NR2F1 database: 112 variants and 84 patients support refining the clinical synopsis of Bosch-Boonstra-Schaaf optic atrophy syndrome

Billiet, B.; Amati-Bonneau, P.; Desquiret-Dumas, V.; Guehlouz, K.; Milea, D.; Gohier, P.; ... ; Ferre, M.

Received: 21 June 2021 | Revised: 12 October 2021 | Accepted: 16 November 2021

DOI: 10.1002/humu.24305

DATABASES

Human Mutation HGVS HUMAN GENOME VARIATION SOCIETY WILEY

NR2F1 database: 112 variants and 84 patients support refining the clinical synopsis of Bosch–Boonstra–Schaaf optic atrophy syndrome

Benjamin Billiet[1] | Patrizia Amati-Bonneau[2,3] | Valérie Desquiret-Dumas[2,3] | Khadidja Guehlouz[1] | Dan Milea[4] | Philippe Gohier[1] | Guy Lenaers[2] | Delphine Mirebeau-Prunier[2,3] | Johan T. den Dunnen[5] | Pascal Reynier[2,3] | Marc Ferré[2]

[1]Département d'Ophtalmologie, Centre Hospitalier Universitaire d'Angers, Angers, France

[2]Unité MITOVASC, Équipe Mitolab, SFR ICAT, INSERM, CNRS, Université d'Angers, Angers, France

[3]Laboratoire de Biochimie et Biologie moléculaire, Centre Hospitalier Universitaire d'Angers, Angers, France

[4]Singapore Eye Research Institute, Singapore National Eye Centre, Duke-NUS, Singapore

[5]Department of Human Genetics, Department of Clinical Genetics, Leiden University Medical Centre, Leiden, The Netherlands

Correspondence
Marc Ferré, Laboratoire de Biochimie et Biologie moléculaire, 4 rue Larrey, 49100 Angers, France.
Email: marc.ferre@univ-angers.fr

Abstract

Pathogenic variants of the nuclear receptor subfamily 2 group F member 1 gene (*NR2F1*) are responsible for Bosch–Boonstra–Schaaf optic atrophy syndrome (BBSOAS), an autosomal dominant disorder characterized by optic atrophy associated with developmental delay and intellectual disability, but with a clinical presentation which appears to be multifaceted. We created the first public locus-specific database dedicated to *NR2F1*. All variants and clinical cases reported in the literature, as well as new unpublished cases, were integrated into the database using standard nomenclature to describe both molecular and phenotypic anomalies. We subsequently pursued a comprehensive approach based on computed representation and analysis suggesting a refinement of the BBSOAS clinical description with respect to neurological features and the inclusion of additional signs of hypotonia and feeding difficulties. This database is fully accessible for both clinician and molecular biologists and should prove useful in further refining the clinical synopsis of *NR2F1* as new data is recorded.

KEYWORDS

BBSOAS, Bosch–Boonstra–Schaaf optic atrophy syndrome, COUP transcription factor 1 protein, COUP-TF1, database, neurodegenerative disorders, *NR2F1*, nuclear receptor subfamily 2 group F member 1, ontology, optic atrophy

1 | BACKGROUND

The nuclear receptor subfamily 2 group F member 1 gene (*NR2F1*; MIM# 132890), consisting of three exons, encodes the 423 amino acids (aa) of the COUP transcription factor 1 protein (COUP-TF1; Swiss-Prot:COT1_HUMAN). It belongs to the superfamily of the steroid/thyroid hormone receptors and is involved in the development of several brain structures, including the neocortex, hippocampus, and ganglionic eminences, as it has been shown in mice (Alfano et al., 2011; Armentano et al., 2006; Bertacchi et al., 2019). Variants in *NR2F1* are responsible for Bosch–Boonstra–Schaaf optic atrophy syndrome (BBSOAS; MIM# 615722), an autosomal dominant disorder characterized by optic atrophy associated with developmental delay and intellectual disability (Bojanek et al., 2020; Bosch et al., 2014; Walsh et al., 2020).

Most pathogenic variants are de novo and dominant (Al-Kateb et al., 2013; Balciuniene et al., 2019; Bertacchi et al., 2020; Bojanek et al., 2020; Bosch et al., 2014; Bosch et al., 2016; Brown et al., 2009; Chen et al., 2016; Dimassi et al., 2016; Eldomery et al., 2017; Hino-Fukuyo et al., 2017; Hobbs et al., 2020; Jezela-Stanek et al., 2020;

Kaiwar et al., 2017; Martín-Hernández et al., 2018; Mio et al., 2020; Park et al., 2019; Rech et al., 2020; Starosta et al., 2020; Walsh et al., 2020; Zou et al., 2020), with some rare familial cases (Chen et al., 2016). During the follow-up, the neurological features are not degenerative and no death of a patient with BBSAOS was reported. The phenotype of BBSOAS is heterogeneous, as some patients do not suffer from visual impairment and others display no neurodevelopmental delay.

In this article, we describe the construction of the first *NR2F1* locus-specific database (LSDB), listing all patients and genetic variants referenced in the literature and unpublished cases from our laboratory, as well as computed data suggesting refinement of the BBSOAS clinical synopsis.

2 | MATERIALS AND METHODS

2.1 | Nomenclature

All names, symbols, and Online Mendelian Inheritance in Man (OMIM) database numbers were checked for correspondence with current official names indicated by the Human Genome Organization (HUGO) Gene Nomenclature Committee (Gray et al., 2013) and the OMIM database (Hamosh et al., 2000). The phenotype descriptions are based on the Human Phenotype Ontology (HPO), indicating the HPO term name and identifier (Köhler et al., 2019).

NR2F1 variants are described according to both the NCBI genomic reference sequence (RefSeq:NG_034119.1) and transcript reference sequence (RefSeq:NM_005654.4), including the reference sequence for three exons encoding a protein of 423 aa (RefSeq:NP_005645.1) (O'Leary et al., 2016). The numbering of the nucleotides reflects that of the complementary DNA (cDNA), with "+1" corresponding to the "A" of the ATG translation initiation codon in the reference sequence, according to which the initiation codon is codon 1, as recommended by version 2.0 of the nomenclature of the Human Genome Variation Society (HGVS) (den Dunnen et al., 2016).

Although the current official HGVS recommendations prescribe description beyond cDNA, we have indicated them for clarity by following a guideline envisaged for the open issue; the notation "c.−1687_*240{0}" indicates the absence of the entire *NR2F1* gene, from the first nucleotide of the first exon to the last nucleotide of the last exon, the limits of which are beyond and not precisely defined.

Information concerning changes in RNA and protein levels have been added from the original papers or predicted from DNA variants if not experimentally studied. Following the HGVS guidelines, deduced changes are indicated between brackets. Protein domains were predicted according to InterPro version 85.0—April 8, 2021 (Blum et al., 2021).

2.2 | Implementation of the database

The *NR2F1* database belongs to the Global Variome shared Leiden Open-source Variation Database (LOVD), currently running under LOVD v.3.0 Build 26c (I. F. Fokkema et al., 2011), following the guidelines for LSDBs (Vihinen et al., 2012) and hosted under the responsibility of the Global Variome/Human Variome Project (Cotton et al., 2008). The database reviews clinical and molecular data from patients carrying *NR2F1* variants published in peer-reviewed literature as well as unpublished contributions that are directly submitted. If there are inaccuracies or an obsolete convention is used, the *"DNA published"* field of the page dedicated to each variant indicates whether the published name has been modified by the curator. The *NR2F1* LSDB website requires full compliance with the rules set out above for the description of sequence variants to provide uniform and comparable data.

2.3 | Data collection and analysis

The causative variants were collected from the literature published to date (May 2021) using the NCBI PubMed search tool (Sayers et al., 2010). The positions of variants in the reference transcripts were determined and updated according to the HGVS nomenclature version 2.0 (den Dunnen et al., 2016). Correct naming at the nucleotide and amino acid levels was verified and reestablished when necessary using the Mutalyzer 2.0.34 *Syntax Checker* (Wildeman et al., 2008). Information on the number of patients carrying each causative variant, as well as their geographical origins and the homo- or heterozygosity, was taken from the original or review publications, as well as from data collected from our clinical laboratory. If the same patient or variant is reported in more than one article, then it is recorded only once in the database with reference to the first publication. Further information on the genetic origin of the allele, segregation with the disease phenotype, and frequency in the control population was recorded. The criteria of pathogenicity, which depend upon the clinical context and molecular findings, are stated under the heading "Clinical classification" for the classification of the variant based on standardized criteria

The clinical classification of variants is based on standardized criteria directed on the clinical consequences as published or submitted, indicated using an enriched system including inheritance—for example, pathogenic (dominant), likely pathogenic, VUS (variant of unknown significance), likely benign, benign—derived from the American College of Medical Genetics and Genomics (ACMG) standards and guidelines (Richards et al., 2015).

We produced a data set from the *NR2F1* LSDB version NR2F1:211007 (last updated on October 7, 2021) to carry out the statistical analysis. The HPO terms have been checked and prepared using the suite of R packages ontologyIndex version 2.5 and ontologyPlot version 1.4 (Greene et al., 2017), within R version 4.0.5 (R Core Team, 2020), to read in the OBO file version hp/releases/2021-04-13 (Köhler et al., 2019). Hierarchical clustering is performed using the *hclust* function from the R-Core package (R Core Team, 2020). Our data were crossed with the following external sources: gnomAD v.2.1.1 (Karczewski et al., 2020); OMIM version updated May 20, 2021 (Hamosh et al., 2000); Orphanet version 5.46.0 (INSERM, 1997).

