Genotype-phenotype correlation in NF1 individuals: evidence for a more severe phenotype associated with missense mutations affecting *NF1* codons 844-848.

Magdalena Koczkowska,[1] Yunjia Chen,[1] Tom Callens,[1] Alicia Gomes,[1] Angela Sharp,[1] Sherrell Johnson,[1] Meng-Chang Hsiao,[1] Zhenbin Chen,[1] Meena Balasubramanian,[2] Christopher P. Barnett,[3] Troy A. Becker,[4] Shay Ben-Shachar,[5] Debora R. Bertola,[6] Jaishri O. Blakeley,[7] Emma M.M. Burkitt-Wright,[8] Alison Callaway,[9] Melissa Crenshaw,[4] Karin S. Cunha,[10] Mitch Cunningham,[11] Maria D. D'Agostino,[12] Karin Dahan,[13] Alessandro De Luca,[14] Anne Destrée,[13] Radhika Dhamija,[15] Marica Eoli,[16] D. Gareth R. Evans,[8] Patricia Galvin-Parton,[17] Jaya K. George-Abraham,[18] Karen W. Gripp,[19] Jose Guevara-Campos,[20] Neil A. Hanchard,[21] Concepcion Hernández-Chico,[22] LaDonna Immken,[18] Sandra Janssens,[23] Kristi J. Jones,[24] Beth A. Keena,[25] Aaina Kochhar,[26] Jan Liebelt,[3] Arelis Martir-Negron,[27] Maurice J. Mahoney,[28] Isabelle Maystadt,[13] Carey McDougall,[25] Meriel McEntagart,[29] Nancy Mendelsohn,[30] David T. Miller,[31] Geert Mortier,[32] Jenny Morton,[33] John Pappas,[34] Scott R. Plotkin,[35] Dinel Pond,[30] Kenneth Rosenbaum,[36] Karol Rubin,[37] Laura Russell,[12] Lane S. Rutledge,[1] Veronica Saletti,[38] Rhonda Schonberg,[36] Allison Schreiber,[39] Meredith Seidel,[35] Elizabeth Siqveland,[30] David W. Stockton,[11] Eva Trevisson,[40] Nicole J. Ullrich,[41] Meena Upadhyaya,[42] Rick van Minkelen,[43] Helene Verhelst,[44] Margaret R. Wallace,[45] Yoon-Sim Yap,[46] Elaine Zackai,[25] Jonathan Zonana,[47] Vickie Zurcher,[39] Kathleen Claes,[23] Yolanda Martin,[22] Bruce R. Korf,[1] Eric Legius,[48] and Ludwine M. Messiaen,[1*]

[1]*Department of Genetics, University of Alabama at Birmingham, Birmingham, Alabama, 35294, USA;* [2]*Sheffield Clinical Genetics Service, Sheffield Children's NHS Foundation Trust, Sheffield, S102TH, United Kingdom;* [3]*Women's and Children's Hospital/SA Pathology, North Adelaide, South Australia, SA 5006, Australia;* [4]*Medical Genetics, John Hopkins All Children's Hospital, St. Petersburg, Florida, 33701, USA;* [5]*The Genetic Institute, Tel-Aviv Sourasky Medical Center and Sackler Faculty of Medicine, Tel-Aviv, 6997801, Israel;* [6]*Department of Pediatrics, University of São Paulo, São Paulo, 05403-000, Brazil;* [7]*Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, Maryland, 21287, USA;* [8]*Genomic Medicine, Division of Evolution and Genomic Sciences, Manchester Academic Health Sciences Centre, University of Manchester, Central Manchester University Hospitals NHS Foundation Trust, Manchester, M13 9WL, United Kingdom;* [9]*Wessex Regional Genetics Laboratory, Salisbury NHS Foundation Trust, Salisbury, SP2 8BJ, United Kingdom;* [10]*Department of Pathology, School of Medicine, Universidade Federal Fluminense, Niterói, 24220-900, Brazil;* [11]*Division of Genetic, Genomic and Metabolic Disorders, Children's Hospital of Michigan, Detroit Medical Center, Detroit, Michigan, 48201, USA;* [12]*Department of Medical Genetics, McGill University Health Centre, Montréal, Québec, H4A 3J1, Canada;* [13]*Center for Human Genetics, Institute of Pathology and Genetics (IPG), Gosselies, B-6041, Belgium;* [14]*Molecular Genetics Unit, Casa Sollievo della Sofferenza Hospital, IRCCS, San Giovanni Rotondo, 71013, Italy*; [15]*Department of Clinical Genomics and Neurology, Mayo Clinic, Phoenix, Arizona, 85259, USA;* [16]*Unit of Molecular Neuro-Oncology, IRCCS Foundation, Carlo Besta Neurological Institute, Milan, 20133, Italy;* [17]*Department of Genetics, Stony Brook Children's, Stony Brook, New York, 11794, USA;* [18]*Dell Children's Medical Center of Central Texas, Austin, Texas, 78723, USA;* [19]*Division of Medical Genetics, Al DuPont Hospital for Children, Wilmington, Delaware, 19803, USA;*

[20]Pediatrics Service, Felipe Guevara Rojas Hospital, University of Oriente, El Tigre-Anzoátegui, Venezuela, 6034, Spain; [21]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas, 77030, USA; [22]Department of Genetics, Hospital Universitario Ramón y Cayal, Institute of Health Research (IRYCIS), Madrid, 28034, Spain and Center for Biomedical Research-Network of Rare Diseases (CIBERER); [23]Center for Medical Genetics, Ghent University Hospital, Ghent, 9000, Belgium; *[24]Department of Clinical Genetics, the Children's Hospital at Westmead, Westmead, NSW 2145, Australia; [25]Division of Human Genetics, Children's Hospital of Philadelphia, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, 19104, USA; [26]Department of Genetics, Valley Children's Healthcare, Madera, California, 93636, USA; [27]Department of Pediatrics, Advocate Children's Hospital, Park Ridge, Illinois, 60068, USA and Division of Clinical Genetics & Metabolic Disorders, Palm Beach Gardens Outpatient Center, Nicklaus Children's Hospital, Miami, Florida, 33155, USA; [28]Department of Genetics, Yale University, New Haven, Connecticut, 06520, USA; [29]St George's University Hospitals NHS Foundation Trust, London, SW17 0QT, United Kingdom; [30]Genomics Medicine Program, Children's Hospital Minnesota, Minneapolis, Minnesota, 55404, USA; [31]Multidisciplinary Neurofibromatosis Program, Boston Children's Hospital, Boston, Massachusetts, 02115, USA; [32]Department of Medical Genetics, University of Antwerp and Antwerp University Hospital, Antwerp, 2650, Belgium; [33]Birmingham Women's and Children's NHS Foundation Trust, Birmingham, B15 2TG, United Kingdom; [34]Department of Pediatrics, Clinical Genetic Services, NYU School of Medicine, New York, New York, 10016, USA; [35]Department of Neurology and Cancer Center, Massachusetts General Hospital, Boston, Massachusetts, 02114, USA; [36]Division of Genetics and Metabolism, Children's National Health System, Washington, DC, 20010, USA; [37]University of Minnesota, Health, Minneapolis,*

Minnesota, 55404, USA; [38]*Developmental Neurology Unit, IRCCS Foundation, Carlo Besta Neurological Institute, Milan, 20133, Italy;* [39]*Genomic Medicine Institute, Cleveland Clinic, Cleveland, Ohio, 44195, USA;* [40]*Clinical Genetics Unit, Department of Women's and Children's Health, University of Padova, Padova, Italy and Italy Istituto di Ricerca Pediatria, IRP, Città della Speranza, Padova, 35128, Italy;* [41]*Department of Neurology, Boston Children's Hospital, Boston, Massachusetts, 02115, USA;* [42]*Division of Cancer and Genetics, Cardiff University, Cardiff, CF14 4XN, United Kingdom;* [43]*Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, 3015 GE, the Netherlands;* [44]*Department of Paediatrics, Division of Paediatric Neurology, Ghent University Hospital, Ghent, 9000, Belgium;* [45]*Department of Molecular Genetics & Microbiology, University of Florida College of Medicine, Gainesville, Florida, 32610, USA;* [46]*Division of Medical Oncology, National Cancer Centre, Singapore, 169610, Singapore;* [47]*Department of Molecular and Medical Genetics, Oregon Health and Science University, Portland, Oregon, 97239, USA;* [48]*Department of Human Genetics, KU Leuven - University of Leuven, Leuven, B-3000, Belgium.*

***Corresponding author:**

Ludwine M. Messiaen, PhD; Medical Genomics Laboratory, Department of Genetics, University of Alabama at Birmingham, 720 20th St. S., Birmingham, AL 35294, USA; Tel: (1) 205 934 5562, Fax: (1) 205 996 2929; lmessiaen@uabmc.edu

Abstract

Neurofibromatosis type 1 (NF1), a common genetic disorder with a birth incidence of 1:2000-3000, is characterized by a highly variable clinical presentation. To date, only two clinically relevant intragenic genotype-phenotype correlations have been reported for *NF1* missense mutations affecting p.Arg1809 and a single amino acid deletion p.Met922del. Both variants predispose to a distinct mild NF1 phenotype with neither externally visible cutaneous/plexiform neurofibromas nor other tumors. Here, we report 162 individuals (129 unrelated probands and 33 affected relatives) heterozygous for a constitutional missense mutation affecting one of five neighboring *NF1* codons Leu844, Cys845, Ala846, Leu847 and Gly848, located in the Cysteine-Serine-Rich Domain (CSRD). Collectively, these recurrent missense mutations affect ~0.8% of unrelated *NF1* mutation-positive probands in the University of Alabama at Birmingham (UAB) cohort. Major superficial plexiform neurofibromas and symptomatic spinal neurofibromas were more prevalent in these individuals compared with classic NF1 cohorts (both $p<0.0001$). Nearly half of the individuals had symptomatic or asymptomatic optic pathway gliomas and/or skeletal abnormalities. Additionally, variants in this region seem to confer a high predisposition to develop malignancies compared with the general NF1 population ($p=0.0061$). Our results demonstrate that these *NF1* missense mutations, although located outside the GAP-related domain, may be an important risk factor for a severe presentation. A genotype-phenotype correlation at the *NF1* region 844-848 exists and will be valuable in the management and genetic counseling of a significant number of individuals.

Introduction

Neurofibromatosis type 1 (NF1 [MIM: 162200]), one of the most common genetic disorders with a birth incidence of 1 in 2000-3000 [1-3], is characterized by a highly variable inter- and intrafamilial expressivity [4]. It is caused by loss-of-function genetic variants in *NF1* (MIM: 613113), located on chromosome 17q11.2. *NF1* encodes neurofibromin, a GTPase activating protein (GAP) that down-regulates the RAS signal transduction pathway through its GAP-related domain (GRD) [5, 6]. The most common first signs of NF1 are multiple café-au-lait macules (CALMs) in >95% of infants and skinfold freckling in >80% of children by the age of 7 years [7]. Other clinical features observed in >90% of adults with NF1 are iris Lisch nodules and cutaneous neurofibromas [8]. Individuals with a more severe phenotype present with plexiform and/or spinal neurofibromas, symptomatic optic pathway gliomas (OPGs) as well as specific osseous lesions, such as sphenoid wing or tibial dysplasia. Approximately 50% of NF1 cases have *de novo* mutations, while the remaining individuals inherit the disorder from an affected parent [4]. According to the National Institutes of Health (NIH) diagnostic criteria at least two of the aforementioned features are required to classify a person as having the clinical diagnosis of NF1 [9].

Due to the variability in clinical presentation, age-dependency of most manifestations, the timing and number of second hits in specific cells, and the wide *NF1* allelic heterogeneity, identification of specific genotype-phenotype correlations is extremely challenging. To date, over 2800 *different* germline *NF1* pathogenic variants have been identified in the University of Alabama at Birmingham (UAB) cohort with only 31 unique pathogenic variants present in ≥0.5% of all unrelated individuals (L.M.M, unpublished data). Moreover, a mild NF1 phenotype, including

only CALMs and skinfold freckles, overlaps with Legius syndrome (MIM: 611431), caused by mutations in *SPRED1* (MIM: 609291) [10, 11].

