European Journal of Medical Genetics 65 (2022) 104632

Contents lists available at ScienceDirect

European Journal of Medical Genetics

journal homepage: www.elsevier.com/locate/ejmg

Genotype-phenotype associations in a large PTEN Hamartoma Tumor Syndrome (PHTS) patient cohort

Linda A.J. Hendricks [a,b], Nicoline Hoogerbrugge [a,c], Hanka Venselaar [d], Stefan Aretz [e,f], Isabel Spier [e,f], Eric Legius [g], Hilde Brems [g], Robin de Putter [h], Kathleen B.M. Claes [h], D. Gareth Evans [i], Emma R. Woodward [i], Maurizio Genuardi [j,k], Fulvia Brugnoletti [j], Yvette van Ierland [l,m], Kim Dijke [l], Emma Tham [n,o], Bianca Tesi [n,o], Janneke H.M. Schuurs-Hoeijmakers [a], Maud Branchaud [p], Hector Salvador [q], Arne Jahn [r,s,t,u,v], Simon Schnaiter [w], Violetta Christophidou Anastasiadou [x], Joan Brunet [y], Carla Oliveira [z,aa], Laura Roht [ab,ac], Ana Blatnik [ad], Arvids Irmejs [ae,af], Arjen R. Mensenkamp [a,c], Janet R. Vos [a,b,*], PTEN Study Group

[a] *Department of Human Genetics, Radboudumc Expert Center for PHTS, Radboud university medical center, Nijmegen, the Netherlands*
[b] *Radboud university medical center, Radboud Institute for Health Sciences, Nijmegen, the Netherlands*
[c] *Radboud university medical center, Radboud Institute for Molecular Life Sciences, Nijmegen, the Netherlands*
[d] *Centre for Molecular and Biomolecular Informatics, Radboud Institute for Molecular Life Sciences, Radboud university medical center, Nijmegen, the Netherlands*
[e] *Institute of Human Genetics, Medical Faculty, University of Bonn, Bonn, Germany*
[f] *Center for Hereditary Tumor Syndromes, University Hospital Bonn, Bonn, Germany*
[g] *Department of Human Genetics, University of Leuven, Leuven, Belgium*
[h] *Center for Medical Genetics, Ghent University Hospital, Ghent, Belgium*
[i] *Manchester Centre for Genomic Medicine, St Mary's Hospital, Division of Evolution and Genomic Sciences, School of Biological Sciences, University of Manchester, Manchester, UK*
[j] *Department of Laboratory and Infectious Diseases, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy*
[k] *Medical Genetics Section, Department of Life Sciences and Public Health, Università Cattolica del Sacro Cuore, Rome, Italy*
[l] *Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, the Netherlands*
[m] *ENCORE Expertise Center, Erasmus University Medical Center, Sophia Children's Hospital, Rotterdam, the Netherlands*
[n] *Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden*
[o] *Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden*
[p] *Department of Genetics, Normandy Center for Genomic and Personalized Medicine, Normandie Univ, UNIROUEN, Inserm U1245 and Rouen University Hospital, Rouen, France*
[q] *Department of Oncology, Hospital Sant Joan de Déu, Universitat de Barcelona, Barcelona, Spain*
[r] *Institute for Clinical Genetics, Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany*
[s] *Hereditary Cancer Syndrome Center Dresden, Dresden, Germany*
[t] *German Cancer Consortium (DKTK), Dresden, Germany*
[u] *National Center for Tumor Diseases (NCT), Partner Site Dresden, Germany*
[v] *German Cancer Research Center (DKFZ), Heidelberg, Germany*
[w] *Institute of Human Genetics, Medical University of Innsbruck, Innsbruck, Austria*
[x] *Karaiskakio Foundation, Nicosia Cyprus and Archbishop Makarios III Children's Hospital, Cyprus*
[y] *Hereditary Cancer Program, Catalan Institute of Oncology, ONCOBELL-IDIBELL-IDIBGI-IGTP, CIBERONC, Barcelona, 08908, Spain*
[z] *Instituto de Investigação e Inovação em Saúde & Institute of Molecular Pathology and Immunology of the University of Porto, Porto, Portugal*
[aa] *Department of Pathology, University of Porto, Porto, Portugal*
[ab] *Department of Clinical Genetics, United Laboratories, Tartu University Hospital, Tartu, Estonia*
[ac] *Department of Clinical Genetics, Institute of Clinical Medicine, University of Tartu, Tartu, Estonia*
[ad] *Department of Clinical Cancer Genetics, Institute of Oncology Ljubljana, Ljubljana, Slovenia*
[ae] *Institute of Oncology, Riga Stradins University, Riga, Latvia*
[af] *Breast Unit, Pauls Stradins Clinical University Hospital, Riga, Latvia*

* Corresponding author. Department of Human Genetics, Radboud university medical center, P.O. Box 9101, 6500 HB, Nijmegen, the Netherlands.
E-mail address: janet.vos@radboudumc.nl (J.R. Vos).

https://doi.org/10.1016/j.ejmg.2022.104632
Received 20 May 2022; Received in revised form 5 September 2022; Accepted 28 September 2022
Available online 18 October 2022

ARTICLE INFO

Keywords:
Genetic variation
Genetic association studies
Human genetics
Medical oncology
Phenotype

ABSTRACT

Background: Pathogenic *PTEN* germline variants cause PTEN Hamartoma Tumor Syndrome (PHTS), a rare disease with a variable genotype and phenotype. Knowledge about these spectra and genotype-phenotype associations could help diagnostics and potentially lead to personalized care. Therefore, we assessed the PHTS genotype and phenotype spectrum in a large cohort study.

Methods: Information was collected of 510 index patients with pathogenic or likely pathogenic (LP/P) *PTEN* variants (n = 467) or variants of uncertain significance. Genotype-phenotype associations were assessed using logistic regression analyses adjusted for sex and age.

Results: At time of genetic testing, the majority of children (n = 229) had macrocephaly (81%) or developmental delay (DD, 61%), and about half of the adults (n = 238) had cancer (51%), macrocephaly (61%), or cutaneous pathology (49%). Across *PTEN*, 268 LP/P variants were identified, with exon 5 as hotspot. Missense variants (n = 161) were mainly located in the phosphatase domain (PD, 90%) and truncating variants (n = 306) across all domains. A trend towards 2 times more often truncating variants was observed in adults (OR = 2.3, 95%CI = 1.5–3.4) and patients with cutaneous pathology (OR = 1.6, 95%CI = 1.1–2.5) or benign thyroid pathology (OR = 2.0, 95%CI = 1.1–3.5), with trends up to 2–4 times more variants in PD. Whereas patients with DD (OR = 0.5, 95%CI = 0.3–0.9) or macrocephaly (OR = 0.6, 95%CI = 0.4–0.9) had about 2 times less often truncating variants compared to missense variants. In DD patients these missense variants were often located in domain C2.

Conclusion: The PHTS phenotypic diversity may partly be explained by the *PTEN* variant coding effect and the combination of coding effect and domain. PHTS patients with early-onset disease often had missense variants, and those with later-onset disease often truncating variants.

1. Introduction

PTEN Hamartoma Tumor Syndrome (PHTS), comprising Cowden, Bannayan-Riley-Ruvalcaba and Proteus-like syndromes, is a very rare genetic predisposition disorder that is characterized by both multi-system tissue and cell overgrowth. PHTS is caused by pathogenic germline variants in the tumor suppressor gene Phosphatase and Tensin homolog *(PTEN)*. The prevalence is currently estimated at 1 in 200,000, though this is likely an underestimation as it is presumed that the majority of patients are not recognized (Nelen et al., 1999). Patients have a diverse phenotypic presentation, including macrocephaly, developmental delay, cutaneous, oral, colorectal and thyroid pathology, and benign and malignant tumors across several sites (Hendricks et al., 2021; Yehia et al., 2020). Patients are diagnosed at diverse ages where males are more often diagnosed in childhood and females more often in adulthood (Tan et al., 2011). Macrocephaly, autism spectrum disorder (ASD) or other developmental delay related features are often indications for PHTS diagnosis in childhood. In contrast, diagnosis in adulthood is often related to cancer (Tan et al., 2011). The phenotypic variation is yet unexplained and limits the recognition of PHTS patients. Moreover, it impairs clinical care and increases uncertainty for patients and their family members, because it is hard to predict which symptoms a patient will develop.

Various efforts have been made to elucidate phenotypic diversity in PHTS. Some studies reported associations between truncating *PTEN* germline variants and cancer or breast fibroadenomas (Marsh et al., 1999), between location and nature of the variant and number of involved organ sites (Marsh et al., 1998), and between non-missense variants and thyroid cancer (Nieuwenhuis et al., 2014). Others did not identify any association and found that even within families the same variants leads to diverse phenotypes (Nelen et al., 1999; Lachlan et al., 2007; Bubien et al., 2013; Leslie and Longy, 2016). However, most studies had a limited sample size which hampered the detection of smaller but potentially relevant associations (Nelen et al., 1999; Marsh et al., 1998, 1999; Nieuwenhuis et al., 2014; Lachlan et al., 2007; Bubien et al., 2013).

The diverse phenotypic presentation might be related to the PTEN protein activity. PTEN's lipid and protein phosphatase activities are reported to be critical for tumor suppressor function and neuronal development, respectively (Mondal and Sen, 2020; Yehia et al., 2019). PTEN is also responsible for inhibition of cell proliferation, survival and migration by phosphatidylinositol-3,4,5-triphosphate to phosphatidylinositol-4,5-bisphosphate conversion (Han et al., 2000). Communication and interaction between PTEN's two major domains is proposed for optimal protein functioning. The phosphatase domain (PD) harbors the active site, and is connected to the C2 domain (C2) that is involved in sub-cellular localization (Mondal and Sen, 2020; Lee et al., 1999). Previously, structural and dynamic differences in the active site loop and interdomain region between ASD and/or cancer patients have been suggested (Smith et al., 2019). Furthermore, a relation between PTEN protein dosage and phenotype load has been observed (Tan et al., 2011).

