation entstandenen Publikationen wurden die Richtlinien des ICMJE (International Committee of Medical Journal Editors; www.icmje.og) zur Autorenschaft eingehalten. Ich erkläre ferner, dass mir die Satzung der Charité – Universitätsmedizin Berlin zur Sicherung Guter Wissenschaftlicher Praxis bekannt ist und ich mich zur Einhaltung dieser Satzung verpflichte.

Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§156,161 des Strafgesetzbuches) sind mir bekannt und bewusst.“

Datum Unterschrift

Anteilserklärung an etwaigen erfolgten Publikationen

Sara Ersözlü hatte folgenden Anteil an der folgenden Publikation:

Magno L, Kretz O, Bert B, Ersözlü S, Vogt J, Fink H, Kimura S, Vogt A, Monyer H, Nitsch R, Naumann T. The integrity of cholinergic basal forebrain neurons depends on expression of Nkx2-1. Eur J Neurosci. 2011 Dec;34(11):1767-82

Beitrag im Einzelnen: Zucht und Kreuzung der Mutantenmauslinien, Genotypisierung, X-Gal Färbung, immunhistochemische Färbungen, quantitative real-time-PCR und die entsprechenden Vorbereitungen des Materials, Datenaufbereitung und -analyse. Aus den eigenständigen Experimenten sind in der Publikation Anteile der Abbildung 1 E-I (siehe Abbildung 9 A–E in der Dissertation) und der Abbildung 5 K–L (siehe Abbildung 13 A–B in der Dissertation) hervorgegangen. Durch die Mitarbeit im Team hat Frau Ersözlü auch einen Beitrag an den Acetylcholinesterase-Untersuchungen geleistet (in der Dissertation nicht weiter erwähnt).

_____

Unterschrift, Datum und Stempel des betreuenden Hochschullehrers

_____

Unterschrift der Doktorandin

Curriculum Vitae

My curriculum vitae does not appear in the electronic version of my paper for reasons of data protection.

Publication List

Publications in peer-reviewed journals:

1) Magno L, Kretz O, Bert B, Ersözlü S, Vogt J, Fink H, Kimura S, Vogt A, Monyer H, Nitsch R, Naumann T. The integrity of cholinergic basal forebrain neurons depends on expression of Nkx2-1. Eur J Neurosci. 2011 Dec;34(11):1767-82. – *Featured article* –
2) Ersözlü S, Hofmänner D, Keller DI. Stabbing pain in the throat after teeth cleaning. BMJ Case Rep. 2017 Aug 11;2017.
3) Ersözlü S, Desnick RJ, Huynh-Do U, Canaan-Kühl S, Barbey F, Genitsch V, Müller T, Cheetham M, Flammer A, Schaub S, Nowak A. Long-Term Outcomes of Kidney Transplantation in Fabry Disease. Transplantation. 2018 Apr 24.
4) Ersözlü S, Nowak A. Morbus Fabry kann die Ursache einer unklaren Linksherzhypertrophie sein. info@herz+gefäss _ 01_ 2019

Abstracts:

1) Ersözlü S, Magno L, Naumann T. Effects of pre- and postnatal deletion of the transcription factor Nkx2-1 on the expression of NGF, trkA, trkB, and $p75^{NTR}$ in mice. Abstract Book First International Student Congress of Austria. 2013 July;6:45-46.

   ISC 2013 – International Student Congress, Medical University of Graz, Austria, July 4–6, 2013 *("1st Place of all oral presentations")*

2) Ersözlü S, Magno L, Naumann T. Effects of pre- and postnatal deletion of the transcription factor Nkx2-1 on the expression of NGF, trkA, trkB, and $p75^{NTR}$ in mice. BMC Proceedings. 9. A6-A6. 10.1186/1753-6561-9-S1-A6.

   ICHAMS 2013 – International Conference for Healthcare and Medical Students, Royal College of Surgeons in Ireland, Dublin, October 11–12, 2013 *("1st Overall oral presentation winner")*

3) Ersözlü S, Desnick RJ, Huynh-Do U, Canaan-Kühl S, Barbey F, Genitsch V, Müller T, Cheetham M, Flammer A, Schaub S, Nowak A. Long-Term Outcomes of Kidney Transplantation in Fabry Disease.