So far, only three clinically significant genotype-phenotype correlations have been reported. First, individuals with a constitutional *NF1* microdeletion usually show a more severe phenotype compared to the general NF1 population. The *NF1* microdeletion syndrome (MIM: 613675) is typically characterized by a large number of neurofibromas at a young age, dysmorphic facial features (hypertelorism, downslanted palpebral fissures, broad nasal bridge, low set ears, micrognathia, coarse face, facial asymmetry) and developmental delay and/or intellectual disability. Individuals may present with cardiac defects as well as growth and skeletal abnormalities. *NF1*-microdeletions have been associated with an increased lifetime risk for malignant peripheral nerve sheath tumors (MPNSTs). The constitutional co-deletion of *SUZ12* (MIM: 606245) within the common *NF1*-microdeletion region is thought to be a risk factor for the malignant neoplasms [12]. Second, individuals with a specific single amino acid *NF1* deletion (c.2970_2972del; p.Met992del) present with a milder phenotype. These individuals have multiple CALMs with/without freckles, but no externally visible cutaneous or plexiform neurofibromas [13]. A third genotype-phenotype correlation involving *NF1* missense mutations affecting arginine at position 1809 is also associated with a distinct presentation [14, 15], including developmental delay and/or learning disabilities, pulmonic stenosis and Noonan-like features, but no external plexiform neurofibromas or symptomatic OPGs. Both of these affected amino acids reside outside the GRD domain.

Another distinct form of NF1 is familial spinal neurofibromatosis (FSNF [MIM: 162210]) originally described by Pulst et al. (1991) [16] in six affected members from two unrelated families. It is characterized by bilateral and histologically proven neurofibromas of all spinal

dorsal roots with a paucity or absolute lack of cutaneous manifestations [17, 18]. So far, only ~100 individuals (both familial and sporadic) have been reported with this form [18]. It has been suggested that individuals with the severe subtype of FSNF more frequently carry an *NF1* missense or splicing mutation [19-21]. Of particular interest are two families: a two-generation family with three first-degree relatives reported by Pascual-Castroviejo et al. (2007) [22] and a three-generation family with three first-degree relatives reported by Burkitt-Wright et al. (2013) [17]. Specific *NF1* missense mutations c.2542G>C (p.Gly848Arg) and c.2543G>A (p.Gly848Glu), located in the Cysteine-Serine-Rich Domain (CSRD), were present in all individuals affected by multiple spinal dorsal root neurofibromas. Despite the evidence that c.2542G>C (p.Gly848Arg) is a clearly pathogenic mutation, two recent studies using mouse models did not recapitulate the phenotype identified in humans [23, 24]. Genetically engineered mice with c.2542G>C (p.Gly848Arg) mutation developed neither OPGs [24] nor plexiform neurofibromas [23], demonstrating phenotypic divergence between NF1 individuals and mice.

In this study, we report a cohort of 129 unrelated probands and 33 affected relatives heterozygous for a constitutional missense mutation affecting one of five neighboring *NF1* codons Leu844, Cys845, Ala846, Leu847 and Gly848. These individuals have a high prevalence of a severe phenotype, including plexiform and symptomatic spinal neurofibromas, symptomatic optic pathway gliomas, other malignant neoplasms, as well as bone abnormalities. The current findings clearly demonstrate that missense mutations outside the GRD are not solely associated with a mild phenotype.

Materials and methods

Individuals and phenotypic data

A total of 162 individuals heterozygous for a missense mutation affecting one of five neighboring *NF1* codons Leu844, Cys845, Ala846, Leu847 and Gly848 were included in the study. Blood samples from seventy-eight individuals (67 probands and 11 relatives) were originally sent to the UAB Medical Genomics Laboratory for molecular *NF1* genetic testing to establish or confirm the diagnosis for NF1. This initial study was expanded to include an additional eighty-four individuals (62 probands and 22 relatives), molecularly diagnosed in collaborating institutions (as detailed in Table S1).

All individuals included in this study were clinically assessed using the standardized phenotypic checklist form as previously reported [15] (Figure S1). The clinical data were collected at the time of mutation analysis and re-verified for accuracy by referring physicians co-authoring this paper at the time of this study. Additionally, referring physicians updated the phenotypic data at the time of this genotype-phenotype study, when available, i.e. when the individual had been seen and followed at their institution after genetic testing results were reported. The phenotypic data and age provided correspond to the latest clinical evaluation. The phenotypic checklist form consists of two parts: i/ general information including gender, date of birth, ethnicity, height, head circumference (HC), weight, fulfillment of the NIH diagnostic criteria and mode of inheritance and ii/ NF1 signs and symptoms, including CALMs, skinfold freckling, Lisch nodules, cutaneous and subcutaneous, plexiform and spinal neurofibromas, OPGs and other neoplasms, skeletal and cardiac abnormalities, development and education levels, presence/absence of Noonan features and segmental phenotype.

Fifteen major clinical features of NF1 were selected for the genotype-phenotype correlation study (Tables 1-3). Individuals with missing data for a particular sign and/or symptom were

197 classified as “unknown” or “not specified” and consequently excluded from that part of the
198 genotype-phenotype analysis. Most features were identified by physical examination;
199 ophthalmologic examination for Lisch nodules and imaging to detect asymptomatic OPGs and
200 spinal neurofibromas was not performed in most individuals. Brain and spine/whole body MRI
201 was done mainly in individuals with signs and/or symptoms indicative of OPGs or internal/spinal
202 neurofibromas; however, depending on institutional policies, some individuals were screened by
203 MRI despite the absence of symptoms. Noonan phenotype was diagnosed if at least two of the
204 following features were observed: short stature, hypertelorism, low set ears, webbed neck, ptosis,
205 midface hypoplasia or pulmonic stenosis. To evaluate short stature and macrocephaly, the World
206 Health Organization (WHO) and the Center for Disease Control (CDC) growth charts and the
207 Gerhard Nellhaus’ curve [25] were used as previously described [15]. Short stature and
208 macrocephaly were defined as height below or equal to the 3rd percentile (PC≤3) and as head
209 circumference equal or above the 98th percentile (PC≥98), respectively. For cognitive
210 impairment/learning disabilities, individuals with attention deficit disorder (ADD) and/or
211 attention deficit hyperactivity disorder (ADHD) but normal development were classified as
212 normal.

213 To establish a genotype-phenotype association we used the same approach as previously
214 described [15]. We compared the phenotypes of individuals with missense mutations affecting
215 codons 844-848 with the cohort of 169 individuals with missense mutations affecting p.Arg1809
216 [14, 15, 26-28], 47 individuals heterozygous for c.2970_2972del (p.Met992del) mutations [13]
217 and previously described large scale NF1 individual cohorts with “classic” NF1 [8, 29-41].

218 This study was approved by the Institutional Review Boards of all participating institutions
219 offering clinical genetic testing.

Molecular analysis

In the Medical Genomics Laboratory at UAB comprehensive *NF1* mutation screening using an RNA-based approach complemented by DNA-dosage analysis was performed as previously described [42, 43]. The status of the specific familial mutation in relatives was ascertained by bidirectional Sanger sequencing (ABI PRISM 3730, Life Technologies).

The nomenclature of the mutations is based on *NF1* mRNA sequence NM_000267.3 according to the recommendations of the Human Genome Variation Society (HGVS). For exon numbering we used the NCBI numbering, followed by the historical numbering in square brackets originally developed by the NF1 community [43].

In silico prediction of effect of missense mutations

Seven software programs were used to predict the effects of missense variants: two online *in silico* prediction tools (CADD v1.3 and PolyPhen-2) and five complementary tools (Grantham Difference, SIFT v4.0.3, SpliceSiteFinder-like, MaxEntScan, NNSplice v0.9 and Human Splicing Finder v2.4.1) embedded in Alamut visual software v2.9.0 (Interactive Biosoftware). The presence or absence of the variants was checked in population databases, including the Genome Aggregation Database (gnomAD), 1000 Genomes and the Exome Variant Server (EVS) as well as in disease databases: the Leiden Open Variation Database (LOVD), ClinVar and the Human Gene Mutation Database (HGMD) (last accessed May 2017). Evolutionary conservation for human neurofibromin NP_000258.1 residues 804-950 was evaluated using Clustal software

v2.0.12. The palindromic sequences and quadruplex forming G-Rich sequences (QGRS) were identified by Palindrome search and QGRS Mapper, respectively.

Interpretation of variant pathogenicity was performed based on the American College of Medical Genetics (ACMG) recommendations [44].

Statistical analysis

For univariate analysis, two-tailed Fisher's exact test was used to compare categorical variables with a p-value <0.05 considered as statistically significant. The resulting p-values were adjusted for multiple comparisons using Benjamini-Hochberg (B-H) procedure with false discovery rates (FDRs) of 0.05 and 0.01. The 95% confidence interval (CI) was also calculated when appropriate. All statistical analyses were performed with GraphPad and VassarStats softwares.

Results

Description of missense mutations affecting codons 844-848

Exon 21 [16] is the largest *NF1* exon (441 nucleotides), and in it we identified, besides the missense variants affecting the codons 844-848, a total of 19 different missense variants in 35 unrelated individuals from the UAB cohort. Fourteen of these alterations were classified as variants of uncertain significance (8/19) or likely benign (6/19) and reported 1-3 times in the UAB cohort (Figure S2). Only five variants were classified as pathogenic (4/19) or likely pathogenic (1/19) according to the current recommendations [44]. Region 844-848 in exon 21 [16] stood out due to its high frequency of variants compared with the neighboring codons,

indicating functional importance (Figures S2 and S3). A similar distribution and spectrum of missense alterations in the *NF1* exon 21 [16] was observed in the publicly available databases (ClinVar, LOVD and HGMD). Besides a clear cluster of recurrent variants in codons 844-848, other alterations spread over the entire exon 21 [16] were mostly classified as variants of uncertain significance and reported 1-2 times in these databases (Figure S2). The frequency of this cluster of variants in AA844-848 is ~0.8% (67/8400) in unrelated *NF1* mutation-positive individuals from the UAB cohort, second only to the p.Arg1809 (~1.2%), and therefore represents a significant hotspot for missense mutations within *NF1*.

In the 129 unrelated individuals reported here, we identified 12 different *NF1* missense alterations affecting one of five neighboring codons in exon 21 [16] (Table 1 and Figure 1). Within the group of individuals with p.Gly848Arg, two different substitutions were observed: c.2542G>A (6/14) and c.2542G>C (8/14). Detailed characteristics of the identified missense mutations are shown in Tables S2 - S4 and Figure 1. All variants identified in this study with confirmed origin of the variant were submitted to the LOVD and ClinVar databases. Based on the data accumulated in this report (Table S1 and Table S2), these variants can all be classified as pathogenic (Table S4) according to current recommendations [44].