Despite these efforts, the diverse spectrum of PHTS remains largely unexplained. Identification of genotype-phenotype associations could help diagnostics of new patients and potentially lead to more personalized patient care in the future. Therefore the aim of this study was to assess the PHTS phenotype and variant spectrum and genotype-phenotype associations in a large, well defined PHTS cohort.

2. Methods

2.1. Study design

Adult and pediatric PHTS index patients, who were diagnosed through routine clinical care, were retrospectively accrued from clinical genetic centers across Europe. Patients with either a *PTEN* germline variant of uncertain significance (VUS) or a pathogenic or likely pathogenic (LP/P) variant, according to the reporting genetic laboratories, were included. Only index patients, PHTS patients who were the first in their family to be diagnosed, were included to have a similar degree of ascertainment bias across the cohort and to prevent familial clustering.

2.2. Phenotypic information

Information on the patient's genotype and phenotype is routinely recorded at time of genetic testing as part of clinical care. For this study, data was systematically collected from all collaborating centers using a data dictionary which included *PTEN* variant notation, family history, and the following phenotypic features: cancer, macrocephaly, developmental delay, cutaneous pathology, benign thyroid pathology, and other PHTS-related phenotypic information. Genotype-phenotype associations were assessed for the most commonly reported phenotypic characteristics.

To assess the phenotype-spectrum and lipid phosphatase activity of the *PTEN* variant, five groups of variants were defined: 1) variants only observed in patients with developmental delay (DD), but without

cancer; 2) variants only observed in patients without DD and cancer; 3) variants observed in patients with cancer, but without DD; 4) variants observed in both patients with cancer and no DD and in patients with DD and no cancer; 5) variants observed in patients with both a history of DD and cancer. These groups were defined using phenotypic information from included index patients, their family history, and literature (Yehia et al., 2020; Bubien et al., 2013; Smith et al., 2019; Spinelli et al., 2015; Portelli et al., 2021; Mighell et al., 2020; Varga et al., 2009).

2.3. Variant classification

Patient selection was based on the initial variant pathogenicity classification provided by the genetic laboratories. This selection comprised 306 different variants in 510 patients, of which 268 variants were classified as LP/P and 38 as VUS (Supplementary Table 1).

All variant classifications provided by the local genetic laboratories were re-evaluated to include latest (molecular) insights and to ensure uniform classification. This was performed using adjusted *PTEN* specific American College of Medical Genetics and Genomics (ACMG) criteria (Mester et al., 2018).

The initial variant classification was compared to the re-evaluated variant classification. Both classifications are presented in Supplementary Table 1. Because sensitivity analyses did not show differences between the results (data not shown), the initial variant classification provided by the genetic laboratories was used for the main analyses and results in this study.

2.4. Genotypic information

Variants were assessed and annotated according to GRCh37/hg19 (NM_000314.4) and were submitted to LOVD (www.lovd.nl/3.0/home). Variants were further analyzed *in silico* using software packages SpliceSiteFinder-like, MaxEntScan, NNSPLICE, and GeneSplicer for potential splice-site variants, integrated in the AlamutVisual software package (v2.13, Interactive Biosoftware). Coding effects were classified in three categories: (1) truncating, including frameshift, nonsense, start-loss, large structural variants (deletions of (the majority) of ≥1 exon), splice-site variants (± 5 base pairs (bp) of intron/exon boundary), and cryptic splice site variants; (2) missense, including missense, and in-frame variants; (3) other, including variants in 5′- and 3′-UTR regions, and intronic variants exceeding intron/exon boundaries with more than 5bp.

Variants were spatially categorized according to intron/exon boundaries, protein domains and their orientation on the 3D crystal structure. Protein domain boundaries in amino acids (AA) were defined as: Phosphatase Binding Domain (PBD) AA 1–6, PD AA 7–185, C2 AA 186–351, C-terminal Region (CTR) AA 352–400, and PDZ domain AA 401–403 (Nieuwenhuis et al., 2014; Yehia et al., 2019; Lee et al., 1999). For statistical analyses were variants that spanned multiple domains, were not located in a domain, or were located in PBD, CTR or PDZ grouped together.

In addition, the predicted PTEN missense variant tolerance at each AA position was obtained from the MetaDome web server. PTEN lipid phosphatase activity after missense variant introduction (excluding in-frame variants) was assessed by using the fitness scores and categories, including truncating-like, hypomorphic, and wildtype-like, calculated by Mighell et al., 2018, 2020. Furthermore, the expected occurrence of nonsense mediated mRNA decay (NMD) was assessed per phenotype. NMD was classified as such when a premature stop codon presented ≥50 nucleotides upstream of the last exon junction (i.e. 5' of c.977). (Nagy and Maquat, 1998).

2.5. PTEN 3D protein analysis

The human wildtype PTEN 3D crystal structure (PDB: 1D5R) was obtained from the RSCB Protein Data Bank, and visualized using the in-silico prediction tool YASARA (V.20.10.4) (Lee et al., 1999; Berman et al., 2000). In this crystal structure remained AA at the N-terminus (1–13), C-terminus (352–403) and the flexible loop (282–312) unresolved. PTEN protein structural stability for missense variants was assessed by the web-based program INPS3D. Stability scores were expressed as changed unfolding free Gibbs energy between the mutant and wildtype protein (DDG). DDG indicates increased (DDG>0) or decreased (DDG<0) stability. Additional spatio-functional classification of PTEN for membrane binding (MB), active site (AS), and interaction surface between domains (IS), and general folding was performed according to information present in literature (Supplementary Table 2). However, insufficient information on general protein folding was available. For spatio-functional classification, phenotypic information of patients with missense variants on the same AA position was combined.

2.6. Statistical analyses

Descriptive statistics were performed using the appropriate measures depending on data distribution. Results were presented with 95% confidence intervals (95%CI) or interquartile ranges (IQR, i.e. 25th – 75th percentile), when applicable. Differences in continuous and categorial outcomes were assessed using the Wilcoxon rank-sum test and the chi-squared test, respectively. Genotype-phenotype associations (OR, 95% CI) were assessed using logistic regression with and without adjustment for age and sex. Two-sided p-values below 0.05 were considered statistically significant. Analyses were performed with and without Bonferroni correction for multiple testing. All analyses were performed using RStudio (V.3.6.2).

3. Results

3.1. Study population

A total of 467 PHTS patients with LP/P *PTEN* variants were recruited from 26 different European clinical genetic centers in Austria, Belgium, Cyprus, Estonia, France, Germany, Italy, Latvia, the Netherlands, Portugal, Slovenia, Spain, Sweden, and the United Kingdom. Forty-nine percent of patients were pediatric (<18 years) of which 66% were male, while of the adults 70% was female (Table 1).The median age of last clinical information was 5 years (IQR 2–10) for pediatric patients and 44 years (IQR 33–52) for adult patients.

3.2. PHTS phenotype spectrum

More than half of pediatric patients had both macrocephaly and DD (57%), and only 6 patients had cancer (3%) (Table 1, Fig. 1). About half of the adult patients had cancer (51%), macrocephaly (61%), cutaneous pathology (49%), or benign thyroid pathology (42%). Overall, other PHTS-related features such as oral lesions, benign breast pathology, Lhermitte-Duclos Disease, vascular abnormalities and benign gastrointestinal pathology were less often reported with 5–16% each (Table 1). In total 204 cancers were reported in 128 patients, of which 54 patients had multiple or bilateral cancers. The majority of these were female adults (n = 102) who presented mainly with breast (72%), thyroid (23%), endometrial (17%), and skin cancer (8%). Other cancers, including gastrointestinal and renal cancers, were reported less frequently (1–6%). Adult male patients with cancer (n = 20, 28%) presented mainly with colorectal (45%), thyroid (30%) and skin cancer (15%). Cancers not associated with the PHTS spectrum were reported in 5–10% and included prostate and pancreatic cancer.

In this cohort, 268 different LP/P *PTEN* variants were observed. About 26% (n = 70) of the variants were observed in patients with DD but without cancer, 22% (n = 60) in patients with cancer but without DD, and 24% (n = 63) in patients without DD and cancer. About 28% (n = 75) of the variants were reported in patients with DD and cancer, including 22% (n = 59) where the same variant, of which 44%

Table 1
Study population.[a]