2.4 | Data access and submission

The *NR2F1* database is an open database allowing any researcher or clinician to consult the contents freely without prior registration or to contribute new data after registration to ensure traceability. The database can be accessed on the World Wide Web at: https://www.lovd.nl/NR2F1 (through the Global Variome shared LOVD server or through the MITOchondrial DYNamics variation portal at: http://nr2f1.mitodyn.org/). The data can also be retrieved via an application programming interface (API), that is, a web service allowing simple queries and retrieval of basic genes and variants information (documentation available on the web page of the database), as well as serving as a public beacon in The Global Alliance for Genomics and Health Beacon Project (Global Alliance for Genomics and Health, 2016). General information is available on the database home page. The process for submitting data begins by clicking the "Submit" tab. Data concerning patients and variants may be retrieved using the standard LOVD tabs, named "Individuals" and "Variants," respectively.

3 | RESULTS AND DISCUSSION

The *NR2F1* database contains three main interconnected tables: the "Individuals" table contains details of the patient examined, including gender, geographic origin, and patient identification, if applicable; the "Phenotype" table indicates the clinical phenotypic features, described according to the root of the phenotypic abnormality subontology (HPO# HP:0000118); and the "Variants" table includes information about the sequence variations at the genomic (DNA) and transcript (cDNA) levels, as well as the clinical classification for each variant.

3.1 | Genotypic data

To date, the database contains 112 sequence variants records. Excluding duplicates, 83 unique variants are listed (Figures 1 and S1 and Table 1), of which 83% (69) are considered pathogenic sequence variants in a dominant condition, 12% (10) are of unknown significance and 5% (4) are benign (Al-Kateb et al., 2013; Balciuniene et al., 2019; Bertacchi et al., 2020; Bojanek et al., 2020; Bosch et al., 2014, 2016; Brown et al., 2009; Chen et al., 2016; Dimassi et al., 2016; Eldomery et al., 2017; I. Fokkema et al., 2019; Hino-Fukuyo et al., 2017; Hobbs et al., 2020; Jezela-Stanek et al., 2020; Kaiwar et al., 2017; Martín-Hernández et al., 2018; Mio et al., 2020; Park et al., 2019; Rech et al., 2020; Starosta et al., 2020; Vissers et al., 2017; Walsh et al., 2020; Zou et al., 2020). The variants considered pathogenic affect exclusively the coding sequence of the gene—no intron being involved except in the case of whole gene deletions—and are particularly overrepresented in the DNA-binding domain of the protein that is composed of two C4-type zinc fingers (NR C4-type). This domain, which represents 18% of the protein (76 aa) and spans the second half of the coding sequence of exon 1, accounts for half of the pathogenic variants, whereas the nuclear hormone receptor ligand-binding domain (NR LBD), which represents more than half of the protein (54%, 227 aa), together with the remaining nonspecific regions of the protein (28%, 120 aa), accounts for as many pathogenic variants, highlighting the importance of the NR C4-type domain in COUP-TF1 function (Figure 2a). In contrast, variants considered benign were reported only in the NR LBD or nonspecific regions.

Among the most frequently observed pathogenic effects on the protein encoded by *NR2F1*, 61% are missense variants, 14% are nonsense and 12% are frameshift variants leading to a premature protein truncation, that is, a quarter of the variations would result in truncated proteins or their absence due to the nonsense-mediated

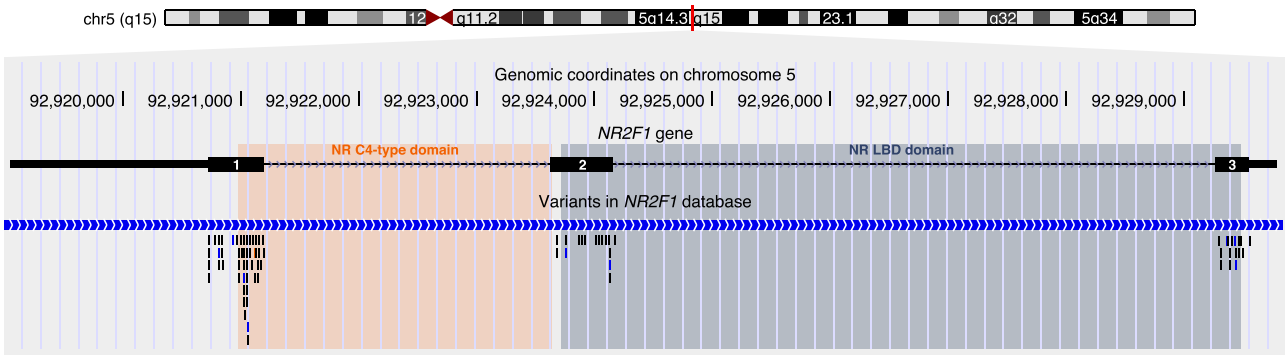

FIGURE 1 Distribution of the 83 unique genomic variants in the *NR2F1* database (compact view). Complete deletions of the gene that extends beyond the exons are shown as an extended bar with rafters, substitutions as black bars, deletions as blue bars, and the duplication as an orange bar. From the top are reported: an ideogram showing the cytogenetic localization (5q15); the genomic coordinates on human chromosome 5 (region shown extending over 10,846 bp, between positions 92,918,993 and 92,929,838 according to assembly GRCh37/hg19); and *NR2F1* gene structure including exon numbering and domains. The full view detailing the names of each variant is available in Figure S1. NR C4-type domain: DNA-binding domain that is composed of two C4-type zinc fingers; NR LBD domain: nuclear hormone receptor ligand-binding domain. Adapted from UCSC Genome Browser (http://genome.ucsc.edu) with the *NR2F1* database custom track; data as of October 7, 2021

TABLE 1 Unique variants listed in the *NR2F1* database (count: 83)

Region	DNA variant[a]	Effect on protein[a,b]	Classification of variant[c]	Domain affected[d]	Rep.[e]	Database-ID[f]	References[g]
Whole gene/ beyond	c.−506041_*1935315del	p.0	Pathogenic (dominant)	NR C4-type, NR LBD	1	NR2F1_000063	RE20
Whole gene/ beyond	c.−325733_*639854del	p.0	Pathogenic (dominant)	NR C4-type, NR LBD	1	NR2F1_000064	RE20
Whole gene/ beyond	c.−1687_*240{0}	p.0	Pathogenic (dominant)	NR C4-type, NR LBD	9	NR2F1_000006, NR2F1_000043, NR2F1_000045, NR2F1_000049, NR2F1_000050, NR2F1_000062	AL13, BO14, BR09, CH16
Exon 1/start codon	c.1A>G	p.0	Pathogenic (dominant)	-	1	NR2F1_000065	RE20
Exon 1/start codon	c.2T>C	p.0	Pathogenic (dominant)	-	4	NR2F1_000048	BE20, CH16, RE20
Exon 1/start codon	c.2T>G	p.0	Pathogenic (dominant)	-	2	NR2F1_000052	CH16
Exon 1/start codon	c.2_4delinsGGA	p.0	Pathogenic (dominant)	-	1	NR2F1_000042	CH16
Exon 1	c.49G>C	p.(Gly17Arg)	VUS	-	1	NR2F1_000015	*FO19*
Exon 1	c.73C>G	p.(Pro25Ala)	VUS	-	1	NR2F1_000016	*FO19*
Exon 1	c.78_96del	p.(Gln28Alafs*85)	Likely pathogenic (dominant)	-	1	NR2F1_000033	-
Exon 1	c.82C>T	p.(Gln28*)	Likely pathogenic (dominant)	-	1	NR2F1_000037	BO20
Exon 1	c.103_113delinsCGCCGCCGC	p.(Gly35Argfs*361)	Likely pathogenic (dominant)	-	1	NR2F1_000053	CH16
Exon 1	c.107G>C	p.(Gly36Ala)	VUS	-	1	NR2F1_000022	*FO19*
Exon 1	c.115G>T	p.(Glu39*)	Likely pathogenic (dominant)	-	1	NR2F1_000054	BE20
Exon 1	c.192del	p.(Gly65Alafs*54)	Likely pathogenic (dominant)	-	1	NR2F1_000092	-
Exon 1	c.237G>C	p.(Gln79His)	Likely pathogenic (dominant)	-	3	NR2F1_000008	*FO19*
Exon 1	c.253G>T	p.(Glu85*)	Likely pathogenic (dominant)	NR C4-type	1	NR2F1_000084	HO20
Exon 1	c.256T>C	p.(Cys86Arg)	Likely pathogenic (dominant)	NR C4-type	1	NR2F1_000067	RE20

(Continues)

TABLE 1 (Continued)