Among the aforementioned variants, 8/12 were present in the LOVD database with 5/8 classified as pathogenic [c.2533T>C (p.Cys845Arg), c.2536G>C (p.Ala846Pro), c.2537C>A (p.Ala846Asp), c.2540T>C (p.Leu847Pro), c.2543G>A (p.Gly848Glu)] and 3/8 as variants of uncertain significance [c.2534G>A (p.Cys845Tyr), c.2540T>G (p.Leu847Arg), c.2542G>C (p.Gly848Arg)]. Eight of the 12 were present in ClinVar, including 3/8 classified as pathogenic [c.2531T>G (p.Leu844Arg), c.2540T>C (p.Leu847Pro), c.2542G>C (p.Gly848Arg)], 1/8 as likely pathogenic [c.2534G>A (p.Cys845Tyr)], 1/8 as a variant of uncertain significance

[c.2533T>C (p.Cys845Arg)], and 3/5 with no significance provided [c.2530C>T (p.Leu844Phe), c.2531T>C (p.Leu844Pro), c.2543G>A (p.Gly848Glu)] (Table S2 and Table S3). One individual (UAB-R4444) with c.2531T>A (p.Leu844His) carried another novel alteration c.2524G>A; assuming both variants reside in cis, this alteration should be described as c.2524_2531delinsAGCTTCCA (p.Gly842_Leu844delinsSerPheHis). None of these variants, except for c.2531T>G (p.Leu844Arg), has been reported in 129,639 unrelated controls of the gnomAD and EVS databases or in the 1000 Genomes Project; c.2531T>G (p.Leu844Arg) was reported once in Latino (the variant's frequency in all populations is 0.00041%). Based on *in silico* analysis all alterations are predicted to be deleterious (SIFT) and probably or possibly damaging (PolyPhen-2). Additionally, CADD classified all variants as more likely to have deleterious effects (range: 22.6 to 31). In contrast to results of *in silico* analysis, suggesting a possible effect of two identified alterations (c.2542G>A and c.2543G>A) on splicing through creation of a novel exonic splice acceptor sequence, transcript analysis and sequencing showed a minor effect on splicing only for c.2542G>A in three individuals (UAB-R9493, UAB-R1474 and UAB-R0008), i.e. low levels of r.2410_2543del. The other individuals with c.2542G>A screened with an RNA-based approach (UAB-R3513 and UAB-R4476) in whom no missplicing was observed, also carried the nearby benign variant c.2544G>A (p.Gly848=) (rs17883704) with both variants proven to reside in cis through next-generation sequencing. As missplicing was only observed in individuals carrying c.2542G>A in the absence of rs17883704 (Figure S4), rs17883704 is hypothesized to have a modifying effect. All missense mutations, except for c.2536G>C (p.Ala846Pro) were proven to be *de novo* in at least one proband; a total of 26 probands with unaffected parents were proven to have a *de novo* mutation, but formal confirmation of paternity/maternity by identity testing was only pursued for individuals tested in

the Netherlands (ROT-R02233, ROT-R22853 and ROT-R17435). Additionally, 7/12 missense mutations [c.2530C>T (p.Leu844Phe), c.2533T>C (p.Cys845Arg), c.2536G>C (p.Ala846Pro), c.2537C>A (p.Ala846Asp), c.2540T>C (p.Leu847Pro), c.2542G>C (p.Gly848Arg) and c.2543G>A (p.Gly848Glu)] segregated with the phenotype (at least one individual per family) in 23 affected first-degree relatives from 15 families (Table S1, Table S2 and Figure S5). Finally, all missense mutations affecting amino acids 844-848 are located in a highly conserved region of the CSRD (amino acids 543-909; Figure S6). Besides cysteine at position 845 that is conserved up to Zebrafish, all remaining amino acids are evolutionarily conserved up to *Drosophila melanogaster* (Ala846 and Gly848) and even to yeast IRA1 and/or IRA2 (Leu844 and Leu847). In chimpanzee, rat and mouse all amino acids from 775 to 856 are fully evolutionarily conserved. None of these variants has been functionally characterized.

Demographic and clinical characterization of the studied cohort

A total of 162 individuals from 129 unrelated families were enrolled in the study, including 37/129 (28.7%) familial and 89/129 (69%) sporadic cases; 3/129 (2.3%) individuals had an unknown family history (ROT-R13734, ROT-R89874 and CAR-R8012M6). Detailed demographic and clinical descriptions of the individuals included in the study are shown in Table 1, Table S1 and Figure S5.

The complete phenotypic checklist forms were collected from 151/162 individuals (93.2%). Of these, 125/151 (82.8%) fulfilled the NIH diagnostic criteria and 118/151 (78.2%) fulfilled the NIH diagnostic criteria if family history was excluded as a criterion. Among 26/151 individuals who did not fulfill the NIH diagnostic criteria (with 20/26 being ≤8 years), multiple CALMs-

only (>5) were present in 16/26, <6 CALMs-*only* were present in 8/26 and 2/26 did not have any pigmentary manifestations, but had externally visible plexiform neurofibromas (UAB-R9135 and UG-R5831) (Table S5). CALMs-*only* (<6) were observed mostly in individuals with a missense mutation at codon 848 [6/8 with c.2542G>C (p.Gly848Arg), 1/8 with c.2543G>A (p.Gly848Glu) and 1/8 with c.2534G>A (p.Cys845Tyr)].

Among 102 individuals ≥9 years, more than 5 CALMs and skinfold freckling were present in 79.8% (79/99) and 80% (76/95), respectively (Table 1). Both clinical features were found in 71.6% (68/95) of cases. Out of 20 individuals ≥9 years with only few or absolute lack of CALMs (Table S1), 11 cases fulfilled the NIH diagnostic criteria based on presence of other clinical signs, such as skinfold freckles, Lisch nodules, neurofibromas and/or osseous lesions (UG-R0781, UAB-R3618-M, MIL-R192/982-F, UAB-R4476, MIL-R999/399, MIL-R999/399-M, ROT-R95424, UG-R923-S, UAB-R3237, MAN-R95417G, MAN-R95417G-C). Among these individuals, 8/11 (72.7%) carried a missense mutation at codon 848. Lisch nodules were reported less frequently (42/98 all ages, but in 34/60 ≥9 years).

Cutaneous and subcutaneous neurofibromas were found in 68.1% (47/69 ≥19 years) and 50.8% (33/65 ≥19 years) of the cases, respectively. Thirty adults had both types of tumors (30/64 ≥19 years, 46.9%). Ten individuals ≥17 years had >100 cutaneous and/or subcutaneous nodules, including a 47-year-old man previously reported [45] with >1,400 neurofibromas (individual counts of externally visible neurofibromas; BRA-R38) and a 17-year-old young woman (ROT-R1CMUL) with >500 cutaneous neurofibromas, >100 subcutaneous neurofibromas and >100 intradermal neurofibromas. Nine out of ten individuals with a very high number of neurofibromas carried a missense mutation at codon 847: c.2540T>G (p.Leu847Arg) [2/9] or c.2540T>C (p.Leu847Pro) [7/9, including two individuals with metastasized MPNSTs]. In 16

cases with "several" neurofibromas a more precise estimated number was not reported. Eight individuals (UAB-R5776, UAB-R3618, UAB-R4624, UAB-R7447, UAB-R1002; UAB-R1037-M, UAB-R3237, PAD-R500-C1) were reported to have a single cutaneous or subcutaneous nodule (none histopathologically confirmed); these individuals were considered as "negative for the criterion of neurofibromas" as ≥2 cutaneous/subcutaneous neurofibromas are required according to the NIH clinical criteria.

Forty-five percent of the individuals ≥9 years had known plexiform neurofibromas (41/92 ≥9 years; 47/143 all ages), including externally visible (n=36) and internal (n=5) tumors. For six cases, the information was not provided whether plexiform neurofibromas were identified clinically or by MRI. Among all individuals with plexiform neurofibromas, 31/47 presented with one plexiform tumor and 16/47 with ≥2 plexiform neurofibromas. Plexiform tumors were found in the head, face and neck area (35.7%, 25/70), limbs (34.3%, 24/70), trunk (17.1%, 12/70), back (n=3), abdomen (n=3), pelvis (n=2) and chest (n=1).

Symptomatic spinal neurofibromas visible by MRI were found in 15.2% of individuals (12/79 ≥9 years; 13/127 all ages). Forty asymptomatic individuals received MRI screening, leading to the identification of another seven cases with spinal tumors (Table S6). Approximately one-third of the individuals with spinal tumors (6/20) had fewer than 6 CALMs and no skinfold freckling, whereas in 60% (12/20) plexiform neurofibromas were observed (with 11/12 being externally visible).

Symptomatic OPGs, confirmed by MRI imaging, were found in 11/104 of individuals older than 5 years (10.6%), whereas asymptomatic OPGs were present in 16/52 additional individuals who underwent MRI examination (30.8% ≥5 years). In 19 of 27 symptomatic and asymptomatic OPGs, the detailed information about the tumor's location was collected, involving optic nerves

(2 symptomatic OPGs and 7 asymptomatic OPGs), chiasm (1 symptomatic OPG and 1 asymptomatic OPG) or both locations (6 symptomatic OPGs and 2 asymptomatic OPGs). Three children were diagnosed with a symptomatic OPG (PAD-R300) or asymptomatic OPGs (UAB-R3714 and UAB-R3513) before age 4 years (Table S7).

Skeletal abnormalities were frequently reported (48/144 all ages) and included scoliosis (27/144 all ages, but 20/64 ≥19 years) and pectus anomalies (10/144 all ages: pectus carinatum 6/10 and excavatum 4/10). In addition, long bone dysplasia (n=4), pseudarthrosis (n=2), tibial dysplasia (n=1), bone cysts (n=2), sphenoid wing dysplasia (n=2), ulnar aplasia, likely representing the severe end of ulnar pseudarthrosis with bone resorption and absence of ulnar bone (n=1), dural ectasia (n=1), 4^{th} lumbar vertebrae fragmentation (n=1), bowed long bones (n=1), clinodactyly (n=1), postaxial polydactyly (n=1) and cherubism (n=1) were observed in the studied group.

Noonan syndrome features were observed in 10/134 (7.5%) individuals. One previously reported individual [46] (UAB-R624) with a family history of *PTPN11*-positive (MIM: 176876) Noonan syndrome (MIM: 163950) had a severe phenotype of pulmonic stenosis and aortic coarctation, dysmorphic features (high forehead, hypertelorism, downslanting palpebral fissures, short neck with a low posterior hair line), short stature, pectus carinatum, >5 CALMs, axillary and inguinal freckling, plexiform and cutaneous neurofibromas, symptomatic OPG with signs of hydrocephalus. Besides the familial *PTPN11* c.1529A>G (p.Gln510Arg) inherited from the individual's father, the *NF1* missense mutation c.2531T>G (p.Leu844Arg) was found *de novo* in the proband (Figure S5). In other individuals with Noonan syndrome features (UAB-R2696, UAB-R5001, UAB-R3725 and UAB-R4676) no pathogenic or likely pathogenic variants in Noonan-related disorders genes (*PTPN11* [MIM:176876], *SPRED1* [MIM:609291], *BRAF* [MIM: 164757], *CBL* [MIM: 165360], *HRAS* [MIM: 190020], *KRAS* [MIM: 190070], *MAP2K1*

399 [MIM: 176872], *MAP2K2* [MIM: 601263], *NRAS* [MIM: 164790], *RAF1* [MIM: 164760],
400 *SHOC2* [MIM: 602775], *SOS1* [MIM: 182530], *RIT1* [MIM: 609591], *RASA2* [MIM: 601589]
401 and *SOS2* [MIM: 601247]) were identified. Cardiovascular abnormalities observed in the studied
402 group included hypertension (n=7, one related to renal artery stenosis), pulmonic stenosis (n=2),
403 mitral valve stenosis, atrial septal defect, ventricular septal defect, Moyamoya disease,
404 pericarditis carcinomatosa, mitral valve insufficiency, mild pulmonic insufficiency and
405 hypertrophic cardiomyopathy (each observed in a single individual). Short stature (PC≤3) and
406 macrocephaly (PC≥98) were found in 15/91 (16.5%) and 36/98 (36.7%), respectively. Of the 138
407 cases with provided developmental data, 56 individuals had abnormal development presenting
408 with at least one of the following forms of cognitive impairment and/or learning difficulties:
409 learning disabilities (n=43), developmental delay (n=30), speech delay (n=8), ADD (n=8),
410 ADHD (n=10), motor delay (n=5), autism spectrum (n=2), Asperger syndrome (n=1). Seven
411 individuals had significant global developmental delay with/without speech delay, learning
412 difficulties and/or AD(H)D, including one with a Full Scale Intelligence Quotient (FSIQ) score
413 59. Additionally, three individuals were reported to have frequent migraine headaches, two had
414 epilepsy and/or psychiatric problems.

415 For 139/162 individuals, data on the presence or absence of tumors other than neurofibromas and
416 OPGs was available. Thirteen of 139 (9.4%) individuals were diagnosed with malignant
417 neoplasms (Table S8), including embryonal rhabdomyosarcoma (3/13), MPNST (7/13, including
418 one woman with MPNST and *BRCA1/2*-negative breast cancer), colon cancer (1/13), medullary
419 thyroid carcinoma (1/13) and juvenile myelomonocytic leukemia (JMML) (1/13). Individuals
420 ≥14 years old with c.2540T>C (p.Leu847Pro) had a higher number of malignant neoplasms
421 compared to individuals carrying other missense mutations in the studied region (p=0.0448;

Table S9). Moreover, this mutation was present in most cases with MPNST (5/7), except for one each carrying c.2543G>A (p.Gly848Glu) or c.2530C>T (p.Leu844Phe). Four of seven individuals with MPNST died before age 30 years (Table S8). Hypothalamic glioma (n=1), lipoma (n=1), cerebral tumors (n=3), non-ossifying fibroma (n=2) and odontogenic fibroma (n=1) were also reported.