	Total population			Pediatric population			Adult population		
	Total	*Male*	*Female*	*Total*	*Male*	*Female*	*Total*	*Male*	*Female*
Demographics									
Population *N (%)*	467	222 (48)	245 (52)	229 (49)	150 (66)	79 (34)	238 (51)	72 (30)	166 (70)
Median age *y*	18	8	34	5	4	6	44	44	44
(IQR)	(5–44)	(3–30)	(11–47)	(2–10)	(2–8)	(3–11)	(33–52)	(31–54)	(34–50)
Inheritance *N (%)*									
De novo	114 (24)	57 (26)	57 (23)	82 (36)	50 (33)	32 (41)	32 (13)	7 (10)	25 (15)
Not de novo	101 (22)	55 (25)	46 (19)	66 (29)	44 (29)	22 (28)	35 (15)	11 (15)	24 (14)
Unknown	252 (54)	110 (50)	142 (58)	81 (35)	56 (37)	25 (32)	171 (72)	54 (75)	117 (70)
Coding effect *N (%)*									
Truncating[b]	306 (66)	141 (64)	165 (67)	130 (57)	88 (59)	42 (53)	176 (74)	53 (74)	123 (74)
Missense	161 (34)	81 (36)	80 (33)	99 (43)	62 (41)	37 (47)	62 (26)	19 (26)	43 (26)
Domain *N (%)*									
Phosphatase	252 (54)	125 (56)	127 (52)	126 (55)	82 (55)	44 (56)	126 (53)	43 (60)	83 (50)
C2	126 (27)	59 (27)	67 (27)	64 (28)	45 (30)	19 (24)	62 (26)	14 (19)	48 (29)
Other/multiple[c]	89 (19)	38 (17)	51 (21)	39 (17)	23 (15)	16 (20)	50 (21)	15 (21)	35 (21)
Phenotype[d] *N (%)*									
Cancer	128 (27)	24 (11)	104 (42)	6 (3)	4 (3)	2 (3)	122 (51)	20 (28)	102 (61)
Macrocephaly	332 (71)	171 (77)	161 (66)	186 (81)	124 (83)	62 (78)	146 (61)	47 (65)	99 (60)
Developmental delay	164 (35)	105 (47)	59 (24)	139 (61)	94 (63)	45 (57)	25 (11)	11 (15)	14 (8)
Cutaneous pathology	177 (38)	87 (39)	90 (37)	60 (26)	46 (31)	14 (18)	117 (49)	41 (57)	76 (46)
Benign thyroid pathology	112 (24)	27 (12)	85 (35)	12 (5)	7 (5)	5 (6)	100 (42)	20 (28)	80 (48)
Oral lesions	51 (11)	12 (5)	39 (16)	7 (3)	5 (3)	2 (2)	44 (18)	7 (10)	37 (22)
Benign breast pathology	23 (5)	1 (0.4)	22 (9)	2 (1)	0 (1)	2 (3)	21 (9)	1 (1)	20 (12)
Lhermitte-Duclos Disease	35 (7)	9 (4)	26 (11)	2 (1)	0 (0)	2 (3)	33 (14)	9 (13)	24 (14)
Vascular abnormalities	57 (12)	25 (11)	32 (13)	21 (9)	9 (6)	12 (15)	36 (15)	16 (22)	20 (12)
Benign gastrointestinal pathology	76 (16)	38 (17)	38 (16)	9 (4)	7 (5)	2 (3)	67 (28)	31 (43)	36 (22)

[a] Characteristics of the PHTS study population with a pathogenic or likely pathogenic *PTEN* germline variant are presented.
[b] The category 'other' contained only 1 variant. This variant was included in analyses in the group 'truncating'.
[c] Variants spanning both phosphatase and C2 domains, located in intronic regions or located in other domains than phosphatase or C2 were included in the group 'other/multiple'.
[d] Phenotypic categories include the following phenotypes: Macrocephaly: head circumference of > 58.5 cm for adult females and >61.5 for adult males, head circumference of >2SD, head circumference > p97, or according to providing doctor without specified head circumference; Developmental delay: autism, autism spectrum disorder, ADHD, learning difficulties, cognitive delay, psychomotor retardation, motor delay, mental retardation, intellectual disability, verbal apraxia, behavioral issues, unspecified developmental delay, unspecified neurodevelopmental disorder, dyspraxia, speech/language delay; Cutaneous pathology: lipoma, papilloma, trichilemmoma, skin hamartoma, keratosis, hyperkeratosis, hypochromic patch, café-au-lait, fibroma, cutaneous pathology unspecified, naevus flammeus, maculae pigmentation penis, penile freckling, trichodiscoma, unspecified mucocutaneous lesions, naevi, pits, papules, sensitive skin, acanthosis nigricans; Benign thyroid pathology: (auto-immune) thyroiditis, Hashimoto's thyroiditis, benign follicular tumor, unspecified benign thyroid pathology, cold nodules, nodes, follicular adenoma, goiter, hyperthyroidism, hypothyroidism, struma, cysts, hyperplasia; Oral lesions: abnormal tongue, cobblestones oral mucosa, unspecified bumps oral mucosa, fibroma, hamartoma, papilloma, unspecified papular lesions, dental crowding, gingival hypertrophy, high palate, inflammation and deformities of gums, mucosal lesions; Benign breast pathology: fibroadenoma, cysts, mastopathy, nodules, hamartoma, intraductal papilloma, papillomatosis, unspecified benign breast lesions; Vascular abnormalities: hemangioma, masson tumor, angioma, arteriovenous malformation, cerebral aneurysm, capillary malformations, developmental venous anomaly, undefined vascular abnormalities; Benign gastrointestinal pathology: gastrointestinal polyps, gastrointestinal hyperplasia, chronic gastritis, gastric hamartomas, (colique) polyposis, ganglioneuroma, lipoma, glycogenic acanthosis esophagus, Meckel's diverticulum, micronodules colon, duodenitis, gastric polyps.
N=number; y=years; IQR = interquartile range, i.e. 25th – 75th percentile.

truncating, was observed in different patients with DD or cancer, and 6% (n = 16) where the same variants, of which 63% truncating, was present in patients with both DD and cancer.

3.3. PTEN variant spectrum

Overall, 175 different LP/P truncating variants were identified in 306 patients, and 93 different LP/P missense variants in 161 patients. Of 306 patients with truncating variants had 21 patients a deletion of ≥1 exon, including 7 patients with a total deletion of *PTEN* and 2 patients with a deletion of both *PTEN* and *BMPR1A* (Fig. 2). LOVD variant numbers are presented in Supplementary Table 1. Several variants were identified in more than 10 patients: p.Arg130* (c.388C>T; n = 34, 7%), p.Arg335* (c.1003C>T; n = 18, 4%), and p.Arg233* (c.697C>T; n = 15, 3%). Most patients had a variant located in exon 5–8 (65%), with exon 5 being the hotspot (30%). After correction for exon-length, exon 5 still contained the most variants per base pairs, although not statistically significantly different from other exons. No deep-intronic variants were reported. Of patients with exonic variants (n = 400, 86%), 0.5% was located in PBD, 63% in PD, 32% in C2, 0.3% in CTR, and 5% spanned multiple domains. Missense variants were predominantly located in PD (n = 145, 90%), and truncating variants were located in both C2 (n = 110, 36%) and PD (n = 107, 35%). None of the missense variants in either C2 or PD were located in tolerant regions according to MetaDome.

3.4. Genotype-phenotype associations

3.4.1. Variant coding effects

Patients with macrocephaly or DD, had about two times less often truncating variants than missense variants (OR = 0.6 (95%CI 0.4–0.9) and OR = 0.5 (95%CI 0.3–0.9), respectively). On the contrary, in patients with cancer, cutaneous pathology or benign thyroid pathology truncating variants were 1.2–2.0 times more often observed than missense variants (OR = 1.2 (95%CI 0.7–2.1), OR = 1.6 (95%CI 1.1–2.5), OR = 2.0 (95%CI 1.1–3.5), respectively; Table 2). Truncating variants were 2.3 times more often observed in adult patients than in pediatric patients (OR = 2.3 (95%CI 1.5–3.4)). After correction for multiple testing, none of the results adjusted for sex and age were statistically significant (Supplementary Table 3).

Among patients with a truncating variant, 68% (n = 207) had a variant introducing a premature stop codon due to stop-gain or frameshift. In 87% of these 207 patients, NMD pathway activation was

Fig. 1. Phenotype presentation of each individual index patient, with patients grouped by variant type and age. Schematic phenotypic presentation of all index patients with a pathogenic or likely pathogenic *PTEN* germline variant. Each vertical bar represents one patient. Patients are grouped by variant coding effect (truncating = left; missense = right). The results for the total population (N = 467) are presented on the top, where grey represents that the phenotypic category is absent, and blue that the category is present. The results are presented separately for adults (N = 238) and children (N = 229) below, where grey represents that the phenotypic category is absent, and pink that the category is present.

predicted. However, no phenotypic differences were observed for NMD and non-NMD truncating variants introducing a premature stop codon (data not shown).

3.4.2. Spatial variant orientation

Missense variants were predominantly located in PD. Those missense variants in C2 (n = 16) were observed in patients with macrocephaly (n = 15), DD (n = 14), cutaneous pathology (n = 2), or thyroid cancer (n = 1), but not in patients with benign thyroid pathology or other cancers (Supplementary Fig. 1). In patients with macrocephaly, 15% of the missense variants were located at an affected AA residue located in the active site and membrane binding site. This was 11% in patients with DD. For patients with cancer, cutaneous pathology or benign thyroid pathology this was higher with 26%, 30% and 31% respectively. For the other sites the frequencies of affected AA residues were similar across all phenotypes: 4–15% only in the interaction surface, 0–4% in the membrane binding site, and 2–13% in both the interaction surface and membrane binding site. No statistically significant correlations were observed between phenotype and domains or exons/introns (data not shown).

3.4.3. Variant coding effect and spatial variant orientation

When combining both coding effect and domain location, the fit of the regression models significantly improved. In patients with cancer, cutaneous pathology or benign thyroid pathology, an increased trend of truncating variants located in C2 was observed and even more so for truncating variants in domain PD (Fig. 3). In patients with DD, a decreased trend of missense variants in PD was observed compared to missense variants in C2, while results for macrocephaly were similar for missense variants in both domains. After correction for multiple testing, no results were statistically significant.

3.4.4. Protein structural stability

The vast majority (94%) of missense variants that underwent structural stability analysis (n = 90) were predicted to reduce the structural stability (Supplementary Fig. 2). Across the phenotypic patient groups, the median structural stability scores ranged from −1.3 to −1.1 and were not significantly different.