Region	DNA variant[a]	Effect on protein[a,b]	Classification of variant[c]	Domain affected[d]	Rep.[e]	Database-ID[f]	References[g]
Exon 1	c.257G>T	p.(Cys86Phe)	Likely pathogenic (dominant)	NR C4-type	1	NR2F1_000039	KA17
Exon 1	c.262G>A	p.(Val88Met)	Likely pathogenic (dominant)	NR C4-type	1	NR2F1_000068	RE20
Exon 1	c.284G>T	p.(Gly95Val)	Likely pathogenic (dominant)	NR C4-type	2	NR2F1_000009	*FO19*, RE20
Exon 1	c.286A>G	p.(Lys96Glu)	Pathogenic (dominant)	NR C4-type	1	NR2F1_000040	MA18
Exon 1	c.289C>T	p.(His97Tyr)	Likely pathogenic (dominant)	NR C4-type	2	NR2F1_000029	-
Exon 1	c.290A>C	p.(His97Pro)	Likely pathogenic (dominant)	NR C4-type	1	NR2F1_000069	RE20
Exon 1	c.291del	p.(Tyr98Thrfs*21)	Likely pathogenic (dominant)	NR C4-type	1	NR2F1_000055	CH16
Exon 1	c.292T>C	p.(Tyr98His)	Likely pathogenic (dominant)	NR C4-type	1	NR2F1_000056	BE20
Exon 1	c.293A>G	p.(Tyr98Cys)	Likely pathogenic (dominant)	NR C4-type	1	NR2F1_000070	RE20
Exon 1	c.311A>G	p.(Glu104Gly)	Likely pathogenic (dominant)	NR C4-type	1	NR2F1_000071	RE20
Exon 1	c.313G>A	p.(Gly105Ser)	Likely pathogenic (dominant)	NR C4-type	2	NR2F1_000034	MI20
Exon 1	c.314G>A	p.(Gly105Asp)	Pathogenic (dominant)	NR C4-type	2	NR2F1_000017	*FO19*, VI17
Exon 1	c.319A>G	p.(Lys107Glu)	Pathogenic (dominant)	NR C4-type	1	NR2F1_000035	ST20
Exon 1	c.320A>G	p.(Lys107Arg)	VUS	NR C4-type	1	NR2F1_000023	*FO19*
Exon 1	c.323G>A	p.(Ser108Asn)	Likely pathogenic (dominant)	NR C4-type	2	NR2F1_000080	-
Exon 1	c.323G>T	p.(Ser108Ile)	Likely pathogenic (dominant)	NR C4-type	1	NR2F1_000072	RE20
Exon 1	c.328_330del	p.(Phe110del)	Likely pathogenic (dominant)	NR C4-type	2	NR2F1_000057	CH16, DI16
Exon 1	c.335G>A	p.(Arg112Lys)	Pathogenic (dominant)	NR C4-type	1	NR2F1_000004	BO14

TABLE 1 (Continued)

Region	DNA variant[a]	Effect on protein[a,b]	Classification of variant[c]	Domain affected[d]	Rep.[e]	Database-ID[f]	References[g]
Exon 1	c.339C>A	p.(Ser113Arg)	Pathogenic (dominant)	NR C4-type	2	NR2F1_000001	BO14, BO16, *FO19*
Exon 1	c.344G>C	p.(Arg115Pro)	Pathogenic (dominant)	NR C4-type	1	NR2F1_000002	BO14
Exon 1	c.365G>C	p.(Cys122Ser)	Likely pathogenic (dominant)	NR C4-type	1	NR2F1_000073	RE20
Exon 1	c.366C>G	p.(Cys122Trp)	Likely pathogenic (dominant)	NR C4-type	1	NR2F1_000086	BA19
Exon 1	c.380dup	p.(Asn127Lysfs*270)	Likely pathogenic (dominant)	NR C4-type	1	NR2F1_000074	RE20
Exon 1	c.382T>C	p.(Cys128Arg)	Likely pathogenic (dominant)	NR C4-type	1	NR2F1_000041	CH16
Exon 1	c.403C>A	p.(Arg135Ser)	Likely pathogenic (dominant)	NR C4-type	1	NR2F1_000075	CH16
Exon 1	c.403C>T	p.(Arg135Cys)	Likely pathogenic (dominant)	NR C4-type	2	NR2F1_000018	*FO19*, HI17
Exon 1	c.413G>A	p.(Cys138Tyr)	Pathogenic (dominant)	NR C4-type	2	NR2F1_000007	CH16, EL17
Exon 1	c.417A>T	p.(Gln139His)	Likely pathogenic (dominant)	NR C4-type	1	NR2F1_000076	RE20
Exon 1	c.425G>A	p.(Arg142His)	Likely pathogenic (dominant)	NR C4-type	1	NR2F1_000059	BE20
Exon 1	c.425G>T	p.(Arg142Leu)	Likely pathogenic (dominant)	NR C4-type	1	NR2F1_000058	CH16
Exon 1	c.436T>C	p.(Cys146Arg)	Likely pathogenic (dominant)	NR C4-type	2	NR2F1_000010	CH16, *FO19*
Exon 1	c.452T>C	p.(Met151Thr)	VUS	NR C4-type	1	NR2F1_000011	*FO19*
Exon 1	c.463G>A	p.(Ala155Thr)	Likely pathogenic (dominant)	NR C4-type	2	NR2F1_000024	CH16, *FO19*
Exon 2	c.513C>A	p.(Tyr171*)	Likely pathogenic (dominant)	-	1	NR2F1_000012	*FO19*
Exon 2	c.513C>G	p.(Tyr171*)	Likely pathogenic (dominant)	-	1	NR2F1_000038	PA19
Exon 2	c.602C>A	p.(Ser201*)	Pathogenic (dominant)	NR LBD	1	NR2F1_000085	ZO20

(Continues)

TABLE 1 (Continued)

Region	DNA variant[a]	Effect on protein[a,b]	Classification of variant[c]	Domain affected[d]	Rep.[e]	Database-ID[f]	References[g]
Exon 2	c.603_606del	p.(Arg202Thrfs*154)	Pathogenic (dominant)	NR LBD	1	NR2F1_000088	*FO19*
Exon 2	c.708C>T	p.(Asn236=)	Likely benign	NR LBD	1	NR2F1_000030	*FO19*
Exon 2	c.729_730delinsCT	p.(Gln244*)	Likely pathogenic (dominant)	NR LBD	1	NR2F1_000060	BE20
Exon 2	c.755T>C	p.(Leu252Pro)	Pathogenic (dominant)	NR LBD	1	NR2F1_000003	BO14
Exon 2	c.854C>A	p.(Ser285*)	Pathogenic (dominant)	NR LBD	1	NR2F1_000090	-
Exon 2	c.883T>C	p.(Phe295Leu)	VUS	NR LBD	1	NR2F1_000081	-
Exon 2	c.909G>C	p.(Gln303His)	VUS	NR LBD	1	NR2F1_000005	BO14
Exon 2	c.931G>C	p.(Ala311Pro)	Likely pathogenic (dominant)	NR LBD	1	NR2F1_000077	RE20
Exon 2	c.954G>C	p.(Glu318Asp)	Pathogenic (dominant)	NR LBD	2	NR2F1_000019	*FO19*, RE20
Exon 2	c.965T>A	p.(Leu322His)	Likely pathogenic (dominant)	NR LBD	1	NR2F1_000091	-
Exon 2	c.968A>C	p.(Lys323Thr)	Likely pathogenic (dominant)	NR LBD	1	NR2F1_000020	*FO19*
Exon 2	c.968_969del	p.(Lys323Serfs*73)	Likely pathogenic (dominant)	NR LBD	1	NR2F1_000061	BE20
Intron 2	c.991+19G>A	p.(=)	Benign	NR LBD	1	NR2F1_000013	*FO19*
Exon 3	c.1016C>T	p.(Ala339Val)	Likely benign	NR LBD	1	NR2F1_000014	*FO19*
Exon 3	c.1024G>A	p.(Glu342Lys)	Likely pathogenic (dominant)	NR LBD	1	NR2F1_000082	-
Exon 3	c.1025A>G	p.(Glu342Gly)	VUS	NR LBD	1	NR2F1_000025	*FO19*
Exon 3	c.1083del	p.(Asn362Thrfs*33)	Likely pathogenic (dominant)	NR LBD	1	NR2F1_000051	WA20
Exon 3	c.1096C>T	p.(Arg366Cys)	Likely pathogenic (dominant)	NR LBD	1	NR2F1_000026	*FO19*
Exon 3	c.1103G>A	p.(Gly368Asp)	Likely pathogenic (dominant)	NR LBD	1	NR2F1_000078	CH16
Exon 3	c.1115T>C	p.(Leu372Pro)	Likely pathogenic (dominant)	NR LBD	1	NR2F1_000079	KA17

TABLE 1 (Continued)

Region	DNA variant[a]	Effect on protein[a,b]	Classification of variant[c]	Domain affected[d]	Rep.[e]	Database-ID[f]	References[g]
Exon 3	c.1117C>T	p.(Arg373*)	Pathogenic (dominant)	NR LBD	2	NR2F1_000021	*FO19*, RE20
Exon 3	c.1147_1149del	p.(Ser383del)	VUS	NR LBD	2	NR2F1_000031	*FO19*
Exon 3	c.1158G>T	p.(Glu386Asp)	VUS	NR LBD	1	NR2F1_000027	*FO19*
Exon 3	c.1168_1170del	p.(Phe390del)	Likely pathogenic (dominant)	NR LBD	1	NR2F1_000087	-
Exon 3	c.1183G>T	p.(Gly395Cys)	Likely pathogenic (dominant)	NR LBD	1	NR2F1_000032	*FO19*
Exon 3	c.1184G>A	p.(Gly395Asp)	Likely pathogenic (dominant)	NR LBD	1	NR2F1_000089	*FO19*
Exon 3	c.1198G>T	p.(Glu400*)	Likely pathogenic (dominant)	NR LBD	1	NR2F1_000083	-
Exon 3	c.1217T>C	p.(Met406Thr)	Likely pathogenic (dominant)	NR LBD	2	NR2F1_000036	JE20, RE20
3′-UTR[h]	c.*9C>T	p.(=)	Likely benign	-	1	NR2F1_000028	*FO19*

Note: Data as of May 7, 2021; for more information, please refer to the database using Database-ID.