The frequency of clinical features in individuals heterozygous for missense mutations affecting one of five neighboring codons 844-848 is presented in Table 2. A lower number of CALMs, freckling and cutaneous neurofibromas was observed in cases with missense mutations at codon 848 (all $p<0.0001$; Table S9); however, these individuals had a higher prevalence of symptomatic spinal neurofibromas ($p=0.0012$; Table S9).

Taken together, a severe phenotype, including at least one of the following features: plexiform and/or symptomatic spinal neurofibromas, symptomatic OPGs, malignant neoplasm or osseous lesions was observed in 75% of adult NF1 individuals (56/75 ≥19 years; Table 2).

Comparison of clinical features observed in the studied cohort with individuals heterozygous for p.Arg1809 and p.Met992del mutations and cohort of individuals with "classic" NF1 phenotype

Comparison of clinical features of the studied group with the *NF1* p.Arg1809 and p.Met992del cohorts as well as previously described large-scale cohorts of individuals with "classic" NF1 is shown in Table 3. The complete list of adjusted p-values with FDRs at 0.05 and 0.01 after B-H correction for multiple testing is presented in Table S10. All p-values ≤0.0125 and p-values

443 ≤0.0012 remained statistically significant after applying the B-H correction at FDRs of 0.05 and
444 0.01, respectively.

445 In the current study, we observed a significantly higher number of major external plexiform
446 neurofibromas compared with the *NF1* p.Arg1809 and the *NF1* p.Met992del cohorts, as well as
447 "classic" NF1 population (all p<0.0001; statistically significant after B-H correction at FDR of
448 0.01). Importantly, while none of the individuals carrying the p.Arg1809 and p.Met992del had
449 external plexiform, cutaneous and/or subcutaneous neurofibromas, ~71% of the individuals ≥19
450 years with a missense mutation affecting codons 844-848 had cutaneous and/or subcutaneous
451 neurofibromas (p<0.0001; statistically significant after B-H correction at FDR of 0.01) and
452 ~39% of the individuals ≥9 years had externally visible plexiform neurofibromas (p<0.0001;
453 statistically significant after B-H correction at FDR of 0.01). Compared with p.Arg1809,
454 p.Met992del and "classic" NF1 cohorts, at least 5-fold greater prevalence of symptomatic spinal
455 neurofibromas was reported in the studied group (0-2.1% vs. 10.2%) which was statistically
456 significant at FDR of 0.01 for the general NF1 population (p<0.0001) and at FDR of 0.05 for the
457 p.Arg1809 cohort (p=0.0022).

458 Symptomatic and asymptomatic OPGs were more frequent compared to individuals with
459 p.Arg1809, p.Met992del and "classic" NF1, with symptomatic and asymptomatic OPGs
460 statistically increased after B-H correction at FDR of 0.05 in the 844-848 cohort compared to the
461 "classic" NF1 cohorts (p=0.0125 and p=0.0043, respectively) and at FDR of 0.01 compared with
462 the p.Arg1809 cohort (p=0.0002 and p<0.0001, respectively). The overall prevalence of
463 malignant neoplasms, other than neurofibromas and OPGs, was also higher in the studied group
464 compared to a large cohort of "classic" NF1 individuals (9.4% vs. 3.4%; p=0.0061, statistically
465 significant at FDR of 0.05 after B-H correction).

Additionally, the AA844-848 cohort had a significantly increased frequency of skeletal abnormalities compared to individuals with p.Arg1809 and "classic" NF1 phenotypes (both statistically significant after B-H correction at FDR of 0.05), regardless of the age. Scoliosis was reported more frequently compared with p.Arg1809 individuals (31.3% vs. 12.5% in ≥19 years), but this difference was not statistically significant after B-H correction.

The prevalence of CALMs was lower than in p.Arg1809 and p.Met992del cohorts (both significant at FDR of 0.05 after B-H correction), while skinfold freckles occurred more common in "classic" NF1 cohorts than in the studied group (significant at FDR of 0.01 after B-H correction). Noonan syndrome features were significantly less frequent in the studied group compared to individuals with p.Arg1809 (significant at FDR of 0.01 after B-H correction). In line with this finding, pulmonic stenosis was very rarely observed in the cohort (1.8% vs. 10.6% in the p.Arg1809 cohort; significant at FDR of 0.05 after B-H correction). All cohorts, except for the p.Met992del, shared a similar frequency of cognitive impairment and/or learning difficulties (~45%).

Discussion

We present 162 individuals heterozygous for a constitutional *NF1* missense mutation in one of five neighboring codons 844-848 who have a high prevalence of a severe NF1 phenotype, including plexiform and/or symptomatic spinal neurofibromas, symptomatic OPGs, other malignant neoplasms, as well as bone abnormalities. The frequency of the cluster of these mutations is ~0.8% (67/8400) in unrelated *NF1* mutation-positive individuals from the UAB cohort, second only to the p.Arg1809 (~1.2%) among the missense variants.

One of the most severe complications in NF1 individuals are clinically apparent plexiform neurofibromas affecting 15-30% of the NF1 general population [8, 35, 47-50]. In this study, externally visible plexiform neurofibromas were found in ~39% of individuals ≥9 years, therefore significantly higher compared with p.Arg1809 and p.Met992del and "classic" NF1 cohorts (significant at FDR of 0.01 after B-H correction; Table 3 and Table S10). Individuals in this study did not undergo whole body MRI; therefore the frequency provided here is a likely underestimate, as internal asymptomatic plexiform neurofibromas were not accounted for.

As plexiform neurofibromas have been suggested to be associated with a higher lifetime risk for the development of MPNSTs [50-53], the finding of MPNSTs in 5% (7/139) of the affected in our cohort, which is twice as high as reported by Huson et al. (1989) in the South-East Wales cohort [29, 30], is in line with expectations.

Approximately 24-40% of NF1 individuals develop spinal neurofibromas [36, 40, 52], but they are most often asymptomatic and not detectable by physical examination. The estimated prevalence of *symptomatic* spinal neurofibromas in the general NF1 population is less than 2% [8, 35, 36]. In the current study, a high number of individuals with *symptomatic* spinal neurofibromas was reported, compared to the "classic" NF1 cohorts (statistically significant at FDR of 0.01 after B-H correction): 13/127 (10.2%) for all ages and 12/79 (15.2%) for ≥ 9 years. Kluwe et al. (2003) [19] suggested that spinal neurofibromas cause symptoms mainly in older cases (mean age 32.8 years), but four of thirteen symptomatic individuals in our cohort were below age 18 (range: 7-17 years). In 40 individuals who underwent MRI examination, an additional seven cases with asymptomatic spinal neurofibromas were found. Among all affected individuals, five persons belonged to two previously reported multi-generation families (UG-R923 and MAN-R95417G) where the spinal tumors segregated within the family [17, 22]. For

two relatives of these probands the spinal neurofibromas were only recognized after MRI, although the tumor burden was extensive. None of the individuals had >5 CALMs, including 2/5 who had <6 CALMs and 3/5 had none. This rare form of NF1 is called "familial spinal neurofibromatosis" (FSNF).

Plexiform and spinal tumors as well as subcutaneous neurofibromas are associated with a severe NF1 phenotype and may result in significant morbidity in children and adults [54, 55]. OPGs, the most common brain tumors in children, are another complication in the general NF1 population [56]. The *overall* prevalence of OPGs in the NF1 population is ~11-20% [40, 50, 57]; however, only ~30% of these individuals have clinically *symptomatic* OPGs and present with impaired visual acuity, visual field loss, abnormal color vision, squint, proptosis and/or hypothalamic dysfunction [49]. Most symptomatic OPGs are diagnosed before age 7 years [57] with the mean age of 5 years [58]. In the studied group, symptomatic OPGs were found in 11/104 (10.6%) of individuals ≥ 5 years, which is more frequent compared with p.Arg1809 and p.Met992del cohorts (none of the individuals had OPGs) and with "classic" NF1 population (3.9%); however, after applying the B-H correction only the result of comparison with p.Arg1809 cohort and the general NF1 population remained statistically significant at FDR of 0.05 (Table 3 and Table S10). Furthermore, there was a higher prevalence of asymptomatic OPGs in 16/52 (30.8%) individuals ≥5 years who underwent MRI examination (statistically significant at FDR of 0.01).

Individuals with NF1 are at higher risk to develop specific malignancies compared with the general population, significantly increasing mortality [59, 60]. Besides the high-grade gliomas, the most common malignancies in NF1 children are rhabdomyosarcomas, JMML, and neuroblastomas, but accurate estimates on prevalence are not available due to the rarity of these tumors [61, 62]. Based on the data provided by Sung et al. (2004) [63] and Crucis et al. (2015)

[64], the prevalence of rhabdomyosarcomas in children with NF1 is estimated at 0.4-0.5%, while Chang and Shannon (2012) [65] reported that the individual risk of JMML in NF1 is ~0.04%. In the studied group, three NF1 children younger than 5 years developed embryonal rhabdomyosarcomas, including one individual, now >26 years, who survived both a rhabdomyosarcoma and astrocytoma grade II, diagnosed at the age two and 15 years, respectively. Furthermore, one five-year-old girl (out of 50 children ≤ 8 years) presented with <6 CALMs and JMML. This girl was heterozygous for two pathogenic *NF1* mutations in the blood, c.2542G>A (p.Gly848Arg), as well as c.1246C>T (p.Arg416*), with p.Gly848Arg being the first hit given the absence of p.Arg416* in buccal swabs, indicating somatic mosaicism for p.Arg416*. An UK population-based hospital admission and death certificate study found that individuals with NF1 have, after excluding the well-established risks of nervous systems tumors, a 2.7-fold increased risk of developing cancers of the esophagus, stomach, colon, liver, lung, bone, thyroid, malignant melanoma, non-Hodgkin lymphoma, chronic myeloid leukemia, breast and ovary [66]. In the current study, we noted recurrent malignant tumors, such as MPNSTs (7/139; 5%) (Table S1 and Table S8). Among these individuals, one 44-year-old woman previously described [67] with the missense mutation c.2540T>C (p.Leu847Pro) had MPNST, *BRCA1/2*-negative (MIM: 113705 and MIM: 600185) breast cancer as well as a high number of cutaneous neurofibromas (>100). In addition, one individual developed a medullary thyroid carcinoma and three first-degree relatives of a Belgian proband with c.2540T>C (p.Leu847Pro) died from malignancies (a metastasized colon adenocarcinoma and two MPNSTs, both deceased before age 26 year). Taken together, the *overall* prevalence of malignant neoplasms in the studied group was substantially higher than in the published datasets of the general NF1 population (significant at FDR of 0.05 after B-H correction; Table 3 and Table S10).

Furthermore, specifically mutation p.Leu847Pro seems to confer a high predisposition to develop malignant tumors compared to other missense variants reported in this study ($p<0.0448$; Table S9), although the CADD score of this variant is not the highest among the studied region (only 26.1). Given the predominance of the p.Leu847Pro mutations in the studied cohort (70/162 individuals), larger datasets are required to further refine the increased tumor risk associated with the other mutations within the studied region.

Skeletal abnormalities, including long bone dysplasia with or without pseudarthrosis, scoliosis, sphenoid wing dysplasia, bone cysts, including cherubism, non-ossifying fibromas and osseous giant cell lesions, hand anomalies, anterior chest wall anomalies and short stature, can lead to serious clinical consequences and significant morbidity [68]. We observed a clear overall increase in the number of skeletal anomalies compared with p.Arg1809 (FDR of 0.05 after B-H correction) and the general NF1 population (FDR of 0.01 after B-H correction). As many as 33.3% of the NF1 individuals (48/144) presented with one or more osseous lesion, scoliosis (n=27) and pectus anomalies (n=10) being most frequent (18.8% and 6.9%, respectively). The *overall* frequency would be higher if individuals with short stature (40.3%; 58/144) are included. Rarely reported complications possibly associated with *NF1* status included cherubism, chronic arthritis of multiple joints with elbow contractures, clinodactyly of the 3-5th toes, postaxial polydactyly and ulnar aplasia, likely representing the severe end of ulnar pseudarthrosis with bone resorption and absence of the ulnar bone. Interestingly, the latter has been reported only in two NF1 cases [69]. Mild to moderate scoliosis was reported in only 18% of *NF1*-positive individuals with bilateral neurofibromas of all spinal roots [18]; however, in our study we observed co-occurrence of scoliosis and spinal tumors in 45% (9/20) of individuals with confirmed symptomatic or asymptomatic spinal neurofibromas (not necessarily affecting all

dorsal roots) (Table S6). An additional 11 individuals had scoliosis without evidence of spinal neurofibromas by MRI (Table S1).