3.4.5. Lipid phosphatase activity

Among the 159 patients with missense variants (excluding in-frame variants), 54% had a variant with predicted truncating-like lipid phosphatase activity, 31% with predicted hypomorphic activity, and 14% with predicted wildtype-like activity. Even though, missense variants were observed in only 26% of adult patients compared to 43% in pediatric patients (Table 1), the lipid phosphatase activity distribution of those missense variants was similar between adult patients (n = 61; 64% truncating-like, 26% hypomorphic, and 10% wildtype-like) and pediatric patients (n = 98; 48% truncating-like, 35% hypomorphic, and 17% wildtype-like). Overall trends indicate that variants with disturbed lipid phosphatase activity might be up to 4 times more often present in patients with macrocephaly, cutaneous pathology or benign thyroid pathology, and up to 5 times less often in patients with cancer and DD (Supplementary Table 4). In patients with cancer and without DD (n =

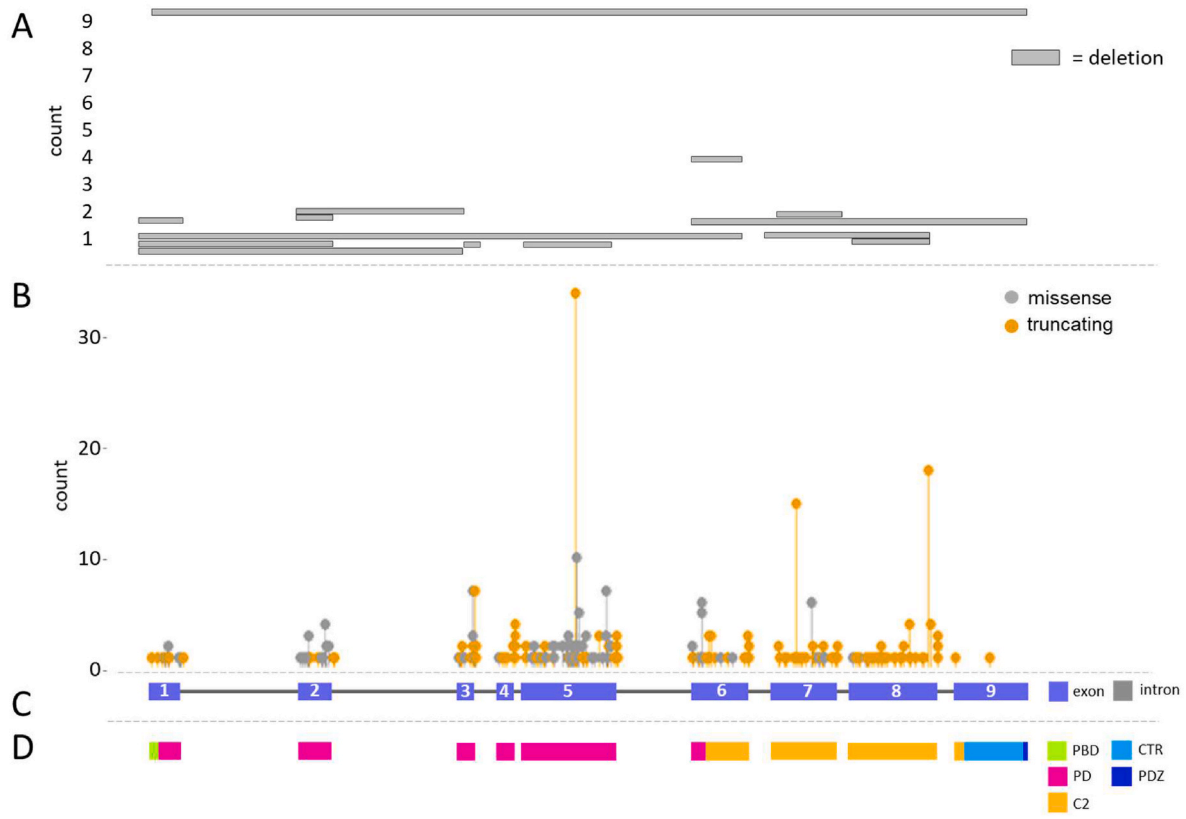

Fig. 2. Schematic variant spectrum of pathogenic and likely pathogenic *PTEN* variants. A) Large structural *PTEN* variants are presented as horizontal bars, with the count of occurrence in the current population on the y-axis. **B)** Variants (except large structural variants) are presented in respect to their location on the linear *PTEN* gene. For small structural variants spanning multiple positions, the first position of the variant is indicated. Colors represent their predicted coding effect. Missense variants are depicted in grey and truncating variants in orange. The variant count in this population is presented on the y-axis. **C)** Scaled presentation of the linear *PTEN* gene, with exonic regions presented in blue and intronic regions in grey. Intronic regions are presented 1:100 to exons. **D)** Protein domain regions are depicted with the Phosphatase Binding Domain (PBD) in green, Phosphatase Domain (PD) in pink, the C2 domain (C2) in orange, the C-terminal Region (CTR) in light blue, and the PDZ domain (PDZ) in dark blue.

14) the predicted lipid phosphatase activity for missense variants tended to be lower compared to a wide range in patients with developmental delay (n = 31) (Supplementary Fig. 3). Results were not statistically significant different.

4. Discussion

This is the largest pediatric and adult PHTS cohort study that assessed the phenotype and variant spectrum and genotype-phenotype associations of this rare syndrome to date, to our knowledge. This cohort study highlighted novel trends of genotype-phenotype associations based on *PTEN's* variant coding effect alone and in combination with domain location. These effects could partly explain the high phenotypic diversity of PHTS, with missense variants more often seen in early-onset disease and truncating variants in later-onset disease in this cohort. Distinct associations of spatial orientation and three-dimensional patterns were not observed. None of the results adjusted for sex and age were statistically significant after correction for multiple testing, which also shows that even in relatively large cohorts, detecting small to moderate effects remain a challenge in rare diseases.

A trend towards two times more truncating variants compared to missense variants was observed in adult patients and in patients with cutaneous and benign thyroid pathology (OR = 2.3 (95%CI 1.5–3.4), OR = 1.6 (95%CI 1.1–2.5) and OR = 2.0 (95%CI 1.1–3.5), respectively). Whereas a trend towards two times fewer truncating variants was observed in patients with macrocephaly and DD (OR = 0.6 (95%CI 0.4–0.9) and OR = 0.5 (95%CI 0.3–0.9), respectively). The higher frequency of truncating variants in patients with a later-onset phenotype was supported by the somewhat higher frequency of truncating-like missense variants in adults compared to children according to lipid-phosphatase activity. These results support the hypothesis that missense variants can act in dominant negative ways and result in similar or more prominent outcomes than truncating variants via haploinsufficiency (Papa et al., 2014). Missense variants were predominantly localized in PD but showed a trend to be less frequent than missense variants in C2 in patients with the early-onset phenotype DD (OR = 0.6 (95%CI 0.3–1.1). Furthermore, in patients with cancer, cutaneous and benign thyroid pathology, a trend towards 1.5 to 2.7 times more C2 truncating variants and 1.7 to 4.1 times more often truncating variants in PD was observed, compared to missense variants in C2. None of these results were statistically significant after correction for multiple testing.

A previous study (n = 146) also indicated that patients with benign thyroid and acral keratosis more often had truncating variants (73% and 76%, not significantly different, Table 3). (Bubien et al., 2013) This is similar to 74% and 79% for patients with benign thyroid pathology and cutaneous pathology in our cohort. Two other studies did not find genotype-phenotype associations for benign pathology but were likely limited by their sample size (n = 22; n = 28, Table 3). (Marsh et al., 1998; Nelen et al., 1999) An enrichment of missense variants in patients with macrocephaly and autism was previously suggested in a small population (n = 10, Table 3) (Leslie and Longy, 2016). This was statistically significant for autism in a larger cohort (n = 295) with similar coding-effect distributions and significance level (without correction for multiple testing) that we observed for DD (Frazier et al., 2015). Additionally, previous studies (n = 43, n = 180) have observed positive

Table 2
Odds ratios for patients with a pathogenic or likely pathogenic variant.[a]