[a]Mutational data are described using the nomenclature of the Human Genome Variation Society (http://www.hgvs.org/mutnomen). Nucleotide numbering reflects cDNA numbering with "+1" corresponding to the A of the ATG translation initiation codon in the *NR2F1* reference sequence (RefSeq:NM_005654.4), according to journal guidelines. The initiation codon is codon 1. The notation out of nomenclature "c.−1687_*240{0}" indicates the absence of the entire *NR2F1* gene with limits not precisely defined.

[b]Predicted effect on protein based on clinical consequences.

[c]Classification of variant based on clinical consequences, using standardized criteria: pathogenic (disease-associated), likely pathogenic (likely disease-associated), VUS (variant of unknown significance), likely benign (likely not disease-associated), benign (not disease-associated); including inheritance (dominant or recessive) if applicable.

[d]Affected domain of the protein: zinc finger, nuclear hormone receptor-type (NR C4-type; from amino acid 83–158, InterPro:IPR001628); nuclear hormone receptor, ligand-binding domain (NR LBD; from amino acid 184–410, InterPro:IPR000536).

[e]Number of times variant has been reported in the database.

[f]Identifier of variant in the *NR2F1* database (https://www.lovd.nl/NR2F1).

[g]Publications describing the variant submitted: AL13 (Al-Kateb et al., 2013), BA19 (Balciuniene et al., 2019), BE20 (Bertacchi et al., 2020), BO14 (Bosch et al., 2014), BO16 (Bosch et al., 2016), BO20 (Bojanek et al., 2020), BR09 (Brown et al., 2009), CH16 (Chen et al., 2016), DI16 (Dimassi et al., 2016), EL17 (Eldomery et al., 2017), *FO19* (I. Fokkema et al., 2019), HI17 (Hino-Fukuyo et al., 2017), HO20 (Hobbs et al., 2020), JE20 (Jezela-Stanek et al., 2020), KA17 (Kaiwar et al., 2017), MA18 (Martín-Hernández et al., 2018), MI20 (Mio et al., 2020), PA19 (Park et al., 2019), RE20 (Rech et al., 2020), ST20 (Starosta et al., 2020), VI17 (Vissers et al., 2017), WA20 (Walsh et al., 2020), ZO20 (Zou et al., 2020); the variants referenced *FO19* are classification records from genomic diagnostic laboratories, that is, there is no associated patient information; the character "-" indicates a submission to the database without publication.

[h]3′-untranslated region.

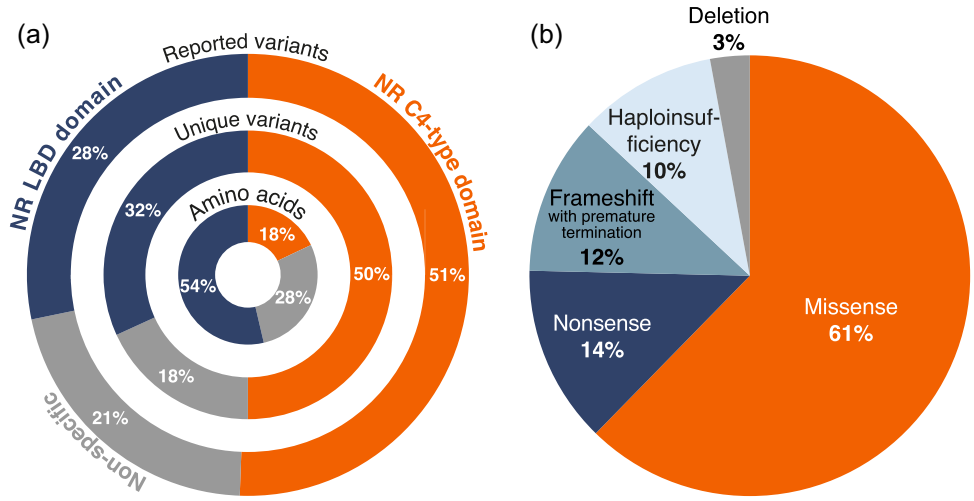

FIGURE 2 Distribution of effects of variants on the protein classified as pathogenic or probably pathogenic in the *NR2F1* database. (a) Comparison for each region of the protein of its size on the protein sequence in amino acids (amino acids), of the count of unique variants (i.e., by counting only once several reported cases of the same variant; Unique variants), and of the count of variants reports in the database (i.e., by counting each of the reported cases of the same variant; Reported variants). (b) Effects of the unique variants on the protein. NR C4-type domain: DNA-binding domain that is composed of two C4-type zinc fingers; NR LBD domain: nuclear hormone receptor ligand-binding domain; Nonspecific: regions of the protein that are not domain specific. Data as of October 7, 2021

mRNA decay surveillance pathway; 10% lead to haploinsufficiency, either due to the complete deletion of the whole gene or as a consequence of a variant in the translation initiation codon on one allele; two variants are a deletion of a single aa (Figure 2b). Although only a few variants are recurrent, the large deletions of *NR2F1* encompassing the cDNA boundaries have been significantly more frequently reported, with 11 records (Table 1) (Al-Kateb et al., 2013; Bosch et al., 2014; Brown et al., 2009; Chen et al., 2016; Rech et al., 2020).

The data from the Genome Aggregation Database (gnomAD), which is the aggregation of the high-quality exome (protein-coding region) DNA sequence data for about 140,000 individuals (Karczewski et al., 2020), has been integrated into the Global Variome shared LOVD server. However, it was decided to only indicate the frequency reported in gnomAD for each variant present in the server, so as to not flood the LSDBs with data not related to a phenotype. This information is particularly useful at the time of curation for assessing the variants' classification. In total, only five of the unique variants in our database (6%) have a frequency assigned in gnomAD: all four variants considered benign that are recorded in the database, with a low frequency ranging from one allele in 248,414 to less than 0.02%; as well as a single variant classified as likely pathogenic with a frequency of 0.04%. This additional information reinforces the importance of the LSDB approach and data sharing to improve the classification of variants for genetic diagnostics.

3.2 | Phenotypic data

To date, the database includes 84 patient records (44 females, 38 males, and two records of unspecified gender), 70 retrieved from publications, along with data from 11, to date, unpublished cases relating to our Molecular Genetics Laboratory and three from another external source. Of these patients' reports, 83 have an extended set of full clinical description; only one case lacked phenotypic description (Balciuniene et al., 2019). For the exhaustive description of phenotypes, use is made exclusively of a standard vocabulary for referencing phenotypic abnormalities, the so-called Human Phenotype Ontology (HPO) (Köhler et al., 2019). A total of 381 unique HPO terms annotate the patients, with each assigned from 1 patient to 67 patients for the most frequent term, optic atrophy (HP:0001138). This led to the analysis of a total of 5251 HPO terms (820 unique) by including the parent terms inferred by the ontological relationships (Figures S2 and S3). More than half of the unique terms (413) are found to be annotated in only one patient, which is considered insufficiently informative. Thus, we focused only on the HPO terms represented in at least 25% of patients, that is, terms annotated in 21 or more patients. Figure 3 provides an overview of this most significant HPO annotation as a grid, highlighting clusters that suggest that *NR2F1* patients share common phenotypic features.