Cohorts of individuals with *NF1* missense mutations affecting codons 844-848 and "classic" NF1 population shared a similar frequency for short stature and macrocephaly. Noonan syndrome (NS) features were rarely observed in the studied group compared with the p.Arg1809 cohort (significant at FDR of 0.01 after B-H correction). In line with previous studies [8, 35, 40, 70], intellectual disability, developmental delay, and/or learning difficulties were frequently observed in the current study (40.6%).

Among the 129 unrelated probands with a missense mutation affecting codons 844-848, p.Leu847Pro and p.Gly848Arg are the most recurrent variants, found in 58 and 14 unrelated individuals, respectively (Table S2 and Figure 1). Both alterations are associated with a severe NF1 phenotype, including a high prevalence of plexiform neurofibromas and skeletal abnormalities, compared to the general NF1 population. However, missense mutations at p.Gly848 predispose with a greater frequency to symptomatic or asymptomatic spinal tumors, which were found in ~70% of probands carrying the p.Gly848Arg or p.Gly848Glu mutations (9/13 ≥9 years, but in 9/10 ≥9 years who received MRI screening), that is slightly higher than in individuals presenting with a severe phenotype caused by a total *NF1* deletion (8/13 ≥9 years) [71]. Several of the severely affected individuals with a missense mutation at p.Gly848 had only few or no pigmentary skin findings. So far, ~100 cases have been reported with the true "spinal NF" phenotype [18] and these individuals more frequently carry a splice site or missense mutation spread over the entire *NF1* coding region. So far, no single mutation has been correlated with this severe clinical presentation. We provide the specific genotype-phenotype association between a particular *NF1* mutation and the spinal phenotype. Individuals with

603 missense mutations at p.Gly848 appear to constitute a distinct group of NF1 individuals with a
604 high prevalence of symptomatic spinal neurofibromas and a clear decrease of pigmentary
605 manifestations (CALMs and skinfold freckles) as well as cutaneous neurofibromas (Table 2 and
606 Table S9). Because of the limited number of individuals ≥9 years old with the missense
607 mutations at codons 844-846, it is still difficult to establish a genotype-phenotype correlation
608 among these cohorts; however, so far these variants also seem to be associated with a severe
609 phenotype, including a high prevalence of plexiform neurofibromas in the p.Cys845 and
610 p.Ala846 cohorts (57.1% and 66.7%, respectively) and OPGs in p.Leu844 cohort (~24% for both
611 symptomatic and asymptomatic OPGs in ≥5 years). At this moment, it cannot be excluded that
612 two specific genotype-phenotype correlations exist within this small region of *NF1* with the *NF1*
613 codon 847 associated with an increased risk for malignant neoplasia and the *NF1* codon 848
614 associated with a high prevalence of symptomatic spinal neurofibromas. The current study,
615 however, intended to show that the whole region of 844-848 codons stood out due to its high
616 frequency of variants compared with the neighboring codons, indicating functional importance.
617 In addition, the cluster of missense mutations here described, although located outside the GRD
618 important for RAS-regulation, is clearly associated with a severe phenotype, not reported so far
619 in literature. As the current study necessarily still underestimates the internal tumor burden, as
620 systematic whole body imaging was not performed, close clinical management seems warranted
621 for individuals presenting with a missense variant affecting the AA844-848.

622 As NF1 is known for its variable expressivity and age-dependency, it is challenging to establish
623 genotype-phenotype correlations. Although we performed a comparative analysis on a large
624 well-described cohort using a standardized phenotypic data collection form, one limitation of the
625 study is that clinical information was collected by physicians from different referral centers,

although all were NF1 specialists. Data in this and the previously reported p.Arg1809 cohort were "double-checked" through verification of the originally submitted phenotypic checklist forms and subsequent update of the clinical notes, so data should be highly accurate.

Clinical variability, both inter- and intrafamilial, has been widely reported in the past two decades [72-74]. Although significant progress has been made over the last twenty years, the mechanisms underlying this phenotypic heterogeneity only gradually start to be unraveled. The factors contributing to the phenotypic variability include: i/ age-dependency of some of the NF1 features [30, 75, 76]; ii/ timing, cell-of-origin and number of second hits in specific cells, resulting in presence and number of CALMs, freckling, tibial dysplasia, neurofibromas and other tumors [77]; iii/ post-zygotic mosaicism for the first *NF1* hit in mosaic individuals [77]; iv/ the enormous *NF1* allelic heterogeneity [78]; v/ occasional presence of two different *NF1* pathogenic variants segregating within a family (see MAD-R9.232; Table S1 and Figure S5) or the occurrence of two independent mutations (one in *NF1* and the other in a different gene) within an individual (see UAB-R624 with the *NF1/PTPN11* mutations and UF-R1 with the *NF1/KIT* mutations; Table S1); vi/ modifying genes [79] and vii/ environmental factors (e.g. number of pregnancies) [80]. To date, two studies have identified potential modifying genes, unlinked to the *NF1* locus, associated with the severity of NF1 presentation [81, 82]. Pasmant et al. (2011) demonstrated that a high number of plexiform neurofibromas has been significantly associated with allele T of SNP rs2151280 of *ANRIL* (MIM: 613149) [81]. Pemov et al. (2014) reported a correlation of two common SNPs (rs4660761 and rs7161), located between *DPH2* (MIM: 603456) and *ATP6V0B* (MIM: 603717), as well as of SNP rs1800934 in *MSH6* (MIM: 600678) with the number of CALMs [82]. Further studies are needed to confirm these findings.

Missense mutations affecting *NF1* codons 844-848 described in this study are clearly pathogenic and individuals with these missense mutations have a statistically higher risk of developing spinal neurofibromas, plexiform neurofibromas and OPGs. Functional studies in mutant mice harboring the missense mutation c.2542G>C (p.Gly848Arg), however, did not recapitulate this human phenotype, as neither optic pathway gliomas [24] nor plexiform neurofibromas [23] developed. Western blot analysis showed that c.2542G>C (p.Gly848Arg) resulted in 38-50% reduction of neurofibromin levels [23, 24]. These mutations reside outside the GRD (amino acids 1203-1549), known to have tumor-suppressor activity through downregulation of members of the Ras family of small GTP-binding proteins. Although *NF1* was cloned in 1990, the cellular functions performed by this huge 2818-amino acid multi-domain protein are still incompletely understood. The cluster of recurrent missense mutations involving AA844-848 described in the current study are located within the CSRD (amino acids 543-909), located N-terminal to the GRD. The CSRD domain, originally described by Fahsold et al (2000) [83], is likely functionally important, which is further implied by the presence of multiple missense variants in this segment of the gene in NF1 individuals. The 3D structure of this region has not been resolved and its precise functions and interactors have not been described. Ras GAP activity is enhanced through phosphorylation by Protein Kinase Cα of serine and threonine residues within this domain [84]. Based on the 2D-modeling of the CSRD using PredictProtein server [85], the region 831-847 might form the C-part of a helix and be buried in the protein. Missense mutations affecting codons 844-848, especially those substituting smaller hydrophobic amino acids to large ones, may result in breaking of the helix and exposure of the buried protein domain, consequently affecting the function of the protein. No functional studies confirming the aforementioned bioinformatics analysis have been performed, however. In any case, missense mutations in this

region seem to act through a loss-of-function mechanism and not gain-of-function or dominant-negative, at least in melanocytes and JMML. Indeed, the c.2540T>C (p.Leu847Pro) was observed as a "second hit" in one CALM, biopsied from a 13.5-year-old girl with >5 CALMs and skinfold freckling carrying the *NF1* constitutional mutation c.5547-1G>A (Table S11), confirming that two hits are required to cause a phenotypic effect. Additionally, we reported a five-year-old girl with JMML (UAB-R9493; Table S1) who carried two pathogenic *NF1* mutations in the blood: c.2542G>A (p.Gly848Arg) as a "first hit" mutation and c.1246C>T (p.Arg416*) as a "second hit". There is a need to improve our understanding of the physiological functions of neurofibromin and to determine how each domain regulates the function of this protein.

Six amino acids in the region AA804-950 are evolutionarily conserved down to yeast (IRA1 and IRA2), Leu844, Gly849, Leu852, Glu923, Leu933 and Phe934 (Figure S6), and would therefore be expected to be of particular functional importance [86]. Only Leu844 and Leu933 have however been observed in NF1 individuals to predispose to recurrent missense mutations (HGMD, LOVD, ClinVar and our cohort). The tumorigenic potential of AA844 is further highlighted by identification of somatic mutations in the COSMIC database: one glioma with p.Leu844Pro, one glioma and four malignant melanomas with p.Leu844Phe.

Palindromic structures belong to the non-B DNA-structures and are often the site of replication errors resulting in substitutions [87]. The *NF1* missense mutation hotspot (AA844-848) is located in the highly conserved amino acid region, suggesting it is functionally important. The genomic sequence encoding the human *NF1* AA845-853 is a part of two palindromic structures (Figure S7); therefore the high rate of recurrent missense mutations affecting Leu847 and Gly848 may partially be due these being both located in the loop of the palindrome. In *NF1* exon

21 [16] other palindromic nucleotide sequences, specifying the amino acid residues AA828-832, AA865-868, AA908-911 and AA933-937 are observed, resulting in four additional stem-loop structures. However, these structures do not predispose to recurrent missense mutations as none were found either in the UAB, HGMD or LOVD cohort, except for c.2798T>C (p.Leu933Pro), whose location does not include the loop of the palindrome. The complex interplay between functional significance and genomic architecture needs to be considered when analyzing the recurrence of mutations.

Although only few clear genotype-phenotype correlations have been so far reported [12-15], the data here presented show that additional clinically relevant *NF1* genotype-phenotype correlations exist. A renewed interest in such studies is needed to come to a timely unfolding of additional correlations, as so far only the surface has been scratched. This will require close collaboration between NF1 clinicians and molecular geneticists. The lack of discovery of more specific genotype-phenotype correlations may be partly due to the methodological approach, including lumping mutations in large categories (truncating versus microdeletion, splice, missense mutations) [88, 89]. Identification of *mutation-specific* genotype-phenotype correlations depends on the datasets size with a large number of individuals, preferentially postpubertal, carrying the *same* non-truncating constitutional mutation, with the associated phenotype recorded in a standardized way. As there are only a limited number of truly recurrent non-truncating mutations, prioritization on individuals carrying such recurrent mutations is indicated. Although each of the recurrent mutation affects only a small percentage of *NF1* individuals (3-8% with the microdeletion type I, ~0.8% with p.Met992del, ~1.2% with the p.Arg1809 missense mutation and ~0.8% for the cluster of missense mutations affecting codons 844-848), together they may affect counseling and surveillance in a significant fraction of the *NF1* population.

In conclusion, the present findings indicate that missense mutations affecting one of five neighboring codons 844-848 located outside the GAP-related domain are an important risk factor for a severe phenotype in NF1 individuals. We report that these individuals have a high prevalence of plexiform and/or spinal neurofibromas, symptomatic and asymptomatic OPGs, malignant neoplasms and skeletal abnormalities. A severe phenotype was observed in 75% of adult NF1 individuals with these mutations, clearly demonstrating that missense mutations outside the GRD can be associated with a severe clinical presentation. The current study identified a genotype-phenotype correlation in this region that may be valuable in the management and genetic counseling of a significant number of NF1 individuals. These data suggest there is a potential need for increased disease surveillance in individuals with these mutations enabling genotype driven personalized medicine.