	Genotype		Total population (n = 467)					Pediatric population (n = 229)					Adult population (n = 238)				
	Category/ref	Category	N (%)[b]	OR (95%CI)	*p* OR	$OR_{adj.}$(95%CI)	*p* OR_{adj}	N (%)[b]	OR (95%CI)	*p* OR	$OR_{adj.}$(95%CI)	*p* OR_{adj}	N (%)[b]	OR (95%CI)	*p* OR	$OR_{adj.}$(95%CI)	*p* OR_{adj}
Cancer			**128 (27)**					**6 (3)**					**122 (51)**				
Coding	–	Mis	33 (26)	–	–	–	–	0 (0)	–	–	–	–	33 (27)	–	–	–	–
	Trunc/Mis	Trunc	95 (74)	1.7 (1.1–2.8)	0.016	1.2 (0.7–2.1)	0.541	6 (100)	NA	NA	NA	NA	89 (73)	0.9 (0.5–1.6)	0.719	0.9 (0.5–1.6)	0.647
Domain	PD/C2	C2	38 (30)	0.9 (0.5–1.4)	0.572	0.9 (0.5–1.7)	0.816	3 (50)	NA	NA	NA	NA	35 (29)	0.8 (0.5–1.6)	0.599	1.02 (0.5–2.0)	0.947
	PD/Other	PD	69 (54)	1.2 (0.7–2.2)	0.487	1.7 (0.8–3.4)	0.144	3 (50)	NA	NA	NA	NA	66 (54)	1.5 (0.8–3.0)	0.215	1.7 (0.8–3.5)	0.154
	C2/Other	Other	21 (16)	1.4 (0.8–2.6)	0.289	1.8 (0.8–4.0)	0.137	0 (0)	NA	NA	NA	NA	21 (17)	1.8 (0.8–3.8)	0.130	1.6 (0.7–3.7)	0.227
Macrocephaly			**332 (71)**					**186 (81)**					**146 (61)**				
Coding	–	Mis	129 (39)	–	–	–	–	87 (47)	–	–	–	–	42 (29)	–	–	–	–
	Trunc/Mis	Trunc	203 (61)	0.5 (0.3–0.8)	0.002	0.6 (0.4–0.9)	0.025	99 (53)	0.4 (0.2–0.9)	0.027	0.5 (0.2–0.97)	0.046	104 (71)	0.7 (0.4–1.3)	0.230	0.7 (0.4–1.3)	0.246
Domain	PD/C2	C2	88 (27)	1.2 (0.8–1.9)	0.416	1.2 (0.7–2.0)	0.430	52 (28)	1.4 (0.6–3.1)	0.098	1.3 (0.6–3.0)	0.490	36 (25)	1.2 (0.6–2.2)	0.613	1.1 (0.6–2.1)	0.704
	PD/Other	PD	186 (56)	1.5 (0.9–2.5)	0.121	1.5 (0.9–2.5)	0.162	108 (58)	3.0 (1.3–6.9)	0.010	3.2 (1.4–7.5)	0.007	78 (53)	0.9 (0.5–1.8)	0.796	0.9 (0.5–1.8)	0.831
	C2/Other	Other	58 (17)	1.2 (0.7–2.2)	0.470	1.2 (0.7–2.2)	0.544	26 (14)	2.2 (0.9–5.5)	0.098	2.4 (0.9–6.3)	0.068	32 (22)	0.8 (0.4–1.7)	0.523	0.8 (0.4–1.8)	0.619
DD			**164 (35)**					**139 (61)**					**25 (11)**				
Coding	–	Mis	78 (48)	–	–	–	–	70 (50)	–	–	–	–	8 (32)	–	–	–	–
	Trunc/Mis	Trunc	86 (52)	0.4 (0.3–0.6)	<0.001	0.5 (0.3–0.9)	0.008	69 (50)	0.5 (0.3–0.8)	0.007	0.4 (0.3–0.8)	0.005	17 (68)	0.7 (0.3–1.9)	0.475	0.8 (0.3–2.0)	0.534
Domain	PD/C2	C2	48 (29)	0.9 (0.6–1.5)	0.763	0.9 (0.5–1.5)	0.711	40 (29)	1.0 (0.6–1.9)	0.893	1.1 (0.6–2.0)	0.867	8 (32)	0.7 (0.3–1.9)	0.481	0.6 (0.2–1.8)	0.370
	PD/Other	PD	92 (56)	1.6 (0.9–2.7)	0.104	1.5 (0.8–2.9)	0.160	80 (58)	1.8 (0.9–3.8)	0.102	1.8 (0.9–3.8)	0.110	12 (48)	0.9 (0.3–3.1)	0.923	0.9 (0.3–3.1)	0.911
	C2/Other	Other	24 (15)	1.7 (0.9–3.0)	0.090	1.7 (0.9–3.4)	0.123	19 (14)	1.8 (0.8–4.0)	0.172	1.7 (0.8–3.9)	0.193	5 (20)	1.3 (0.4–4.7)	0.634	1.5 (0.4–5.2)	0.536
Cutaneous			**177 (38)**					**60 (26)**					**117 (49)**				
Coding	–	Mis	46 (26)	–	–	–	–	22 (37)	–	–	–	–	24 (21)	–	–	–	–
	Trunc/Mis	Trunc	131 (74)	1.9 (1.2–2.8)	0.003	1.6 (1.1–2.5)	0.024	38 (63)	1.5 (0.8–2.7)	0.233	1.2 (0.7–2.3)	0.522	93 (79)	1.8 (0.99–3.2)	0.057	1.8 (1.0–3.3)	0.052
Domain	PD/C2	C2	52 (29)	0.8 (0.5–1.3)	0.369	0.8 (0.5–1.3)	0.328	17 (28)	1.1 (0.6–2.2)	0.770	1.2 (0.6–2.5)	0.551	35 (30)	0.6 (0.3–1.1)	0.123	0.6 (0.3–1.1)	0.082
	PD/Other	PD	92 (52)	0.98 (0.6–1.6)	0.923	0.98 (0.6–1.7)	0.945	36 (60)	1.8 (0.8–4.8)	0.190	1.7 (0.7–4.5)	0.276	56 (48)	0.7 (0.4–1.4)	0.366	0.7 (0.4–1.4)	0.342
	C2/Other	Other	33 (19)	1.2 (0.7–2.1)	0.536	1.2 (0.7–2.2)	0.481	7 (12)	1.7 (0.6–4.7)	0.319	1.4 (0.5–3.9)	0.562	26 (22)	1.2 (0.6–2.5)	0.638	1.3 (0.6–2.7)	0.548
Benign thyroid			**112 (24)**					**12 (5)**					**100 (42)**				
Coding	–	Mis	23 (21)	–	–	–	–	2 (17)	–	–	–	–	21 (21)	–	–	–	–
	Trunc/Mis	Trunc	89 (79)	2.5 (1.5–4.2)	<0.001	2.0 (1.1–3.5)	0.019	10 (83)	4.0 (1.0–26.7)	0.076	2.7 (0.6–18.2)	0.232	79 (79)	1.6 (0.9–2.9)	0.132	1.6 (0.9–3.1)	0.128
Domain	PD/C2	C2	32 (29)	0.9 (0.6–1.5)	0.735	0.97 (0.5–1.7)	0.908	6 (50)	0.3 (0.1–1.2)	0.084	0.3 (0.1–1.2)	0.092	26 (26)	1.1 (0.6–2.1)	0.744	1.3 (0.7–2.4)	0.479
	PD/Other	PD	60 (54)	1.1 (0.6–2.0)	0.798	1.3 (0.7–2.5)	0.476	4 (33)	0.6 (1.1–4.5)	0.573	0.4 (0.1–3.3)	0.350	56 (56)	1.4 (0.7–2.8)	0.307	1.5 (0.7–3.0)	0.271
	C2/Other	Other	20 (18)	1.2 (0.6–2.3)	0.622	1.3 (0.6–2.7)	0.466	2 (17)	1.9 (0.4–13.5)	0.441	1.4 (0.3–10.4)	0.738	18 (18)	1.3)0.6–2.8)	0.523	1.2 (0.5–2.6)	0.688

[a] Odds ratios (ORs) are presented for PHTS index patients with a pathogenic or likely pathogenic variant per phenotypic characteristic with corresponding 95% Confidence Intervals (95%CI). Both unadjusted and adjusted ORs for sex and age are presented (OR and OR_{adj}) with corresponding p-values (*p* OR and *p* OR_{adj}). Results are presented without correction for multiple testing. For the coding effect, ORs are presented for truncating variants (Trunc), with missense (Mis) as reference category (Trunc/Mis). For the domain, ORs are presented for the phosphatase domain (PD), with C2 as reference (PD/C2); for PD with 'other' as reference (PD/Other); for C2 with 'other' as reference (C2/Other). DD = developmental delay; Cutaneous = cutaneous pathology; Benign thyroid = benign thyroid pathology. NA = not available, due to too small subgroups. Results are presented for the total population, pediatric population and adult population separately.

[b] For each phenotypic category, the total number of patients with the phenotype is presented in bold with the corresponding percentage from the population (N (%)). For coding effect and domain, the number of patients with the specific phenotype is presented per category, with the percentage from the total number of patients with the phenotype.

associations for non-missense or truncating variants and thyroid or general cancer, supporting positive associations for truncating variants and cancer observed in our study (Table 3). (Marsh et al., 1999; Nieuwenhuis et al., 2014) These positive associations for thyroid and breast malignancies suggest potential tissue preference for non-missense variants. However, PTEN tissue preference remains indistinct. This was supported by reported multiorgan involvement for missense variants (Table 3) (Marsh et al., 1998), by diverse tissues with similar PTEN RNA expression levels (Uhlén et al., 2015), and by multi-promotor involvement in PTEN regulation (Kent et al., 2002). In this cohort, most cancers developed at early age and fit with the PHTS cancer spectrum, but we cannot exclude that these include some sporadic cancers.

Previous studies did not identify evident genotype-phenotype associations for spatial categorization by introns/exons or protein domains (Table 3). (Marsh et al., 1998, 1999; Tan et al., 2012; Bubien et al., 2013; Nieuwenhuis et al., 2014) However, in this large cohort study we found a novel and strong indication of a combined effect of *PTEN* variant coding effect and domain on phenotypic diversity. Spatially the PD domain harbored most variants, especially missense variants, which confirms previous studies (Tan et al., 2011; Nieuwenhuis et al., 2014). Our results indicated that PD missense variants were present in patients with a more diverse phenotypic spectrum compared to the more rare C2 missense variants that mainly were observed in patients with developmental delay and macrocephaly.

In this cohort, also several aspects related to the functioning and effects of *PTEN* gene variants were assessed using external indicators. Predictions of NMD and protein stability did not associate with phenotypic diversity, and distinct three-dimensional patterns for ASD and/or cancer associated variants were absent as previously mentioned (Smith et al., 2019). Mighell et al. previously showed that disturbed lipid phosphatase activity contributed to development of macrocephaly (Mighell et al., 2020). Our study, in combination with the previous obtained lipid phosphatase activity scores by Mighell et al., confirmed that disturbed lipid phosphatase activity contributes to development of macrocephaly, but also to cutaneous pathology and benign thyroid pathology, while it does not or to a lesser extent contribute to development of developmental delay. This is in line with previous results for macrocephaly. Patterns of fitness scores in cancer and/or DD categories in our study, were similar to the patterns that were previously observed by Mighell et al. However, our results suggest that the level of disturbance in lipid phosphatase activity is important for cancer development instead of only the presence or absence of disturbed activity as discussed previously (Leslie and Longy, 2016). Though, some caution is warranted as the lipid-phosphatase activity scoring was only available for missense variants and not for truncating variants and different methods are available to obtain these scores (Han et al., 2000; Mighell et al., 2018, 2020).