We carried out a study of these same most-represented data as a diagram to integrate the ontological links together with the information on the frequency of the HPO terms (Figure 4). Patients with a pathogenic variant in *NR2F1* almost always harbor at least one abnormality of the eye (HP:0000478; 79 patients or 95%), most often an optic atrophy (HP:0000648; 67 patients or 81%), associated with at least one abnormality of the nervous system (HP:0000707; 76 patients or 92%), most often manifests as delayed speech and language development (HP:0000750; 48 patients or 58%). The vast majority of these phenotypes have been reported isolated (sporadic, HP:0003745; 64 patients or 77%), with this HPO term being used to specify the apparent absence of transmission of the patient's

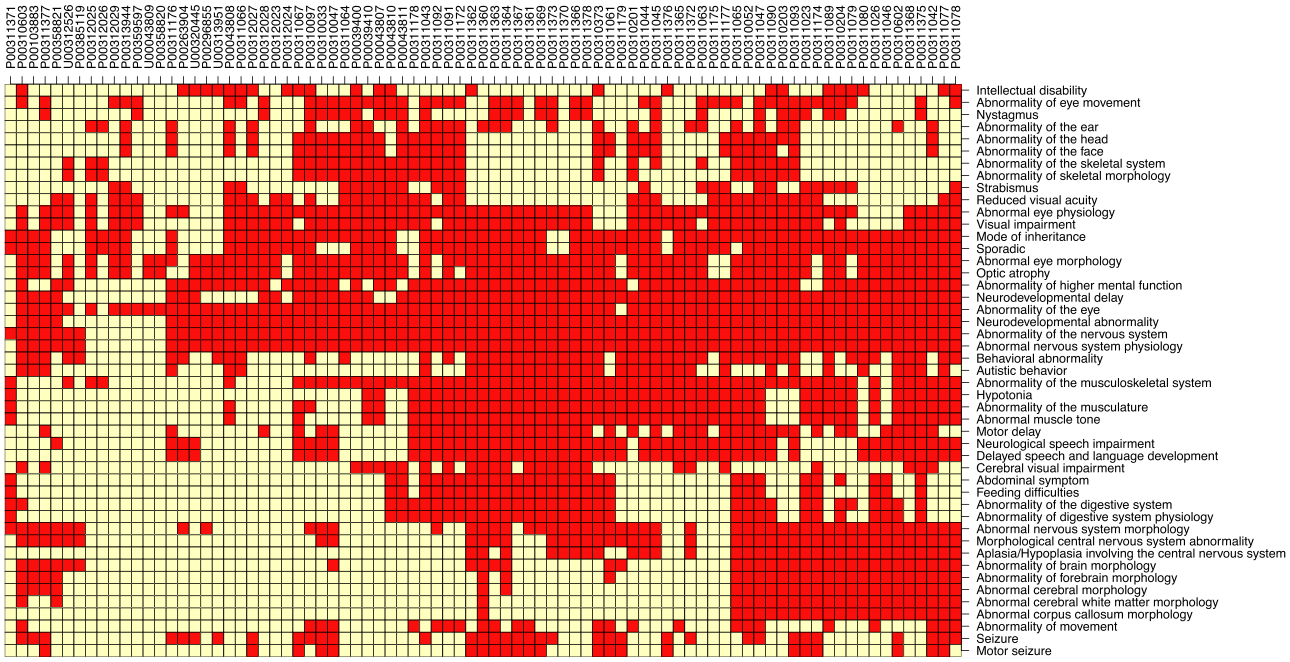

FIGURE 3 Visualization of the Human Phenotype Ontology (HPO) annotation describing the 82 symptomatic patients' reports with an extended set of full clinical description in the *NR2F1* data set. Only the HPO terms represented in more than 25% of the patients are displayed. A red box indicates the presence of the phenotype; columns and rows are clustered using *hclust*; human-readable shortened ontological term names were used (where possible); the term Mode of inheritance (HP:0000005) indicates that the mode of transmission of the patient's phenotypic profile to relatives is known—it appears in the data set as an ancestor of the HPO term Sporadic (HP:0003745; in 79% of patients) or Autosomal dominant inheritance (HP:0000006; in two patients). In columns, the identifiers of the patients (eight digits) are prefixed by an arbitrary letter. The visualization of the full HPO annotation in the *NR2F1* data set is available in Figure S2. Data as of October 7, 2021

phenotypic profile from one generation to the next. Only two independent familial cases in which the variant segregates with the phenotype have been recorded as such: patients #00311064 (father) and #00311065 (son) (Chen et al., 2016), as well as #00310033 and #00310047 (twins) (Mio et al., 2020). In contrast, two cases of inheritance of the variant without transmission of the phenotype are recorded in four patients: #00312024 (mother) and #00312023 (daughter), as well as #00312026 (mother) and #00312025 (daughter) from our center. This stresses the variable expressivity of BBSOAS, emphasizing that the mode of transmission is not necessarily obvious at the time of diagnosis.

3.3 | Clinical synopsis

The comparison of our data set, which reflects the state of the knowledge in the literature as well as recent unpublished patients, with the reference knowledge concerning BBSOAS in disease databases (MIM# 615722 and ORPHA:401777; Figure S4) confirms some clinical signs associated with *NR2F1* and suggests that some others should be added or removed (Table 2). Compared to the OMIM synopsis, we first corroborate that ocular involvement predominates (95% of patients), with mainly optic atrophy (HP:0000648) and visual impairment (HP:0000505) in 81% and 72% of patients respectively. The absence of optic disc pallor (HP:0000543) in both our data set and Orphanet is explained by the fact that this subjective clinical sign (discoloration of optic nerve head, as interpreted by the clinician) represents de facto a form of optic atrophy; some authors explicitly mention it (Bosch et al., 2014; Chen et al., 2016; Rech et al., 2020; Zou et al., 2020) while the majority do not. Ultimately, optic disc pallor is a milder form of optic atrophy.

We also confirm that the brain physiology (HP:0012638) is primarily affected (90% of patients), but with emphasis on the presence of delayed speech and language development (HP:0000750), motor delay (HP:0001270), and seizure (HP:0001250) in about half of the patients registered in the database. Interestingly, autistic behavior (HP:0000729), which is tagged "in some patients" in OMIM, appears as frequently as the previous phenotypes in the *NR2F1* database (42%). Furthermore, we confirm the variability and nonspecificity of the dysmorphic features.

Finally, we suggest considering, in addition to what is already well established, two independent phenotypes: hypotonia (HP:0001252), which is found in 55% of patients in the database, and feeding difficulties (HP:0011968), which is found in 34%. This impaired ability to eat, related to problems gathering food and preparing to suck, chew, and swallow it, may be secondary to hypotonia (Rech et al., 2020).

Overall, our meta-analysis of the state of molecular and clinical knowledge of BBSOAS, representing over 5000 ontological terms in 83 patients, is inspired by big data methodologies, by retaining only

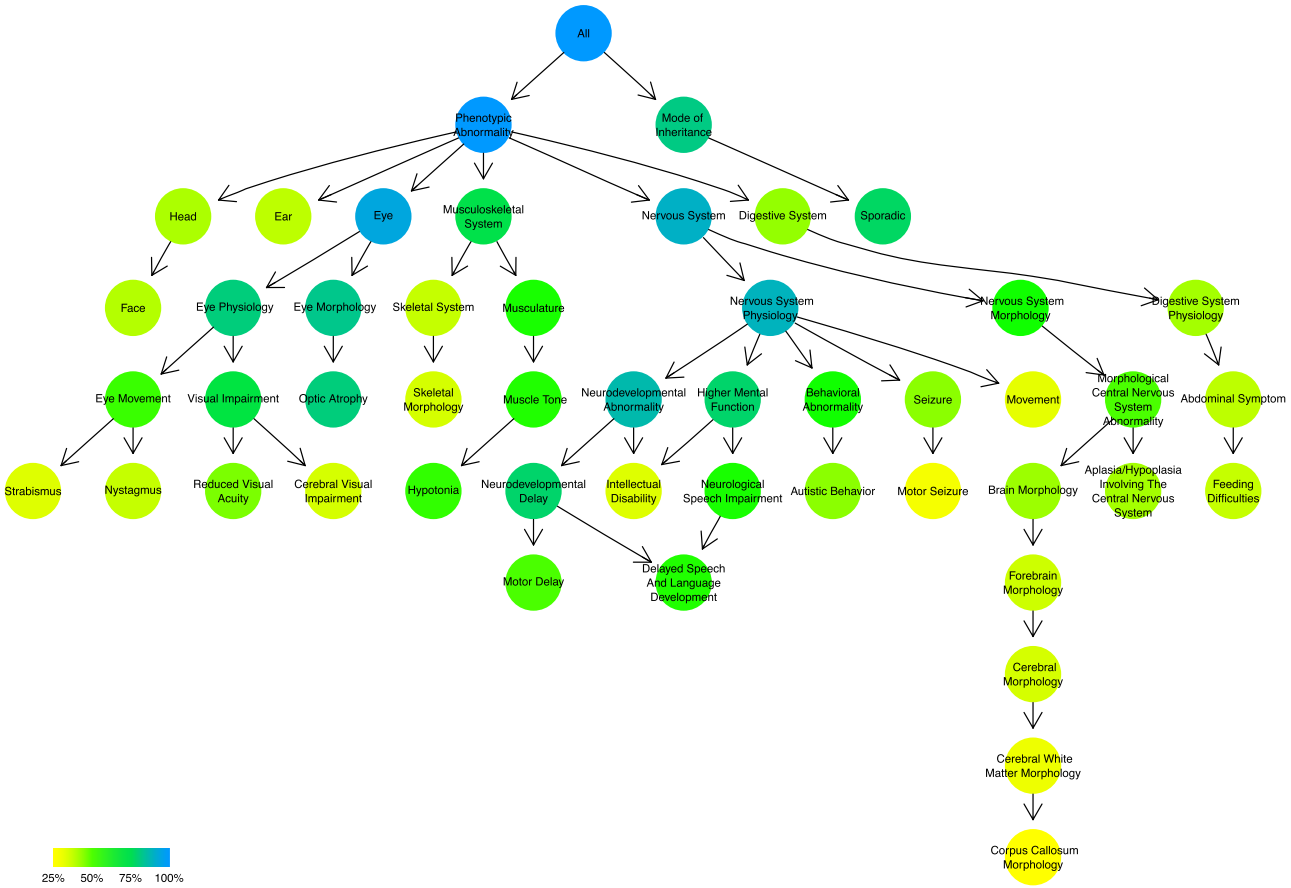

FIGURE 4 Visualization of frequency and relationships of the Human Phenotype Ontology (HPO) terms in the *NR2F1* data set. Only the HPO terms represented in more than 25% of the patients are displayed. The mode of inheritance (HP:0000005) and phenotypic abnormalities (HP:0000118) subgraphs are descending from the root of all terms (All; HP:0000001) in the HPO. Arrows indicate relations between terms in the ontology. Colors correspond to the frequency of the phenotypes, from 25% in yellow to 100% in blue, the light green color corresponding to a term present in half of the patients. Human-readable shortened ontological term names were used (where possible). The visualization of the full frequency and relationships of the HPO terms in the *NR2F1* data set is available in Figure S3. Data as of October 7, 2021

the most frequent phenotypes annotated in at least 25% of patients. Only strong discordance with the OMIM reference guided the suggestion of signs to be added or removed from the clinical synopsis (Table 2). The benefit is to eliminate punctual biases or diagnostic errors and provide a global profile. In contrast, it does not finely reflect the specificities of each individual. The database contains direct submissions from unpublished patients (17%), a practice that is becoming more frequent in our experience as curators. We will replicate our same analysis when new data are recorded in the database with the aim of showing further refinement clinical synopsis of the clinical spectrum BBSOAS.