Supplemental Data

Supplemental Data include seven figures and eleven tables and can be found with this article online at XXXXX.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgements

We thank the individuals and their families for participating in this study. This work was supported by the Children's Tumor Foundation by the Isaac and Sadie Fuchs Genotype-Phenotype Study (to L.M.M.) and by internal funds from the Medical Genomics Laboratory at UAB. Parts of this work were presented during 17th European Neurofibromatosis Meeting (September, 8-11, 2016, Padova-Abano Terme, Italy) and the Children's Tumor Foundation NF Conference (June, 10-13, 2017, Washington, DC, USA).

Dr. Magdalena Koczkowska is also affiliated with the Department of Biology and Medical Genetics at the Medical University of Gdansk in Poland.

Web Resources

1000 Genomes: http://www.1000genomes.org

CADD: http://cadd.gs.washington.edu/

ClinVar: https://www.ncbi.nlm.nih.gov/clinvar/

Clustal software v2.0.12: http://www.clustal.org/clustal2/

COSMIC: http://cancer.sanger.ac.uk/cosmic

EVS: http://evs.gs.washington.edu/EVS/

gnomAD: http://gnomad.broadinstitute.org/

GraphPad: http://graphpad.com

HGMD: http://www.hgmd.cf.ac.uk/ac/index.php

HGVS: http://varnomen.hgvs.org

LOVD: http://www.lovd.nl/NF1

OMIM: https://www.omim.org/

Palindrome search: http://bioinfo.cs.technion.ac.il/projects/Engel-Freund/new.html

PolyPhen-2: http://genetics.bwh.harvard.edu/pph2

QGRS Mapper: http://bioinformatics.ramapo.edu/QGRS/index.php

SIFT: http://sift.jcvi.org

VassarStats: http://vassarstats.net

References

1. Lammert, M., Friedman, J.M., Kluwe, L., Mautner, V.F. (2005). Prevalence of neurofibromatosis 1 in German children at elementary school enrollment. Arch. Dermatol. *141*, 71-74.
2. Evans, D.G., Howard, E., Giblin, C., Clancy, T., Spencer, H, Huson, S.M., Lalloo, F. (2010). Birth incidence and prevalence of tumor-prone syndromes: estimates from a UK family genetic register service. Am. J. Med. Genet. A *152A*, 327-332.
3. Uusitalo, E., Leppävirta, J., Koffert, A., Suominen, S., Vahtera, J., Vahlberg, T., Pöyhönen, M., Peltonen, J., Peltonen, S. (2015). Incidence and mortality of neurofibromatosis: a total population study in Finland. J. Invest. Dermatol. *135*, 904-906.
4. Friedman, J.M. (2014). Neurofibromatosis 1. In GeneReviews®, M.P. Adam, H.H. Ardinger, R.A. Pagon et al., ed. (Seattle WA: University of Washington, Seattle), http://www.ncbi.nlm.nih.gov/books/NBK1109/.

5. Ballester, R., Marchuk, D., Boguski, M., Saulino, A., Lechter, R., Wigler, M., Collins, F. (1990). The NF1 locus encodes a protein functionally related to mammalian GAP and yeast IRA proteins. Cell *63*, 851-859.
6. DeClue, J.E., Cohen, B.D., Lowy, D.R. (1991). Identification and characterization of the neurofibromatosis type 1 protein product. Proc. Natl. Acad. Sci. U.S.A. *88*, 9914-9918.
7. Peltonen, S., and Pöyhönen, M. (2012). Clinical diagnosis and atypical forms of NF1. In Neurofibromatosis type 1. Molecular and cellular biology, M. Upadhyaya, D.N. Cooper, ed. (Springer-Verlag Berlin Heidelberg), pp. 17-30.
8. Huson, S.M., Harper, P.S., Compston, D.A. (1988). Von Recklinghausen neurofibromatosis. A clinical and population study in south-east Wales. Brain. *111*, 1355-1381.
9. National Institutes of Health Consensus Development Conference. (1988). Neurofibromatosis. Conference statement. Arch. Neurol. *45*, 575-578.
10. Brems, H., Chmara, M., Sahbatou, M., Denayer, E., Taniguchi, K., Kato, R., Somers, R., Messiaen, L., De Schepper, S., Fryns, J.P., et al. (2007). Germline loss-of-function mutations in SPRED1 cause a neurofibromatosis 1-like phenotype. Nat. Genet. *39*, 1120-1126.
11. Messiaen, L., Yao, S., Brems, H., Callens, T., Sathienkijkanchai, A., Denayer, E., Spencer, E., Arn, P., Babovic-Vuksanovic, D., Bay, C., et al. (2009). Clinical and mutational spectrum of neurofibromatosis type 1-like syndrome. JAMA. *302*, 2111-2118.
12. Kehrer-Sawatzki, H., Mautner, V.F., Cooper, D.N. (2017). Emerging genotype-phenotype relationships in patients with large NF1 deletions. Hum. Genet. *136*, 349-376.
13. Upadhyaya, M., Huson, S.M., Davies, M., Thomas, N., Chuzhanova, N., Giovannini, S., Evans, D.G., Howard, E., Kerr, B., Griffiths, S., et al. (2007). An absence of cutaneous neurofibromas associated with a 3-bp inframe deletion in exon 17 of the NF1 gene (c.2970-

2972 delAAT): evidence of a clinically significant NF1 genotype-phenotype correlation. Am. J. Hum. Genet. *80*, 140-151.

14. Pinna, V., Lanari, V., Daniele, P., Consoli, F., Agolini, E., Margiotti, K., Bottillo, I., Torrente, I., Bruselles, A., Fusilli, C., et al. (2015). p.Arg1809Cys subsitution in neurofibromin is associated with a distinctive NF1 phenotype without neurofibromas. Eur. J. Hum. Genet. *23*, 1068-1071.
15. Rojnueangnit, K., Xie, J., Gomes, A., Sharp, A., Callens, T., Chen, Y., Liu, Y., Cochran, M., Abbott, M.A., Atkin, J., et al. (2015). High incidence of Noonan syndrome features including short stature and pulmonic stenosis in patients carrying NF1 missense mutations affecting p.Arg1809: genotype-phenotype correlation. Hum. Mutat. *36*, 1052-1063.
16. Pulst, S.M., Riccardi, V.M., Fain, P., Korenberg, J.R. (1991). Familial spinal neurofibromatosis: clinical and DNA linkage analysis. Neurology. *41*, 1923-1927.
17. Burkitt-Wright, E.M., Sach, E., Sharif, S., Quarrell, O., Carroll, T., Whitehouse, R.W., Upadhyaya, M., Huson, S.M., Evans, D.G.R. (2013). Can the diagnosis of NF1 be excluded clinically? A lack of pigmentary findings in families with spinal neurofibromatosis demonstrates a limitation of clinical diagnosis. J. Med. Genet. *50*, 606-613.
18. Ruggieri, M., Polizzi, A., Spalice, A., Salpietro, V., Caltabiano, R., D'Orazi, V., Pavone, P., Pirrone, C., Magro, G., Platania, N., et al. (2015). The natural history of spinal neurofibromatosis: a critical review of clinical and genetic features. Clin. Genet. *87*, 401-410.
19. Kluwe, L., Tatagiba, M., Fünsterer, C., Mautner, V.F. (2003). NF1 mutations and clinical spectrum in patients with spinal neurofibromas. J. Med. Genet. *40*, 368-371.
20. Messiaen, L., Riccardi, V., Peltonen, J., Maertens, O., Callens, T., Karvonen, S.L., Leisti, E.L., Koivunen, J., Vandenbroucke, I., Stephens, K., et al. (2003). Independent NF1

mutations in two large families with spinal neurofibromatosis. J. Med. Genet. *40*, 122-126.

21. Upadhyaya, M., Spurlock, G., Kluwe, L., Chuzhanova, N., Bennett, E., Thomas, N., Guha, A., Mautner, V. (2009). The spectrum of somatic and germline NF1 mutations in NF1 patients with spinal neurofibromas. Neurogenetics. *10*, 251-263.
22. Pascual-Castroviejo, I., Pascual-Pascual, S.I., Velazquez-Fragua, R., Botella, P., Viaño, J. (2007). Familial spinal neurofibromatosis. Neuropediatrics. *38*, 105-108.
23. Li, K., Turner, A.N., Chen, M., Brosius, S.N., Schoeb, T.R., Messiaen, L.M., Bedwell, D.M., Zinn, K.R., Anastasaki, C., Gutmann, D.H., et al. (2016). Mice with missense and nonsense NF1 mutations display divergent phenotypes compared with human neurofibromatosis type 1. Dis. Model. Mech. *9*, 759-767.
24. Toonen, J.A., Anastasaki, C., Smithson, L.J., Gianino, S.M., Li, K., Kesterson, R.A., Gutmann, D.H. (2016). NF1 germline mutation differentially dictates optic glioma formation and growth in neurofibromatosis-1. Hum. Mol. Genet. *25*, 1703-1713.
25. Nellhaus, G. (1968). Head circumference from birth to eighteen years. Practical composite international and interracial graphs. Pediatrics. *41*, 106-114.
26. Nyström, A.M., Ekvall, S., Strömberg, B., Holmström, G., Thuresson, A.C., Annerén, G., Bondeson, M.L. (2009). A severe form of Noonan syndrome and autosomal dominant café-au-lait spots - evidence for different genetic origins. Acta Paediatr. *98*, 693-698.
27. Ekvall, S., Sjörs, K., Jonzon, A., Vihinen, M., Annerén, G., Bondeson, M.L. (2014). Novel association of neurofibromatosis type 1-causing mutations in families with neurofibromatosis-Noonan syndrome. Am. J. Med. Genet. A. *164A*, 579-587.
28. Santoro, C., Maietta, A., Giugliano, T., Melis, D., Perrotta, S., Nigro, V., Piluso, G. (2015). Arg^{1809} substitution in neurofibromin: further evidence of a genotype-phenotype correlation

847 in neurofibromatosis type 1. Eur. J. Hum. Genet. *23*, 1460-1461.
848 29. Huson, S.M., Compston, D.A., Clark, P., Harper, P.S. (1989a). A genetic study of von
849 Recklinghausen neurofibromatosis in south east Wales. I. Prevalence, fitness, mutation rate,
850 and effect of parental transmission on severity. J. Med. Genet. *26*, 704-711.
851 30. Huson, S.M., Compston, D.A., Harper, P.S. (1989b). A genetic study of von Recklinghausen
852 neurofibromatosis in south east Wales. II. Guidelines for genetic counseling. J. Med. Genet.
853 *26*, 712-721.
854 31. Listernick, R., Charrow, J., Greenwald, M., Mets, M. (1994). Natural history of optic
855 pathway tumours in children with neurofibromatosis type 1: a longitudional study. J. Pediatr.
856 *125*, 63-66.
857 32. Van Es, S., North, K.N., McHugh, K., De Silva, M.. (1996). MRI findings in children with
858 neurofibromatosis type 1: a prospective study. Pediatr. Radiol. *26*, 478-487.
859 33. Friedman, J.M., and Birch, P.H. (1997). Type 1 neurofibromatosis: a descriptive analysis of
860 the disorder in 1,728 patients. Am. J. Hum. Genet. *70*, 138-143.
861 34. Cnossen, M.H., de Goede-Bolder, A., van den Broek, K.M., Waasdorp, C.M., Oranje, A.P.,
862 Stroink, H., Simonsz, H.J., van den Ouweland, A.M., Halley, D.J., Niermeijer, M.F. (1998).
863 A prospective 10 year follow up study of patients with neurofibromatosis type 1. Arch. Dis.
864 Child. *78*, 408-412.
865 35. McGaughran, J.M., Harris, D.I., Donnai, D., Teare, D., MacLoad, R., Westerbeek, R.,
866 Kingston, H., Super, M., Harris, R., Evans, D.G. (1999). A clinical study of type 1
867 neurofibromatosis in north west England. J. Med. Genet. *36*, 197-203.
868 36. Thakkar, S.D., Feigen, U., Mautner, V.F. (1999). Spinal tumours in neurofibromatosis type
869 1: an MRI study of frequency, multiplicity and variety. Neuroradiology. *41*, 625-629.

37. Lin, A.E., Birch, P.H., Korf, B.R., Tenconi, R., Niimura, M., Poyhonen, M., Armfield Uhas, K., Sigorini, M., Virdis, R., Romano, C., et al. (2000). Cardiovascular malformations and other cardiovascular abnormalities in neurofibromatosis 1. Am. J. Med. Genet. *95*, 108-117.