   48th Annual Meeting of the Swiss Society of Nephrology, Interlaken, Switzerland, December 8–9, 2016 *(Poster Presentation and Selection for Oral Presention)*

4) Ersözlü S, Desnick RJ, Huynh-Do U, Canaan-Kühl S, Barbey F, Genitsch V, Müller T, Cheetham M, Flammer A, Schaub S, Nowak A. Long-Term Outcomes of Kidney Transplantation in Fabry Disease.

   15th European Round Table on Fabry Disease, Berlin, March 10–11, 2017 *(1st Poster Prize)*

5) Nowak A, Desnick RJ, Huynh-Do U, Canaan-Kühl S, Barbey F, Genitsch V, Müller T, Cheetham M, Flammer A, Schaub S, Ersözlü S. Long-Term Outcomes of Kidney Transplantation in Fabry Disease. Molecular Genetics and Metabolism. 123. S108. 10.1016/j.ymgme.2017.12.286.

6) Ersözlü S, Rosset-Zufferey S, Genitsch V, Wustmann K, Huynh-Do U. Fabrys disease at a second glance: a case report of a patient presenting with later onset Fabry cardio- and nephropathy.

   51st Annual Meeting of the Swiss Society of Nephrology, Interlaken, Switzerland, December 4-6, 2019 *(Poster Presentation)*

Aknowledgements

This dissertation was begun after the second semester of my medical studies at the Charité, when my anatomy Professor, PD Dr. med Thomas Naumann, gave me the early opportunity to start working in his team researching the transcription factor Nkx2-1 in the mouse brain. At the age of 14 years, I had already participated in a neuroscientific course at the Student's Summer University of the Free University in Berlin, investigating the learning behavior of bees. After this, I knew that I wanted to be involved in neuroscience some day. I am so grateful to my "Doktorvater" Dr. Naumann for giving me this opportunity, for mentoring me both scientifically and as a young physician beyond neuroscience. The many occasions where we also talked about the historical, philosophical, and humanitarian aspects of current and past medical practice have shaped me. I also owe special thanks to Dr. Lorenza Magno, PhD of Medical Neurosciences, who taught me how to perform experiments in the lab, helped me with obtaining the neuroscientific basic knowledge and specific knowledge on Nkx2-1 in the mouse brain. I enjoyed working together with her a lot, until she finished her PhD and moved to London in 2010. I am thankful for the proofreading of my research paper (Hausarbeit in 2010) and this dissertation. I am moreover grateful to Gudrun Thomascheck and Christian Guijarro for technical assistance in the lab.

After the experimental work was completed and I proceeded with the clinical part of my studies I started to get more and more interested in the human brain-lung-thyroid syndrome. With this in mind, I first performed a small review for the introduction and the discussion of my dissertation. Later, as the data turned out to be very interesting, we decided to turn it into a systematic review that I performed in my last years alongside my training in Internal Medicine and Cardiology, where I also started to specialize in rare genetic multi-organ disorders like Morbus Fabry and haemochromatosis. In the course of my research work regarding Morbus Fabry at the University Hospital in Zurich, I got into clinical research involving human genetics and statistical analysis that allowed me to also transfer knowledge to my *NKX2-1* review. I am thankful to PD Dr. med. Albina Nowak for her continuous mentoring me in this field since 2016.

I also want to thank Janic Teutsch, MSc. Biology from the Division for Human Genetics of the University Hospital Bern for checking the nomenclature of the human mutations and the coordinates of the *NKX2-1* domains and cooperation on the review article on *NKX2-1* haploinsufficiency in humans – which is currently in preparation.

I am furthermore grateful to the German Merit Foundation (Studienstiftung) for the scholarschip during my medical studies and the research semester, including travel grants to present my research work at international conferences.

Finally, I am very happy and thankful for all the support of my family and friends, in particular of my mother, during the different phases of this long journey in trying to understand this rare disease.