4 | CONCLUSION

Genomic medicine calls for the precise definition of phenotypic variations (Biesecker, 2004; Deans et al., 2015; Robinson, 2012). Thus, descriptions of human diseases using HPO annotations are key elements of our novel *NR2F1* clinicobiological database. Meanwhile, specialized databases reporting pathogenic variations, the so-called LSDB, have proven to be the most complete (Brookes & Robinson, 2015), as they benefit from the participation of a curator who is a referent specialist for the gene or disease in question. Our systematic and rational approach, based on the computer representation and analysis of data, has led us to propose a refinement of the description of the clinical signs of BBSOAS, particularly with respect to neurological features and the suggestion of additional hypotonia that may result in impaired feeding ability. Interestingly, this analysis could be reproduced in the future to enable further refinement as new data is recorded in the database.

WEB RESSOURCES

The following web resources were used: gnomAD (https://gnomad.broadinstitute.org), HGVS-nomenclature (http://varnomen.hgvs.org), HPO (https://hpo.jax.org), LOVD (https://www.lovd.nl), Mutalyzer (https://mutalyzer.nl), OMIM (https://www.omim.org), Orphanet

TABLE 2 Comparison of phenotype frequency in the *NR2F1* database with clinical synopses referenced in the gold standard disease databases

	HPO id[a]	OMIM:615722[b]	ORPHA:401777[c]	LSDB[d]	Refine[e]
INHERITANCE					
- Autosomal dominant inheritance	HP:0000006	✓		5%	N/A
- Sporadic	HP:0003745			77%	N/A
NEUROLOGIC					
Central Nervous System					
- Abnormal nervous system physiology	HP:0012638			90%	
- Intellectual disability	HP:0001249	✓	Frequent	30%	=
- Global developmental delay	HP:0001263	✓	Frequent	8%	-
- Seizure	HP:0001250		Frequent	42%	++
- Delayed speech and language development	HP:0000750			58%	+++
- Absent speech	HP:0001344		Occasional	19%	
- Motor delay	HP:0001270			52%	+++
- Morphological central nervous system abnormality	HP:0002011			52%	
- Abnormality of brain morphology	HP:0012443			40%	
- Abnormal cerebral morphology	HP:0002060			31%	+
- Hypoplasia of the corpus callosum	HP:0002079		Frequent	13%	
- Abnormal hippocampus morphology	HP:0025100		Very rare	1%	
- Delayed myelination	HP:0012448		Very rare	5%	
Behavioral Psychiatric Manifestations					
- Behavioral abnormality	HP:0000708			60%	
- Autistic behavior	HP:0000729	In some patients		42%	++
- Attention deficit hyperactivity disorder	HP:0007018		Occasional	14%	
- Obsessive-compulsive behavior	HP:0000722		Occasional	4%	
- Repetitive compulsive behavior	HP:0008762		Occasional	4%	
HEAD & NECK					
Eyes					
- Optic atrophy	HP:0000648	✓	Frequent	81%	=
- Optic disc pallor	HP:0000543	✓		Implied	=
- Optic disc hypoplasia	HP:0007766		Occasional		
- Optic nerve hypoplasia	HP:0000609		Occasional	23%	
- Keratoconus	HP:0000563		Very rare	1%	
- Visual impairment	HP:0000505	✓	Frequent	72%	=
- Reduced visual acuity	HP:0007663	✓	Frequent	45%	=
- Amblyopia	HP:0000646		Occasional	11%	
- Cerebral visual impairment	HP:0100704	✓	Occasional	31%	=
- Visual field defects	HP:0001123	✓	Occasional	12%	-
- Nystagmus	HP:0000639	✓	Very rare	34%	=
- Strabismus	HP:0000486	✓	Occasional	30%	=
- Esotropia	HP:0000565		Occasional	8%	
- Exotropia	HP:0000577		Occasional	1%	
- Hypermetropia	HP:0000540		Very rare	11%	
- Myopia	HP:0000545		Very rare	1%	
Ear					
- Hearing impairment	HP:0000365		Occasional	17%	
- Protruding ear	HP:0000411		Occasional	11%	
- Abnormality of the helix	HP:0011039		Occasional	4%	
Face					
- Abnormal facial shape	HP:0001999	Var./nonspecif.	Frequent	2%	
- Epicanthus	HP:0000286	Var./nonspecif.	Occasional	6%	
- Upslanted palpebral fissure	HP:0000582	Var./nonspecif.	Occasional	4%	
- Anteverted nares	HP:0000463	Var./nonspecif.	Occasional	4%	
- Prominent nasal bridge	HP:0000426	Var./nonspecif.	Occasional	1%	
- Short nasal bridge	HP:0003194	Var./nonspecif.	Occasional		
ABDOMEN					
Gastrointestinal					
- Feeding difficulties	HP:0011968			34%	+
MUSCULATURE					
- Hypotonia	HP:0001252		Frequent	55%	++
- Spasticity	HP:0001257		Very rare	4%	
SKELETAL					
- Delayed skeletal maturation	HP:0002750		Very rare	1%	
Hands					
- Tapered finger	HP:0001182	✓	Occasional	2%	--
GROWTH					
Height					
- Short stature	HP:0004322		Very rare	4%	

Abbreviations: HPO, Human Phenotype Ontology; LSDB, locus-specific database; OMIM, Online Mendelian Inheritance in Man database; N/A, not applicable; =, confirmed; +, suggested addition; −, suggested deletion.

[a]HPO term identifier.

[b]Section clinical synopsis of OMIM entry #615722. Var./nonspecif.: dysmorphic features, variable, nonspecific; "✓": term listed in the synopsis.

[c]Section clinical signs and symptoms of orphanet entry #401777.

[d]Frequency in the *NR2F1* LSDB (data as of May 7, 2021).

[e]Proposal to refine the clinical synopsis in relation to the OMIM reference.

(https://www.orpha.net), PROSITE (https://prosite.expasy.org), PubMed (https://pubmed.ncbi.nlm.nih.gov), RefSeq (https://www.ncbi.nlm.nih.gov/refseq/), UniProt/Swiss-Prot (https://www.uniprot.org).

ACKNOWLEDGMENTS

We thankfully acknowledge grants from the following foundations and patients' associations: Association contre les Maladies mitochondriales, Fondation Visio, Kjer France Ouvrir les Yeux.

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

AUTHOR CONTRIBUTIONS

Marc Ferré designed and supervised the project. Johan T. den Dunnen and Pascal Reynier participated in the design and supervision. Benjamin Billiet collected the data with the help of Patrizia Amati-Bonneau, Valérie Desquiret-Dumas, Khadidja Guehlouz, Dan Milea, Philippe Gohier, Delphine Mirebeau-Prunier, and Guy Lenaers. Marc Ferré performed the statistical analysis. Benjamin Billiet, Pascal Reynier, and Marc Ferré wrote the manuscript with inputs from all authors. All authors have read and approved the final manuscript.

DATA AVAILABILITY STATEMENT

Data will be added and updated in the *NR2F1* database, which is freely available in three different ways: (1) on the World Wide Web at: https://www.lovd.nl/NR2F1; (2) via an application programming interface (API); and (3) as a public beacon in The Global Alliance for Genomics and Health Beacon Project (Global Alliance for Genomics and Health, 2016). The data described in this study are also available on request from the corresponding author.