38. Blazo, M.A., Lewis, R.A., Chintagumpala, M.M., Frazier, M., McCluggage, C., Plon, S.E. (2004). Outcomes of systematic screening for optic pathway tumors in children with Neurofibromatosis Type 1. Am. J. Med. Genet. A *127A*, 224-229.

39. Khosrotehrani, K., Bastuji-Garin, S., Riccardi, V.M., Birch, P., Friedman, J.M., Wolkenstein, P. (2005). Subcutaneous neurofibromas are associated with mortality in neurofibromatosis 1: a cohort study of 703 patients. Am. J. Med. Genet. A *132A,* 49-53.

40. Plotkin, S.R., Bredella, M.A., Cai, W., Kassarjian, A., Harris, G.J., Esparza, S., Merker, V.L., Munn, L.L., Muzikansky, A., Askenazi, M., et al. (2012). Quantitative assessment of whole-body tumor burden in adult patients with neurofibromatosis. PLoS One. *7*, e35711.

41. Blanchard, G., Lafforgue, M.P., Lion-François, L., Kemlin, I., Rodriguez, D., Castelnau, P., Carneiro, M., Meyer, P., Rivier, F., Barbarot, S., et al. (2016). Systematic MRI in NF1 children under six years of age for the diagnosis of optic pathway gliomas. Study and outcome of a French cohort. Eur. J. Paediatr. Neurol. *20*, 275-281.

42. Messiaen, L.M., Callens, T., Mortier, G., Beysen, D., Vandenbroucke, I., Van Roy, N., Spelemen, F., Paepe, A.D. (2000). Exhaustive mutation analysis of the NF1 gene allows identification of 95% of mutations and reveals a high frequency of unusual splicing defects. Hum. Mutat. *15*, 541-555.

43. Messiaen, L.M., and Wimmer, K. (2012). Mutation analysis of the *NF1* gene by cDNA-based sequencing of the coding region. In Advances in Neurofibromatosis Research, K.S.G. Cunha and M. Geller, ed. (Nova Science Publishers, Inc.), pp. 89-108.

44. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hedge, M., Lyon, E., Spector, E., et al. (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. Genet. Med. *17*, 405-424.

45. Cunha, K.S., Oliveira, N.S., Fausto, A.K., Cruz de Souza, C., Gros, A., Bandres, T., Idrissi, Y., Merlio, J-P., Soares de Moura Neto, R.S., Silva, R., et al. (2016). Hybridization capture-based next-generation sequencing to evaluate coding sequence and deep intronic mutations in the *NF1* gene. Genes (Basel) *7*, 133.

46. Bertola, D.R., Pereira, A.C., Passetti, F., de Oliveira, P.S., Messiaen, L., Gelb, B.D., Kim, C.A., Krieger, J.E. (2005). Neurofibromatosis-Noonan syndrome: molecular evidence of the concurrence of both disorders in a patient. Am. J. Med. Genet. A *136A*, 242-245.

47. Tonsgard, J.H., Kwak, S.M., Short, M.P., Dachman, A.H. (1998). CT imaging in adults with neurofibromatosis-1: frequent asymptomatic plexiform lesions. Neurology. *50*, 1755-1760.

48. Waggoner, D.J., Towbin, J., Gottesman, G., Gutmann, D.H. (2000). Clinic-based study of plexiform neurofibromas in neurofibromatosis 1. Am. J. Med. Genet. *92*, 132-135.

49. Ferner, R.E., Huson, S.M., Thomas, N., Moss, C., Willshaw, H., Evans, D.G., Upadhyaya, M., Towers, R., Gleeson, M., Steiger, C., et al. (2007). Guidelines for the diagnosis and management of individuals with neurofibromatosis 1. J. Med. Genet. *44*, 81-88.

50. Duong, T.A., Sbidian, E., Valeyrie-Allanore, L., Vialette, C., Ferkal, S., Hadj-Rabia, S., Glorion, C., Lyonnet, S., Zerah, M., et al. (2011). Mortality associated with neurofibromatosis 1: a cohort study of 1895 patients in 1980-2006 in France. Orphanet. J. Rare Dis. *6*, 18.

916 51. Evans, D.G., Baser, M.E., McGaughran, J., Sharif, S., Howard, E., Moran, A. (2002).
917 Malignant peripheral nerve sheath tumours in neurofibromatosis type 1. J. Med. Genet. *39*,
918 311-314.
919 52. Tucker, T., Wolkenstein, P., Revuz, J., Zeller, J., Friedman, J.M. (2005). Association
920 between benign and malignant peripheral nerve sheath tumors in NF1. Neurology. *65*, 205-
921 211.
922 53. McCaughan, J.A., Holloway, S.M., Davidson, R., Lam, W.W. (2007). Further evidence of
923 the increased risk for malignant peripheral nerve sheath tumour from a Scothish cohort of
924 patients with neurofibromatosis type 1. J. Med. Genet. *44*, 463-466.
925 54. Korf, B.R. (1999). Plexiform neurofibromas. Am. J. Med. Genet. *89*, 31-37.
926 55. Prada, C.E., Rangwala, F.A., Martin, L.J., Lovell, A.M., Saal, H.M., Schorry, E.K., Hopkin,
927 R.J. (2012). Pediatric plexiform neurofibromas: impact on morbidity and mortality in
928 neurofibromatosis type 1. J. Pediatr. *160*, 461-467.
929 56. Jahraus, C.D., and Tarbell, N.J. (2006). Optic pathway gliomas. Pediatr. Blood Cancer. *46*,
930 586-596.
931 57. Listernick, R., Ferner, R.E., Liu, G.T., Gutmann, D.H. (2007). Optic pathway gliomas in
932 neurofibromatosis-1: controversies and recommendations. Ann. Neurol. *61*, 189-198.
933 58. Nicolin, G., Parkin, G., Mabbott, D., Hargrave, D., Bartels, U., Tabori, U., Rutka, J., Buncic,
934 J.R., Bouffet, E. (2009). Natural history and outcome of optic pathway gliomas in children.
935 Pediatr. Blood Cancer. *53*, 1231-1237.
936 59. Rasmussen, S.A., and Friedman, J.M. (2000). NF1 gene and neurofibromatosis 1. Am. J.
937 Epidemiol. *151*, 33-40.
938 60. Uusitalo, E., Rantanen, M., Kallionpää, R.A., Pöyhönen, M., Leppävirta, J., Ylä-Outinen, H.,

939 Riccardi, V.M., Pukkala, E., Pitkäniemi, J., Peltonen, S., et al. (2016). Distinctive cancer
940 associations in patients with neurofibromatosis type 1. J. Clin. Oncol. *34*, 1978-1986.
941 61. Brems, H., Beert, E., de Ravel, T., Legius, E. (2009). Mechanisms in the pathogenesis of
942 malignant tumours in neurofibromatosis type 1. Lancet. Oncol. *10*, 508-515.
943 62. Patil, S., and Chamberlain, R.S. (2012). Neoplasms associated with germline and somatic
944 NF1 gene mutations. Oncologist. *17*, 101-116.
945 63. Sung, L., Anderson, J.R., Arndt, C., Raney, R.B., Meyer, W.H., Pappo, A.S. (2004).
946 Neurofibromatosis in children with Rhabdomyosarcoma: a report from the Intergroup
947 Rhabdomyosarcoma study IV. J. Pediatr. *144,* 666-668.
948 64. Crucis, A., Richer, W., Brugières, L., Bergeron, C., Marie-Cardine, A., Stephan, J.L., Girard,
949 P., Corradini, N., Munzer, M., Lacour, B., et al. (2015). Rhabdomyosarcomas in children
950 with neurofibromatosis type 1: a national historical cohort. Pediatr. Blood Cancer. *62*, 1733-
951 1738.
952 65. Chang, T., Shannon, K. (2012). *NF1* mutations in hematologic cancers. In Neurofibromatosis
953 type 1. Molecular and cellular biology, M. Upadhyaya, D.N. Cooper, ed. (Springer-Verlag
954 Berlin Heidelberg), pp. 469-485.
955 66. Seminog, O.O., and Goldacre, M.J. (2013). Risk of benign tumours of nervous system, and
956 of malignant neoplasms, in people with neurofibromatosis: population-based record-linkage
957 study. Br. J. Cancer. *108*, 193-198.
958 67. McPherson, J.R., Ong, C.K., Ng, C.C., Rajasegaran, V., Heng, H.L., Yu, W.S., Tan, B.K.,
959 Madhukumar, P., Teo, M.C., Ngeow, J. et al. (2015). Whole-exome sequencing of breast
960 cancer, malignant peripheral nerve sheath tumor and neurofibroma from a patient with
961 neurofibromatosis type 1. Cancer Med. *4*, 1871-1878.

962 68. Stevenson, D.A., and Yang, F.C. (2011). The musculoskeletal phenotype of the
963 RASopathies. Am. J. Med. Genet. C Semin. Med. Genet. *157C*, 90-103.
964 69. Vargiami, E., Zafeiriou, D.I., Bantouraki, M. (2004). Ulnar hypolasia and neurofibromatosis
965 type 1. J. Pediatr. *145*, 859.
966 70. Krab, L.C., de Goede-Bolder, A., Aarsen, F.K., Pluijm, S.M., Bouman, M.J., van der Geest,
967 J.N., Leguin, M., Catsman, C.E., Arts, W.F., Kushner, S.A., et al. (2008). Effect on
968 simvastatin on cognitive functioning in children with neurofibromatosis type 1: a randomized
969 controlled trial. JAMA. *300*, 287-294.
970 71. Mautner, V.F., Kluwe, L., Friedrich, R.E., Roehl, A.C., Bammert, S., Högel, J., Spöri, H.,
971 Cooper, D.N., Kehrer-Sawatzki, H. (2010). Clinical characterisation of 29 neurofibromatosis
972 type-1 patients with molecularly ascertained 1.4 Mb type-1 NF1 deletions. J. Med. Genet. *47*,
973 623-630.
974 72. Easton, D.F., Ponder, M.A., Huson, S.M., Ponder, B.A. (1993). An analysis of variation in
975 expression of neurofibromatosis (NF) type 1: evidence for modifying genes. Am. J. Hum.
976 Genet. *53*, 305-313.
977 73. Szudek, J., Joe, H., Friedman, J.M. (2002). Analysis of intrafamilial phenotypic variation in
978 neurofibromatosis 1 (NF1). Genet. Epidemiol. *23*, 150-164.
979 74. Sabbagh, A., Pasmant, E., Laurendeau, I., Parfait, B., Barbadot, S., Guillot, B., Combemale,
980 P., Ferkal, S., Vidaud, M., Aubourg, P., et al. (2009). Unravelling the genetic basis of
981 variable clinical expression in neurofibromatosis 1. Hum. Mol. Genet. *18*, 2768-2778.
982 75. DeBella, K., Szudek, J., Friedman, J.M. (2000). Use of the national institutes of health
983 criteria for diagnosis of neurofibromatosis 1 in children. Pediatrics *105*, 608-614.
984 76. Williams, V.C., Lucas, J., Babcock, M.A., Gutmann, D.H., Korf, B., Maria, B.L. (2009).

Neurofibromatosis type 1 revisited. Pediatrics *123*, 124-133.

77. Messiaen, L., and Xie, J. (2012). *NF1* germline and somatic mosaicism. In Neurofibromatosis type 1. Molecular and cellular biology, M. Upadhyaya, D.N. Cooper, ed. (Springer-Verlag Berlin Heidelberg), pp. 151-172.

78. Messiaen, L.M., and Wimmer, K. (2008). NF1 mutational spectrum. In Neurofibromatoses, D. Kaufmann, ed. (Monogr. Hum. Genet. Basel, Karger), pp. 63-77.

79. Pasmant, E., Vidaud, D., Wolkenstein, P. (2012). Modifier genes in NF1. In Neurofibromatosis type 1. Molecular and cellular biology, M. Upadhyaya, D.N. Cooper, ed. (Springer-Verlag Berlin Heidelberg), pp. 269-285.

80. Terry, A.R., Barker, F.G. II, Leffert, L., Bateman, B.T., Souter, I., Plotkin, S.R. (2013). Neurofibromatosis type 1 and pregnancy complications: a population-based study. Am. J. Obstet. Gynecol. *209*, 46.e1-46.e8.