A major strength of this epidemiological study is the large cohort of both pediatric and adult PHTS index patients that were recruited from expert centers across Europe, combined with the rigorous evaluation of various components of the PHTS spectrum and genotype-phenotype effect in PHTS. This overview of the genotypic and phenotypic spectrum and the potential association, is of great use in the diagnostic setting when evaluating new patients and can contribute to diagnostic variant interpretation. Although, not statistically significant, the small-to-moderate, genotype-phenotype associations could potentially be clinically relevant and contribute to more personalized disease expectations and healthcare in the future, after a validation and prediction study. Though despite the size of the cohort, no effect was statistically significant after correction for multiple testing. However, given the rarity and phenotypic diversity of PHTS, it is a challenge to have sufficient power for identifying both clinically relevant and statistically significant associations, especially after corrections.

By selecting only index patients identified through routine care, familial clustering in the data was prevented. The index patients were all highly ascertained through similar diagnostic criteria. The cohorts' sex and age distribution reflect clinical referral patterns, where boys with developmental delay and adult females with cancer (e.g. breast) are more often referred and recognized as having PHTS. The observed variant spectrum, variant types, hotspots and common variants also confirmed previous reports (Tan et al., 2011, 2012; Nieuwenhuis et al., 2014; Yehia et al., 2019; Portelli et al., 2021). In this clinical cohort no deep-intronic variants were identified while these are increasingly being reported in hereditary tumor syndromes. This is likely related to the restricting sequencing techniques routinely applied in diagnostic laboratories today. This study design facilitated insight into the *PTEN* variant spectrum and relative effects such as genotype-phenotype associations, meanwhile the absolute frequencies of phenotypes should not directly be used for patient management and translated to the general PHTS population as they are likely overestimated. They can serve as an upper boundary of the expected frequency ranges. On the contrary, underreporting of phenotypic information at moment of genetic diagnoses could underestimate phenotypic prevalence.

Our understanding of both genetic and phenotypic components plays an important role in unravelling genotype-phenotype associations in PHTS. In our cohort was 74% (n = 200) of the variants only identified in a single patient, even when family history and literature were included in the evaluation. With time, likely the *PTEN* variant spectrum gradually expands with more rare variants, and the phenotype associated with a single variant evolves as more patients will be identified and more patients have follow-up on age-dependent events.

To further unravel the genotype-phenotype spectrum in PHTS, future validation series should extend to non-index patients with additional patient follow-up of especially pediatric patients to enable further in-depth analysis of penetrance variation, intrafamilial variation and age-dependent phenotypic events. Many different mechanisms could potentially contribute to phenotype diversity and multifactorial analyses in addition to variant coding effect and domain are advised to further evaluate this diversity. For example, changes in phosphatase activity, protein stability and intramolecular interactions for missense variants are reported as major drivers for the clinical PHTS phenotype (Portelli et al., 2021), and PTEN tumor suppressor function has been suggested via regulation of anaphase-promoting complex (APC) and E-cadherin in the nucleus independent of lipid phosphatase activity (Song et al., 2011). Also other risk factors, such as lifestyle, epigenetic modification and other genetic variation may contribute to the phenotype spectrum. Investigation of protein levels and presence of a second hit in (affected) tissues might give additional insight to disease mechanisms, because this might contribute to observed phenotypic diversity following the two-hit hypothesis (Knudson, 1971). Future *in vitro* and *in vivo* functional studies should assess biochemical protein and lipid phosphatase activities after both missense and truncating variant introduction in a model system and focus on protein-protein interactions, compensation mechanisms, conformational changes, and membrane and substrate binding in order to elucidate PTEN functions in relation to phenotypic diversity.

Taken together, this large cohort study provided an overview of the phenotypic and variant spectrum of PHTS that can support diagnostic variant interpretation. Moreover, potential genotype-phenotype associations were observed that may be interesting for more personalized healthcare of PHTS, after further evaluation of the predictive value and molecular mechanisms. We highlighted novel trends for *PTEN* variant coding effect alone and combined with PTEN domain that could partly explain the wide phenotypic diversity of PHTS, with missense variants often seen in early-onset disease and truncating variants in later-onset disease.

Funding sources

This work (L.A.J.H. and J.R.V.) was financially supported by the PTEN Research Foundation. E.R.W. and D.G.E. are supported by the NIHR Manchester Biomedical Research Centre (Grant Reference

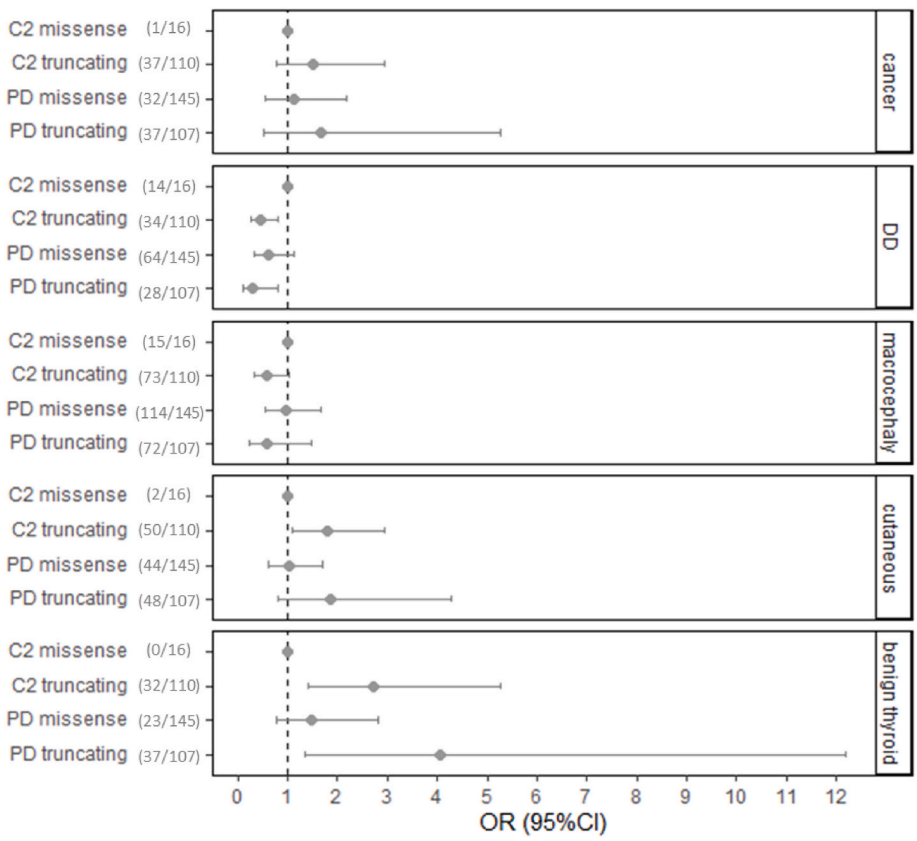

Fig. 3. Coding effect and domain. Odds ratios (OR) are presented with corresponding 95% confidence intervals (95%CI) for the logistic regression including age, sex, coding effect and domain. Results are presented without correction for multiple testing. The vertical dashed line indicates OR = 1.0. For each group, the number of patients with the corresponding phenotype (n) in the group patients with the genotypic characteristics (N) is presented (n/N). Results are presented per phenotypic category for missense variants in the C2 domain (C2 missense), truncating variants in the C2 domain (C2 truncating), missense variants in the phosphatase domain (PD missense) and truncating variants in the phosphatase domain (PD truncating), where C2 missense was the reference category.; DD = developmental delay; cutaneous = cutaneous pathology; benign thyroid = benign thyroid pathology.

Number 1215–200074). E.T. is supported by Region Stockholm (Grant ID, 2020-500306 DS). L.R. is supported by the Estonian Research Council (Grant ID PRG471).

Ethics declaration

This study received approval from the Institutional Review Board from the Radboudumc Nijmegen with a waiver for informed consent. For participating centers no written consent was required either. Informed consent was not required since only minimal personal data was used and data was de-identified for use. The study was performed in accordance with the Declaration of Helsinki.