ORCID

Guy Lenaers http://orcid.org/0000-0003-2736-3349
Johan T. den Dunnen http://orcid.org/0000-0002-6304-1710
Pascal Reynier http://orcid.org/0000-0003-0802-4608
Marc Ferré http://orcid.org/0000-0001-8265-7249

REFERENCES

Alfano, C., Viola, L., Heng, J. I., Pirozzi, M., Clarkson, M., Flore, G., de Maio, A., Schedl, A., Guillemot, F., & Studer, M. (2011). COUP-TFI promotes radial migration and proper morphology of callosal projection neurons by repressing Rnd2 expression. *Development*, *138*(21), 4685–4697. https://doi.org/10.1242/dev.068031

Al-Kateb, H., Shimony, J. S., Vineyard, M., Manwaring, L., Kulkarni, S., & Shinawi, M. (2013). NR2F1 haploinsufficiency is associated with optic atrophy, dysmorphism and global developmental delay. *American Journal of Medical Genetics. Part A*, *161A*(2), 377–381. https://doi.org/10.1002/ajmg.a.35650

Armentano, M., Filosa, A., Andolfi, G., & Studer, M. (2006). COUP-TFI is required for the formation of commissural projections in the forebrain by regulating axonal growth. *Development*, *133*(21), 4151–4162. https://doi.org/10.1242/dev.02600

Balciuniene, J., DeChene, E. T., Akgumus, G., Romasko, E. J., Cao, K., Dubbs, H. A., Mulchandani, S., Spinner, N. B., Conlin, L. K., Marsh, E. D., Goldberg, E., Helbig, I., Sarmady, M., & Abou Tayoun, A. (2019). Use of a dynamic genetic testing approach for childhood-onset epilepsy. *JAMA Network Open*, *2*(4), e192129. https://doi.org/10.1001/jamanetworkopen.2019.2129

Bertacchi, M., Parisot, J., & Studer, M. (2019). The pleiotropic transcriptional regulator COUP-TFI plays multiple roles in neural development and disease. *Brain Research*, *1705*, 75–94. https://doi.org/10.1016/j.brainres.2018.04.024

Bertacchi, M., Romano, A. L., Loubat, A., Tran Mau-Them, F., Willems, M., Faivre, L., Khau van Kien, P., Perrin, L., Devillard, F., Sorlin, A., Kuentz, P., Philippe, C., Garde, A., Neri, F., Di Giaimo, R., Oliviero, S., Cappello, S., D'Incerti, L., Frassoni, C., & Studer, M. (2020). NR2F1 regulates regional progenitor dynamics in the mouse neocortex and cortical gyrification in BBSOAS patients. *EMBO Journal*, *39*(13), e104163. https://doi.org/10.15252/embj.2019104163

Biesecker, L. G. (2004). Phenotype matters. *Nature Genetics*, *36*(4), 323–324. https://doi.org/10.1038/ng0404-323

Blum, M., Chang, H. Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A., Nuka, G., Paysan-Lafosse, T., Qureshi, M., Raj, S., Richardson, L., Salazar, G. A., Williams, L., Bork, P., Bridge, A., Gough, J., Haft, D. H., Letunic, I., Marchler-Bauer, A., … Finn, R. D. (2021). The InterPro protein families and domains database: 20 years on. *Nucleic Acids Research*, *49*(D1), D344–D354. https://doi.org/10.1093/nar/gkaa977

Bojanek, E. K., Mosconi, M. W., Guter, S., Betancur, C., Macmillan, C., & Cook, E. H. (2020). Clinical and neurocognitive issues associated with Bosch–Boonstra–Schaaf optic atrophy syndrome: A case study. *American Journal of Medical Genetics. Part A*, *182*(1), 213–218. https://doi.org/10.1002/ajmg.a.61409

Bosch, D. G., Boonstra, F. N., Gonzaga-Jauregui, C., Xu, M., de Ligt, J., Jhangiani, S., Wiszniewski, W., Muzny, D. M., Yntema, H. G., Pfundt, R., Vissers, L. E., Spruijt, L., Blokland, E. A., Chen, C. A., Baylor-Hopkins Center for Mendelian, G., Lewis, R. A., Tsai, S. Y., Gibbs, R. A., Tsai, M. J., … Schaaf, C. P. (2014). NR2F1 mutations cause optic atrophy with intellectual disability. *American Journal of Human Genetics*, *94*(2), 303–309. https://doi.org/10.1016/j.ajhg.2014.01.002

Bosch, D. G., Boonstra, F. N., de Leeuw, N., Pfundt, R., Nillesen, W. M., de Ligt, J., Gilissen, C., Jhangiani, S., Lupski, J. R., Cremers, F. P., & de Vries, B. B. (2016). Novel genetic causes for cerebral visual impairment. *European Journal of Human Genetics*, *24*(5), 660–665. https://doi.org/10.1038/ejhg.2015.186

Brookes, A. J., & Robinson, P. N. (2015). Human genotype-phenotype databases: aims, challenges and opportunities. *Nature Reviews Genetics*, *16*(12), 702–715. https://doi.org/10.1038/nrg3932

Brown, K. K., Alkuraya, F. S., Matos, M., Robertson, R. L., Kimonis, V. E., & Morton, C. C. (2009). NR2F1 deletion in a patient with a de novo paracentric inversion, inv(5)(q15q33.2), and syndromic deafness. *American Journal of Medical Genetics. Part A*, *149A*(5), 931–938. https://doi.org/10.1002/ajmg.a.32764

Chen, C. A., Bosch, D. G., Cho, M. T., Rosenfeld, J. A., Shinawi, M., Lewis, R. A., Mann, J., Jayakar, P., Payne, K., Walsh, L., Moss, T., Schreiber, A., Schoonveld, C., Monaghan, K. G., Elmslie, F., Douglas, G., Boonstra, F. N., Millan, F., Cremers, F. P., … Schaaf, C. (2016). The expanding clinical phenotype of Bosch–Boonstra–Schaaf optic atrophy syndrome: 20 new cases and possible genotype-phenotype correlations. *Genetics in Medicine*, *18*(11), 1143–1150. https://doi.org/10.1038/gim.2016.18

Cotton, R. G., Auerbach, A. D., Axton, M., Barash, C. I., Berkovic, S. F., Brookes, A. J., Burn, J., Cutting, G., den Dunnen, J. T., Flicek, P., Freimer, N., Greenblatt, M. S., Howard, H. J., Katz, M., Macrae, F. A., Maglott, D., Möslein, G., Povey, S., Ramesar, R. S., … Watson, M. (2008). Genetics. The Human Variome Project. *Science*, *322*(5903), 861–862. https://doi.org/10.1126/science.1167363

Deans, A. R., Lewis, S. E., Huala, E., Anzaldo, S. S., Ashburner, M., Balhoff, J. P., Blackburn, D. C., Blake, J. A., Burleigh, J. G., Chanet, B.,

Cooper, L. D., Courtot, M., Csösz, S., Cui, H., Dahdul, W., Das, S., Dececchi, T. A., Dettai, A., Diogo, R., … Mabee, P. (2015). Finding our way through phenotypes. *PLOS Biology*, *13*(1), e1002033. https://doi.org/10.1371/journal.pbio.1002033

Dimassi, S., Labalme, A., Ville, D., Calender, A., Mignot, C., Boutry-Kryza, N., de Bellescize, J., Rivier-Ringenbach, C., Bourel-Ponchel, E., Cheillan, D., Simonet, T., Maincent, K., Rossi, M., Till, M., Mougou-Zerelli, S., Edery, P., Saad, A., Heron, D., des Portes, V., … Lesca, G. (2016). Whole-exome sequencing improves the diagnosis yield in sporadic infantile spasm syndrome. *Clinical Genetics*, *89*(2), 198–204. https://doi.org/10.1111/cge.12636

den Dunnen, J. T., Dalgleish, R., Maglott, D. R., Hart, R. K., Greenblatt, M. S., McGowan-Jordan, J., Roux, A. F., Smith, T., Antonarakis, S. E., & Taschner, P. E. (2016). HGVS recommendations for the description of sequence variants: 2016 Update. *Human Mutation*, *37*(6), 564–569. https://doi.org/10.1002/humu.22981

Eldomery, M. K., Coban-Akdemir, Z., Harel, T., Rosenfeld, J. A., Gambin, T., Stray-Pedersen, A., Küry, S., Mercier, S., Lessel, D., Denecke, J., Wiszniewski, W., Penney, S., Liu, P., Bi, W., Lalani, S. R., Schaaf, C. P., Wangler, M. F., Bacino, C. A., Lewis, R. A., … Lupski, J. R. (2017). Lessons learned from additional research analyses of unsolved clinical exome cases. *Genome Medicine*, *9*(1), 26. https://doi.org/10.1186/s13073-017-0412-6

Fokkema, I., van der Velde, K. J., Slofstra, M. K., Ruivenkamp, C., Vogel, M. J., Pfundt, R., Blok, M. J., Lekanne Deprez, R. H., Waisfisz, Q., Abbott, K. M., Sinke, R. J., Rahman, R., Nijman, I. J., de Koning, B., Thijs, G., Wieskamp, N., Moritz, R., Charbon, B., Saris, J. J., … van Gijn, M. E. (2019). Dutch genome diagnostic laboratories accelerated and improved variant interpretation and increased accuracy by sharing data. *Human Mutation*, *40*, 2230–2238. https://doi.org/10.1002/humu.23896

Fokkema, I. F., Taschner, P. E., Schaafsma, G. C., Celli, J., Laros, J. F., & den Dunnen, J. T. (2011). LOVD v.2.0: the next generation in gene variant databases. *Human Mutation*, *32*(5), 557–563. https://doi.org/10.1002/humu.21438

Global Alliance for Genomics and Health. (2016). Genomics. A federated ecosystem for sharing genomic, clinical data. *Science*, *352*(6291), 1278–1280. https://doi.org/10.1126/science.aaf6162

Gray, K. A., Daugherty, L. C., Gordon, S. M., Seal, R. L., Wright, M. W., & Bruford, E. A. (2013). Genenames.org: The HGNC resources in 2013. *Nucleic Acids Research*, *41*(Database issue), D545–D552. https://doi.org/10.1093/nar/gks1066

Greene, D., Richardson, S., & Turro, E. (2017). ontologyX: A suite of R packages for working with ontological data. *Bioinformatics*, *33*(7), 1104–1106. https://doi.org/10.1093/bioinformatics/btw763

Hamosh, A., Scott, A. F., Amberger, J., Valle, D., & McKusick, V. A. (2000). Online Mendelian Inheritance in Man (OMIM). *Human Mutation*, *15*(1), 57–61. https://doi.org/10.1002/(SICI)1098-1004(200001)15:1%3C57::AID-HUMU12%3E3.0.CO;2-G

Hino-Fukuyo, N., Kikuchi, A., Yokoyama, H., Iinuma, K., Hirose, M., Haginoya, K., Niihori, T., Nakayama, K., Aoki, Y., & Kure, S. (2017). Long-term outcome of a 26-year-old woman with West syndrome and an nuclear receptor subfamily 2 group F member 1 gene (NR2F1) mutation. *Seizure*, *50*, 144–146. https://doi.org/10.1016/j.seizure.2017.06.018

Hobbs, M. M., Wolters, W. C., & Rayapati, A. O. (2020). Bosch–Boonstra–Schaaf optic atrophy syndrome presenting as new-onset psychosis in a 32-year-old man: A case report and literature review. *Journal of Psychiatric Practice*, *26*(1), 58–62. https://doi.org/10.1097/PRA.0000000000000440

INSERM. (1997). Orphanet: an online database of rare diseases and orphan drugs. Retrieved May 21, 2021, from http://www.orpha.net

Jezela-Stanek, A., Ciara, E., Jurkiewicz, D., Kucharczyk, M., Jędrzejowska, M., Chrzanowska, K. H., Krajewska-Walasek, M., & Żemojtel, T. (2020). The phenotype-driven computational analysis yields clinical diagnosis for patients with atypical manifestations of known intellectual disability syndromes. *Molecular Genetics & Genomic Medicine*, *8*(9), e1263. https://doi.org/10.1002/mgg3.1263

Kaiwar, C., Zimmermann, M. T., Ferber, M. J., Niu, Z., Urrutia, R. A., Klee, E. W., & Babovic-Vuksanovic, D. (2017). Novel NR2F1 variants likely disrupt DNA binding: molecular modeling in two cases, review of published cases, genotype-phenotype correlation, and phenotypic expansion of the Bosch–Boonstra–Schaaf optic atrophy syndrome. *Cold Spring Harbor Molecular Case Studies*, *3*(6), https://doi.org/10.1101/mcs.a002162

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., Gauthier, L. D., Brand, H., Solomonson, M., Watts, N. A., Rhodes, D., Singer-Berk, M., England, E. M., Seaby, E. G., Kosmicki, J. A., … MacArthur, D. G. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, *581*(7809), 434–443. https://doi.org/10.1038/s41586-020-2308-7

Köhler, S., Carmody, L., Vasilevsky, N., Jacobsen, J., Danis, D., Gourdine, J. P., Gargano, M., Harris, N. L., Matentzoglu, N., McMurry, J. A., Osumi-Sutherland, D., Cipriani, V., Balhoff, J. P., Conlin, T., Blau, H., Baynam, G., Palmer, R., Gratian, D., Dawkins, H., … Robinson, P. N. (2019). Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources. *Nucleic Acids Research*, *47*(D1), D1018–D1027. https://doi.org/10.1093/nar/gky1105

Martín-Hernández, E., Rodríguez-García, M. E., Chen, C. A., Cotrina-Vinagre, F. J., Carnicero-Rodríguez, P., Bellusci, M., Schaaf, C. P., & Martínez-Azorín, F. (2018). Mitochondrial involvement in a Bosch–Boonstra–Schaaf optic atrophy syndrome patient with a novel de novo NR2F1 gene mutation. *Journal of Human Genetics*, *63*(4), 525–528. https://doi.org/10.1038/s10038-017-0398-3

Mio, C., Fogolari, F., Pezzoli, L., D'Elia, A. V., Iascone, M., & Damante, G. (2020). Missense NR2F1 variant in monozygotic twins affected with the Bosch–Boonstra–Schaaf optic atrophy syndrome. *Molecular Genetics & Genomic Medicine*, *8*(7), e1278. https://doi.org/10.1002/mgg3.1278

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., … Pruitt, K. D. (2016). Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research*, *44*(D1), D733–D745. https://doi.org/10.1093/nar/gkv1189

Park, S. E., Lee, J. S., Lee, S. T., Kim, H. Y., Han, S. H., & Han, J. (2019). Targeted panel sequencing identifies a novel NR2F1 mutations in a patient with Bosch–Boonstra–Schaaf optic atrophy syndrome. *Ophthalmic Genetics*, *40*(4), 359–361. https://doi.org/10.1080/13816810.2019.1650074

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Rech, M. E., McCarthy, J. M., Chen, C. A., Edmond, J. C., Shah, V. S., Bosch, D., Berry, G. T., Williams, L., Madan-Khetarpal, S., Niyazov, D., Shaw-Smith, C., Kovar, E. M., Lupo, P. J., & Schaaf, C. P. (2020). Phenotypic expansion of Bosch–Boonstra–Schaaf optic atrophy syndrome and further evidence for genotype-phenotype correlations. *American Journal of Medical Genetics. Part A*, *182*(6), 1426–1437. https://doi.org/10.1002/ajmg.a.61580

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., Rehm, H. L., & ACMG Laboratory Quality Assurance, C. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular

Pathology. *Genetics in Medicine*, *17*(5), 405–424. https://doi.org/10.1038/gim.2015.30

Robinson, P. N. (2012). Deep phenotyping for precision medicine. *Human Mutation*, *33*(5), 777–780. https://doi.org/10.1002/humu.22080

Sayers, E. W., Barrett, T., Benson, D. A., Bolton, E., Bryant, S. H., Canese, K., Chetvernin, V., Church, D. M., Dicuccio, M., Federhen, S., Feolo, M., Geer, L. Y., Helmberg, W., Kapustin, Y., Landsman, D., Lipman, D. J., Lu, Z., Madden, T. L., Madej, T., … Ye, J. (2010). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, *38*(Database issue), D5–D16. https://doi.org/10.1093/nar/gkp967

Starosta, R. T., Tarnowski, J., Vairo, F. P. E., Raymond, K., Preston, G., & Morava, E. (2020). Bosch–Boonstra–Schaaf optic atrophy syndrome (BBSOAS) initially diagnosed as ALG6-CDG: Functional evidence for benignity of the ALG6 c.391T>C (p.Tyr131His) variant and further expanding the BBSOAS phenotype. *European Journal of Medical Genetics*, *63*(7), 103941. https://doi.org/10.1016/j.ejmg.2020.103941

Vihinen, M., den Dunnen, J. T., Dalgleish, R., & Cotton, R. G. (2012). Guidelines for establishing locus-specific databases. *Human Mutation*, *33*(2), 298–305. https://doi.org/10.1002/humu.21646

Vissers, L., van Nimwegen, K., Schieving, J. H., Kamsteeg, E. J., Kleefstra, T., Yntema, H. G., Pfundt, R., van der Wilt, G. J., Krabbenborg, L., Brunner, H. G., van der Burg, S., Grutters, J., Veltman, J. A., & Willemsen, M. (2017). A clinical utility study of exome sequencing versus conventional genetic testing in pediatric neurology. *Genetics in Medicine*, *19*(9), 1055–1063. https://doi.org/10.1038/gim.2017.1

Walsh, S., Gösswein, S. S., Rump, A., von der Hagen, M., Hackmann, K., Schröck, E., Di Donato, N., & Kahlert, A. K. (2020). Novel dominant-negative NR2F1 frameshift mutation and a phenotypic expansion of the Bosch–Boonstra–Schaaf optic atrophy syndrome. *European Journal of Medical Genetics*, *63*(10), 104019. https://doi.org/10.1016/j.ejmg.2020.104019

Wildeman, M., van Ophuizen, E., den Dunnen, J. T., & Taschner, P. E. (2008). Improving sequence variant descriptions in mutation databases and literature using the Mutalyzer sequence variation nomenclature checker. *Human Mutation*, *29*(1), 6–13. https://doi.org/10.1002/humu.20654

Zou, W., Cheng, L., Lu, S., & Wu, Z. (2020). A de novo nonsense mutation in the N-terminal of ligand-binding domain of NR2F1 gene provoked a milder phenotype of BBSOAS. *Ophthalmic Genetics*, *41*(1), 88–89. https://doi.org/10.1080/13816810.2020.1719520

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Billiet, B., Amati-Bonneau, P., Desquiret-Dumas, V., Guehlouz, K., Milea, D., Gohier, P., Lenaers, G., Mirebeau-Prunier, D., den Dunnen, J. T., Reynier, P., & Ferré, M. (2022). *NR2F1* database: 112 variants and 84 patients support refining the clinical synopsis of Bosch–Boonstra–Schaaf optic atrophy syndrome. *Human Mutation*, 43, 128–142. https://doi.org/10.1002/humu.24305