CRediT authorship contribution statement

Linda A.J. Hendricks: Data curation, Formal analysis, Investigation, Methodology, Resources, Validation, Visualization, Writing – original draft, Writing – review & editing. **Nicoline Hoogerbrugge:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Supervision, Visualization, Writing – original draft, Writing – review & editing. **Hanka Venselaar:** Resources, Writing – review & editing. **Stefan Aretz:** Resources, Writing – review & editing. **Isabel Spier:** Resources, Writing – review & editing. **Eric Legius:** Resources, Writing – review & editing. **Hilde Brems:** Resources, Writing – review & editing. **Robin de Putter:** Resources, Writing – review & editing. **Kathleen B.M. Claes:** Resources, Writing – review & editing. **D. Gareth Evans:** Resources, Writing – review & editing. **Emma R. Woodward:** Resources, Writing – review & editing. **Maurizio Genuardi:** Resources, Writing – review & editing. **Fulvia Brugnoletti:** Resources, Writing – review & editing. **Yvette van Ierland:** Resources, Writing – review & editing. **Kim Dijke:** Resources, Writing – review & editing. **Emma Tham:** Resources, Writing – review & editing. **Bianca Tesi:** Resources, Writing – review & editing. **Janneke H.M. Schuurs-Hoeijmakers:** Resources, Writing – review & editing. **Maud Branchaud:** Resources, Writing – review & editing. **Hector Salvador:** Resources, Writing – review & editing. **Arne Jahn:** Resources, Writing – review & editing. **Simon Schnaiter:** Resources, Writing – review & editing. **Violetta Christophidou Anastasiadou:** Resources, Writing – review & editing. **Joan Brunet:** Resources, Writing – review & editing. **Carla Oliveira:** Resources, Writing – review & editing. **Laura Roht:** Resources, Writing – review & editing. **Ana Blatnik:** Resources, Writing – review & editing. **Arvids Irmejs:** Resources, Writing – review & editing. **Floor Duijkers:** Resources, Writing – review & editing. **Jacques C. Giltay:** Resources, Writing – review & editing. **Liselotte P. van Hest:** Resources, Writing – review & editing. **Tjitske Kleefstra:** Resources, Writing – review & editing. **Edward M. Leter:** Resources, Writing – review & editing. **Maartje Nielsen:** Resources, Writing – review & editing. **Sebastiaan W.R. Nijmeijer:** Resources, Writing – review & editing. **Maran J.W. Olderode-Berends:** Resources, Writing – review & editing. **Arjen R. Mensenkamp:** Resources, Writing – review & editing. **Janet R. Vos:** Data curation, Conceptualization, Formal analysis, Investigation, Methodology, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors have no conflict of interest to declare.

Data availability

The genetic data is provided in supplementary table 1. The spatio-functional classification per AA is provided in supplementary table 2. Results not shown in text are available upon request.

Table 3
Overview of identified genotype-phenotype associations in this study and literature.

Study (reference)	This study[a]	Marsh et al., 1998	Marsh et al., 1999	Nelen et al., 1999	Lachlan et al., 2007	Tan et al., 2012	Bubien et al., 2013	Nieuwenhuis et al., 2014	Frazier et al., 2015	Leslie and Longy, 2016
Cohort										
Patients, N	467	28 CS families for analyses	43 BRR cases, 37 CS families	22	42	368	146	180	295	10
Remarks	–	Family-level		–	–	–	–			Unpublished data M Longy
Methodology [b]										
Associations	OR, correlation	Correlation	Correlation	Descriptive	Descriptive	OR	Correlation	Correlation	Correlation	Descriptive
Corrections	Age, sex, multiple testing	No	No	No	No	No	No[c]	No	No	No
Genotype definition										
Coding effect	Truncating; Missense	Point missense, point nonsense, frameshift, splice site; truncating vs. non-truncating; splice site vs. non-splice site	Truncating; Non-truncating	Descriptive	Descriptive	Nonsense; Missense; Deletion; Indel; Insertion; Splice Junction; Large deletion; Promotor	Truncated; Non-truncated; Missense; Non-missense	NMD; non-NMD	Missense	Missense
Domain	C2; PD; Other	5′ of and including PTPase motif vs. 3′ of PTPase motif	5′ of or within PTPase core motif; 3′ of core motif	n.a.	n.a.	Catalytic core motif N-terminal PD; Upstream phosphatase core motif	PD; C2	PD; C2	n.a.	n.a.
Exon/intron	All introns and exons	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Other	Adult vs. pediatric	Germline variant presence vs. absence	Germline variant presence vs. absence	Descriptive of location	n.a.	Conservation	n.a.	n.a.	n.a.	n.a.
Outcome measurement	Cancer; Macrocephaly; DD; Cutaneous; Benign thyroid	Number organ sites involved; thyroid disease	Cancer; Breast fibroadenoma; intestinal polyps; speckled penis; CAL; thyroid; hemangioma	Skin; Thyroid; LDD; Breast; macrocephaly; Intestine; urogenital; neurology	CS vs. BRRS	BC; EC; TC; RC; CRC; major organ systems	Mucocutaneous; benign thyroid; breast; digestive; lipovascular; brain; BC; TC; genitourinary; cancer	Any cancer; BC; LDD; TC; EC; Skin cancer; RC; CRC; Lung cancer	Macrocephaly; ASD	Macrocephaly; Autism
Genotype-phenotype [d]	Coding effect: More **truncating** for **cutaneous** and **benign**	Coding effect: **Non-truncating** associated with	Coding effect: Correlation between **truncating** and	Coding effect: **LDD** not observed for **missense**	Coding effect: Same variant in CS and BRRS;	Coding effect: Correlation **promoter** and **BC**	Coding effect: More often benign thyroid and acral keratosis	Coding effect: Positive correlation for	Coding effect: **Missense** in patients in **ASD**	Coding effect: 8/10 mutations in patients with

(*continued on next page*)

Table 3 (*continued*)

Study (reference)	This study[a]	Marsh et al., 1998	Marsh et al., 1999	Nelen et al., 1999	Lachlan et al., 2007	Tan et al., 2012	Bubien et al., 2013	Nieuwenhuis et al., 2014	Frazier et al., 2015	Leslie and Longy, 2016
	thyroid and less for **macrocephaly** and **DD**;	**five-organ involvement** (p = 0.11)	**cancer** or **breast** fibroadenoma in BRRS (p = 0.024)		Variable phenotype within and between families with the same variant	(OR = 4.04); Correlation between **nonsense** mutation and **CRC** (OR = 4.97)	with truncating (p > 0.05)	**non-missense** and **TC** (p = 0.014)	and **macro-cephaly** (p = 0.013);	**macrocephaly** and **autism** are **missense**
	Domain: C2 truncating more often than C2 missense for **cancer**, **cutaneous** and **benign thyroid** and stronger effect for PD truncating; Less PD missense than C2 missense for DD	Domain: None	Domain: None	Domain: n.a	Domain: n.a.	Domain: None	Domain: None	Domain: None	Domain: n.a.	Domain: n.a.
	Exon/intron: None Other: More **truncating** than missense in **adults**	Exon/intron: n.a. Other: Correlation between *PTEN* mutation-positive CS families and **breast** involvement (p = 0.05–0.30)	Exon/intron: n.a. Other: Correlation between *PTEN* mutation presence and **cancer** or **breast** fibroadenoma in BRRS (p = 0.014)	Exon/intron: n.a. Other: None	Exon/intron: n.a. Other: n.a.	Exon/intron: n.a. Other: None	Exon/intron: n.a. Other: n.a.	Exon/intron: n.a. Other: n.a.	Exon/intron: n.a. Other: n.a.	Exon/intron: n.a. Other: n.a.

Abbreviations: OR = odds ratio; C2 = C2 domain; PD = phosphatase domain; n.a. = not assessed; NMD = nonsense mediated mRNA decay; DD = developmental delay; CAL = café-au-lait; LDD = Lhermitte-Duclos Disease; CS = cowden syndrome; BRR(S) = Bannayan-Riley-Ruvalcaba (syndrome); BC = breast cancer; EC = endometrial cancer; TC = thyroid cancer; RC = renal cancer; CRC = colorectal cancer; ASD = autism spectrum disorder; none = no genotype-phenotype association identified.

[a] Odds ratios (OR) are presented in this manuscript (Table 2, Fig. 3).

[b] Descriptive = no statistical tsts performed; the results were descriptive.

[c] Patients <20 years excluded for mucocutaneous analyses.

[d] Only results for identified associations are presented.

Acknowledgements

This research is supported (not financially) by the European Reference Network on Genetic Tumour Risk Syndromes (ERN GENTURIS)—Project ID No 739547. ERN GENTURIS is partly co-funded by the European Union within the framework of the Third Health Programme "ERN-2016—Framework Partnership Agreement 2017–2021".

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejmg.2022.104632.

References

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The protein Data Bank. Nucleic Acids Res. 28 (1), 235–242. https://doi.org/10.1093/nar/28.1.235.

Bubien, V., Bonnet, F., Brouste, V., Hoppe, S., Barouk-Simonet, E., David, A., Edery, P., Bottani, A., Layet, V., Caron, O., Gilbert-Dussardier, B., Delnatte, C., Dugast, C., Fricker, J.P., Bonneau, D., Sevenet, N., Longy, M., Caux, F., 2013. High cumulative risks of cancer in patients with PTEN hamartoma tumour syndrome. J. Med. Genet. 50 (4), 255–263. https://doi.org/10.1136/jmedgenet-2012-101339.

Frazier, T.W., Embacher, R., Tilot, A.K., Koenig, K., Mester, J., Eng, C., 2015. Molecular and phenotypic abnormalities in individuals with germline heterozygous PTEN mutations and autism. Mol. Psychiatr. 20 (9), 1132–1138. https://doi.org/10.1038/mp.2014.125.

Han, S.Y., Kato, H., Kato, S., Suzuki, T., Shibata, H., Ishii, S., Shiiba, K., Matsuno, S., Kanamaru, R., Ishioka, C., 2000. Functional evaluation of PTEN missense mutations using in vitro phosphoinositide phosphatase assay. Cancer Res. 60 (12), 3147–3151.

Hendricks, L.A.J., Hoogerbrugge, N., Schuurs-Hoeijmakers, J.H.M., Vos, J.R., 2021. A review on age-related cancer risks in PTEN hamartoma tumor syndrome. Clin. Genet. 99 (2), 219–225. https://doi.org/10.1111/cge.13875.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., Haussler, D., 2002. The human genome browser at UCSC. Genome Res. 12 (6), 996–1006. https://doi.org/10.1101/gr.229102.

Knudson Jr., A.G., 1971. Mutation and cancer: statistical study of retinoblastoma. Proc. Natl. Acad. Sci. U. S. A. 68 (4), 820–823. https://doi.org/10.1073/pnas.68.4.820.

Lachlan, K.L., Lucassen, A.M., Bunyan, D., Temple, I.K., 2007. Cowden syndrome and Bannayan Riley Ruvalcaba syndrome represent one condition with variable expression and age-related penetrance: results of a clinical study of PTEN mutation carriers. J. Med. Genet. 44 (9), 579–585. https://doi.org/10.1136/jmg.2007.049981.

Lee, J.O., Yang, H., Georgescu, M.M., Di Cristofano, A., Maehama, T., Shi, Y., Dixon, J.E., Pandolfi, P., Pavletich, N.P., 1999. Crystal structure of the PTEN tumor suppressor: implications for its phosphoinositide phosphatase activity and membrane association. Cell 99 (3), 323–334. https://doi.org/10.1016/s0092-8674(00)81663-3.

Leslie, N.R., Longy, M., 2016. Inherited PTEN mutations and the prediction of phenotype. Semin. Cell Dev. Biol. 52, 30–38. https://doi.org/10.1016/j.semcdb.2016.01.030.

Marsh, D.J., Coulon, V., Lunetta, K.L., Rocca-Serra, P., Dahia, P.L., Zheng, Z., Liaw, D., Caron, S., Duboué, B., Lin, A.Y., Richardson, A.L., Bonnetblanc, J.M., Bressieux, J.M., Cabarrot-Moreau, A., Chompret, A., Demange, L., Eeles, R.A., Yahanda, A.M., Fearon, E.R., Fricker, J.P., Gorlin, R.J., Hodgson, S.V., Huson, S., Lacombe, D., LePrat, F., Odent, S., Toulouse, C., Olopade, O.I., Sobol, H., Tishler, S., Woods, C.G., Robinson, B.G., Weber, H.C., Parsons, R., Peacocke, M., Longy, M., Eng, C., 1998. Mutation spectrum and genotype-phenotype analyses in Cowden disease and Bannayan-Zonana syndrome, two hamartoma syndromes with germline PTEN mutation. Hum. Mol. Genet. 7 (3), 507–515. https://doi.org/10.1093/hmg/7.3.507.

Marsh, D.J., Kum, J.B., Lunetta, K.L., Bennett, M.J., Gorlin, R.J., Ahmed, S.F., Bodurtha, J., Crowe, C., Curtis, M.A., Dasouki, M., Dunn, T., Feit, H., Geraghty, M.T., Graham Jr., J.M., Hodgson, S.V., Hunter, A., Korf, B.R., Manchester, D., Miesfeldt, S., Murday, V.A., Nathanson, K.L., Parisi, M., Pober, B., Romano, C., Tolmie, J.L., Trembath, R., Winter, R.M., Zackai, E.H., Zori, R.T., Weng, L., Dahia, P.L., Eng, C., 1999. PTEN mutation spectrum and genotype-phenotype correlations in Bannayan-Riley-Ruvalcaba syndrome suggest a single entity with Cowden syndrome. Hum. Mol. Genet. 8 (8), 1461–1472. https://doi.org/10.1093/hmg/8.8.1461.

Mester, J.L., Ghosh, R., Pesaran, T., Huether, R., Karam, R., Hruska, K.S., Costa, H.A., Lachlan, K., Ngeow, J., Barnholtz-Sloan, J., Sesock, K., Hernandez, F., Zhang, L., Milko, L., Plon, S.E., Hegde, M., Eng, C., 2018. Gene-specific criteria for PTEN variant curation: recommendations from the ClinGen PTEN expert panel. Hum. Mutat. 39 (11), 1581–1592. https://doi.org/10.1002/humu.23636.

Mighell, T.L., Evans-Dutson, S., O'Roak, B.J., 2018. A saturation mutagenesis approach to understanding PTEN lipid phosphatase activity and genotype-phenotype relationships. Am. J. Hum. Genet. 102 (5), 943–955. https://doi.org/10.1016/j.ajhg.2018.03.018.

Mighell, T.L., Thacker, S., Fombonne, E., Eng, C., O'Roak, B.J., 2020. An integrated deep-mutational-scanning approach provides clinical insights on PTEN genotype-phenotype relationships. Am. J. Hum. Genet. 106 (6), 818–829. https://doi.org/10.1016/j.ajhg.2020.04.014.

Mondal, S.K., Sen, M.K., 2020. Loss of phosphatase activity in PTEN (phosphatase and tensin homolog deleted on chromosome ten) results in endometrial carcinoma in humans: an in-silico study. Heliyon 6 (1), e03106. https://doi.org/10.1016/j.heliyon.2019.e03106.

Nagy, E., Maquat, L.E., 1998. A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. Trends Biochem. Sci. 23 (6), 198–199. https://doi.org/10.1016/s0968-0004(98)01208-0.

Nelen, M.R., Kremer, H., Konings, I.B., Schoute, F., van Essen, A.J., Koch, R., Woods, C. G., Fryns, J.P., Hamel, B., Hoefsloot, L.H., Peeters, E.A., Padberg, G.W., 1999. Novel PTEN mutations in patients with Cowden disease: absence of clear genotype-phenotype correlations. Eur. J. Hum. Genet. 7 (3), 267–273. https://doi.org/10.1038/sj.ejhg.5200289.

Nieuwenhuis, M.H., Kets, C.M., Murphy-Ryan, M., Yntema, H.G., Evans, D.G., Colas, C., Møller, P., Hes, F.J., Hodgson, S.V., Olderode-Berends, M.J., Aretz, S., Heinimann, K., Gómez García, E.B., Douglas, F., Spigelman, A., Timshel, S., Lindor, N.M., Vasen, H.F., 2014. Cancer risk and genotype-phenotype correlations in PTEN hamartoma tumor syndrome. Fam. Cancer 13 (1), 57–63. https://doi.org/10.1007/s10689-013-9674-3.

Papa, A., Wan, L., Bonora, M., Salmena, L., Song, M.S., Hobbs, R.M., Lunardi, A., Webster, K., Ng, C., Newton, R.H., Knoblauch, N., Guarnerio, J., Ito, K., Turka, L.A., Beck, A.H., Pinton, P., Bronson, R.T., Wei, W., Pandolfi, P.P., 2014. Cancer-associated PTEN mutants act in a dominant-negative manner to suppress PTEN protein function. Cell 157 (3), 595–610. https://doi.org/10.1016/j.cell.2014.03.027.

Portelli, S., Barr, L., de Sá, A.G.C., Pires, D.E.V., Ascher, D.B., 2021. Distinguishing between PTEN clinical phenotypes through mutation analysis. Comput. Struct. Biotechnol. J. 19, 3097–3109. https://doi.org/10.1016/j.csbj.2021.05.028.

Smith, I.N., Thacker, S., Jaini, R., Eng, C., 2019. Dynamics and structural stability effects of germline PTEN mutations associated with cancer versus autism phenotypes. J. Biomol. Struct. Dyn. 37 (7), 1766–1782. https://doi.org/10.1080/07391102.2018.1465854.

Song, M.S., Carracedo, A., Salmena, L., Song, S.J., Egia, A., Malumbres, M., Pandolfi, P. P., 2011. Nuclear PTEN regulates the APC-CDH1 tumor-suppressive complex in a phosphatase-independent manner. Cell 144 (2), 187–199. https://doi.org/10.1016/j.cell.2010.12.020.

Spinelli, L., Black, F.M., Berg, J.N., Eickholt, B.J., Leslie, N.R., 2015. Functionally distinct groups of inherited PTEN mutations in autism and tumour syndromes. J. Med. Genet. 52 (2), 128–134. https://doi.org/10.1136/jmedgenet-2014-102803.

Tan, M.H., Mester, J., Peterson, C., Yang, Y., Chen, J.L., Rybicki, L.A., Milas, K., Pederson, H., Remzi, B., Orloff, M.S., Eng, C., 2011. A clinical scoring system for selection of patients for PTEN mutation testing is proposed on the basis of a prospective study of 3042 probands. Am. J. Hum. Genet. 88 (1), 42–56. https://doi.org/10.1016/j.ajhg.2010.11.013.

Tan, M.H., Mester, J.L., Ngeow, J., Rybicki, L.A., Orloff, M.S., Eng, C., 2012. Lifetime cancer risks in individuals with germline PTEN mutations. Clin. Cancer Res. 18 (2), 400–407. https://doi.org/10.1158/1078-0432.Ccr-11-2283.

Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C.A., Odeberg, J., Djureinovic, D., Takanen, J.O., Hober, S., Alm, T., Edqvist, P.H., Berling, H., Tegel, H., Mulder, J., Rockberg, J., Nilsson, P., Schwenk, J.M., Hamsten, M., von Feilitzen, K., Forsberg, M., Persson, L., Johansson, F., Zwahlen, M., von Heijne, G., Nielsen, J., Pontén, F., 2015. Proteomics. Tissue-based map of the human proteome. Science 347 (6220), 1260419. https://doi.org/10.1126/science.1260419.

Varga, E.A., Pastore, M., Prior, T., Herman, G.E., McBride, K.L., 2009. The prevalence of PTEN mutations in a clinical pediatric cohort with autism spectrum disorders, developmental delay, and macrocephaly. Genet. Med. 11 (2), 111–117. https://doi.org/10.1097/GIM.0b013e31818fd762.

Yehia, L., Ngeow, J., Eng, C., 2019. PTEN-opathies: from biological insights to evidence-based precision medicine. J. Clin. Invest. 129 (2), 452–464. https://doi.org/10.1172/jci121277.

Yehia, L., Seyfi, M., Niestroj, L.M., Padmanabhan, R., Ni, Y., Frazier, T.W., Lal, D., Eng, C., 2020. Copy number variation and clinical outcomes in patients with germline PTEN mutations. JAMA Netw. Open 3 (1), e1920415. https://doi.org/10.1001/jamanetworkopen.2019.20415.