81. Pasmant, E., Sabbagh, A., Vidaud, M., Bièche, I. (2011). ANRIL, a long, noncoding RNA, is an unexpected major hotspot in GWAS. FASEB J. *25*, 444-448.

82. Pemov, A., Sung, H., Hyland, P.L., Sloan, J.L., Ruppert, S.L., Baldwin, A.M., Boland, J.F., Bass, S.E., Lee, H.J., Jones, K.M., et al. (2014). Genetic modifiers of neurofibromatosis type 1-associated café-au-lait macule count identified using multi-platform analysis. PLoS Genet. *10*, e1004575.

83. Fahsold, R., Hoffmeyer, S., Mischung, C., Gille, C., Ehlers, C., Kücükceylan, N., Abdel-Nour, M., Gewies, A., Peters, H., Kaufmann, D., et al. (2000). Minor lesion mutational spectrum of the entire NF1 gene does not explain its high mutability but points to a functional domain upstream of the GAP-related domain. Am. J. Hum. Genet. *66*, 790-818.

84. Mangoura, D., Sun, Y., Li, C., Singh, D., Gutmann, D.H., Flores, A., Ahmed, M., Vallianatos, G. (2006). Phosphorylation of neurofibromin by PKC is a possible molecular switch in EGF receptor signaling in neural cells. Oncogene. *25*, 735-745.

85. Rost, B., Yachdav, G., Liu, J. (2004). The PredictProtein server. Nucleic Acids Res. *32*, W321-W326.

86. Ng, P.C., and Henikoff, S. (2001). Predicting deleterious amino acid subsitutions. Genome Res. *11*, 863-874.

87. Kamat, M.A., Bacolla, A., Cooper, D.N., Chuzhanova, N. (2016). A role for non-B DNA forming sequences in mediating microlesions causing human inherited disease. Hum. Mutat. *37,* 65-73.

88. Sabbagh, A., Pasmant, E., Imbard, A., Luscan, A., Soares, M., Blanché, H., Laurendeau, I., Ferkal, S., Vidaud, M., Pinson, S., et al. (2013). NF1 molecular characterization and neurofibromatosis type 1 genotype-phenotype correlation: the French experience. Hum. Mutat. *34*, 1510-1518.

89. van Minkelen, R., van Bever, Y., Kromosoeto, J.N., Withagen-Hermans, C.J., Nieuwlaat, A., Halley, D.J., van den Ouweland, A.M. (2014). A clinical and genetic overview of 18 years neurofibromatosis type 1 molecular diagnostics in the Netherlands. Clin. Genet. *85*, 318-327.

90. Zhou, X., Edmonson, M.N., Wilkinson, M.R., Patel, A., Wu, G., Liu, Y., Zhang, Z., Rusch, M.C., Parker, M., et al. (2016). Exploring genomic alteration in pediatric cancer using ProteinPaint. Nat. Genet. *48*, 4-6.

Table 1. Demographic and clinical characterization of the individuals with a missense mutation affecting codons 844-848.

	Codon 844			**Codon 845**			**Codon 846**			**Codon 847**			**Codon 848**			**All codons 844-848**			**Total**
Mutation [Proband:Relative]	c.2530C>T **(p.Leu844Phe)** [10:1] c.2531T>A **(p.Leu844His)** [2:0] c.2531T>C **(p.Leu844Pro)** [7:0] c.2531T>G **(p.Leu844Arg)** [6:0]			c.2533T>C **(p.Cys845Arg)** [3:1] c.2534G>A **(p.Cys845Tyr)** [8:0]			c.2536G>C **(p.Ala846Pro)** [1:2] c.2537C>A **(p.Ala846Asp)** [5:2]			c.2540T>C **(p.Leu847Pro)** [58:12] c.2540T>G **(p.Leu847Arg)** [8:0]			c.2542G>A **(p.Gly848Arg)** [6:0] c.2542G>C **(p.Gly848Arg)** [8:11] c.2543G>A **(p.Gly848Glu)** [7:4]						
Mutation positive individuals [Proband:Relative]	26 [25:1]			12 [11:1]			10 [6:4]			78 [66:12]			36 [21:15]			162 [129:33]			
Age group, years	**≤8**	**9-18**	**≥19**	**≤8**	**9-18**	**≥19**	**≤8**	**9-18**	**≥19**	**≤8**	**9-18**	**≥19**	**≤8**	**9-18**	**≥19**	**≤8**	**9-18**	**≥19**	**All ages**
Total	12	5	9	4	2	6	3	1	6	28	14	36	13	5	18	60	27	75	162
Proband:Relative	12:0	5:0	8:1	4:0	2:0	5:1	2:1	1:0	3:3	27:1	12:2	27:9	6:7	4:1	11:7	51:9	24:3	54:21	129:33
Age range, years	1-8	9-16	24-55	1-2	15-16	19-48	4-5	18	33-69	1-8	9-18	19-72	1-7	10-17	19-74	1-8	9-18	19-74	1-74
Male: Female	6:6	4:1	1:8	1:3	1:1	1:5	2:1	0:1	1:5	10:18	5:9	19:17	9:4	2:3	5:13	28:32	12:15	27:48	67:95
Fulfilling the NIH criteria if the family history is taken into account	10/11	4/5	9/9	2/4	1/2	4/5	3/3	1/1	6/6	17/28	14/14	35/36	4/11	4/5	17/18	36/57	24/27	71/74	131/158
Fulfilling the NIH criteria if solely taking the physical signs into account	10/11	4/5	9/9	2/4	1/2	4/5	2/3	1/1	6/6	17/28	14/14	33/36	4/11	4/5	13/18	35/57	24/27	65/74	124/158
> 5 CALMs	12/12	5/5	8/8	4/4	1/2	4/5	3/3	1/1	4/6	27/28	14/14	32/35	5/11	3/5	7/18	51/58	24/27	55/72	130/157
Freckling	10/10	4/5	6/7	0/4	1/2	4/5	2/2	1/1	5/5	12/23	13/13	31/34	4/10	3/5	8/18	28/49	22/26	54/69	104/144
Lisch nodules	2/9	1/4	4/4	0/1	0/0	1/2	0/1	0/1	2/2	4/19	3/9	17/19	2/8	0/5	6/14	8/38	4/19	30/41	42/98
Skeletal abnormalities [A]	2/11	2/5	5/9	2/4	1/2	2/4	0/2	0/1	0/5	3/25	3/14	17/28	3/11	3/5	5/18	10/53	9/27	29/64	48/144
Plexiform neurofibromas	0/11	2/5	3/9	0/3	2/2	2/5	0/2	1/1	1/2	6/24	3/13	19/33	0/11	1/5	7/17	6/51	9/26	32/66	47/143
Cutaneous neurofibromas [B]	1/11	1/5	7/9	0/4	0/2	3/4	0/2	1/1	4/5	1/26	4/14	28/33	1/11	1/5	5/18	3/54	7/27	47/69	57/150
Subcutaneous neurofibromas [B]	1/9	0/5	6/8	1/4	0/2	1/4	0/2	0/0	3/5	1/26	4/13	17/30	1/11	0/5	6/18	4/52	4/25	33/65	41/142
Cutaneous and subcutaneous [B]	0/9	0/5	5/8	0/4	0/2	1/3	0/2	0/0	3/5	0/25	1/13	17/30	0/11	0/5	4/18	0/51	1/25	30/64	31/140
Symptomatic spinal NF	0/10	0/3	0/8	0/2	1/2	0/4	0/2	0/0	0/2	1/23	1/13	2/27	0/11	1/4	7/16	1/48	3/22	9/57	13/127
Spinal neurofibromas by MRI [C]	0/1	0/0	0/5	0/0	1/2	1/1	0/1	0/0	0/1	1/5	2/6	3/16	0/1	2/3	10/11	1/8	5/11	14/34	20/53
Symptomatic OPGs [D]	1/11	1/5	0/9	0/3	0/2	0/5	1/3	1/1	0/3	2/25	1/13	2/27	1/11	1/5	1/13	5/53	4/26	3/57	12/136
Asymptomatic OPGs [E]	2/6	1/2	2/4	0/1	0/2	0/2	0/1	0/0	0/3	1/8	6/9	4/13	1/4	0/2	1/6	4/20	7/15	7/28	18/63
Other neoplasms [F]	1/11	0/4	1/8	0/2	0/1	0/4	0/2	0/1	0/3	1/24	3/14	11/34	2/11	1/5	1/15	4/50	4/25	13/64	21/139
Cognitive impairment and/or learning disabilities	3/11	3/4	0/6	1/4	0/2	3/4	3/3	0/1	1/5	10/26	7/13	12/26	5/11	5/5	3/17	22/55	15/25	19/58	56/138
Noonan syndrome features	0/9	1/5	1/8	0/2	1/1	0/4	0/2	0/1	0/4	3/27	0/13	3/26	1/10	0/5	0/17	4/50	2/25	4/59	10/134
Short stature [G]	1/7	0/2	0/4	0/3	1/1	0/1	0/2	0/0	1/2	0/11	3/10	4/21	3/10	0/3	2/14	4/33	4/16	7/42	15/91
Macrocephaly	2/11	1/4	1/2	1/3	0/1	0/0	2/2	0/0	0/2	8/21	2/11	10/17	3/11	1/4	5/9	16/48	4/20	16/30	36/98
Pulmonic stenosis	0/8	1/5	0/6	0/2	0/2	1/1	0/3	0/0	0/5	0/23	0/13	0/20	0/8	0/3	0/14	0/44	1/23	1/46	2/113

[A] All bone abnormalities included, that is, scoliosis (n=27), pectus excavatum (n=4), pectus carinatum (n=6), long bone dysplasia (n=4), pseudarthrosis (n=2), bone cysts (n=2), sphenoid wing dysplasia (n=2), ulnar aplasia, dural ectasia, 4th lumbar vertebrae fragmentation, bowed long bones, tibial dysplasia, clinodactyly, postaxial polydactyly and cherubism. [B] At least two cutaneous/subcutaneous neurofibromas were required to be considered as "positive for the criterion of neurofibromas". [C] The frequency of both symptomatic and asymptomatic spinal

neurofibromas in individuals who had done MRI examination. [D] The presence or absence of symptomatic OPGs was determined by ophthalmological examination and confirmed by MRI. [E] Including only individuals without signs of symptomatic OPGs who underwent MRI examination. [F] Including benign and malignant neoplasms, except for OPGs and neurofibromas. [G]As no specific growth curves are available for the Hispanic and Asian populations, Hispanic and Asian individuals were excluded as having short or normal stature.

Table 2. Frequency of clinical features in cohorts of individuals with a missense mutation affecting Leu844, Cys845, Ala846, Leu847 and Gly848.

NF1 feature	Number of individuals (%) [95% Confidence Interval]				
	Leu844	**Cys845**	**Ala846**	**Leu847**	**Gly848**
>5 CALMs	25/25 (100) [86.7-100]	9/11 (81.8) [52.3-94.9]	8/10 (80) [49-94.3]	73/77 (94.8) [87.4-98]	15/34 (44.1) [28.9-60.6]
Skinfold freckling [A]	10/12 (83.3) [55.2-95.3]	5/7 (71.4) [35.9-91.8]	6/6 (100) [61-100]	44/47 (93.6) [82.8-97.8]	11/23 (47.8) [29.2-67]
Lisch nodules	7/17 (41.2) [21.6-64]	1/3 (33.3) [6.2-79.2]	2/4 (50) [15-85]	24/47 (51.1) [37.2-64.7]	8/27 (29.6) [15.9-48.5]
Plexiform neurofibromas [A]	5/14 (35.7) [16.3-61.2]	4/7 (57.1) [25-84.2]	2/3 (66.7) [20.8-93.9]	22/46 (47.8) [34.1-61.9]	8/22 (36.4) [19.7-57]
Cutaneous neurofibromas [B]	7/9 (77.8) [45.3-93.7]	3/4 (75) [30.1-95.4]	4/5 (80) [37.6-96.4]	28/33 (84.9) [69.1-93.4]	5/18 (27.8) [12.5-50.9]
Subcutaneous neurofibromas [B]	6/8 (75) [40.9-92.9]	1/4 (25) [4.6-69.9]	3/5 (60) [23.1-88.2]	17/30 (56.7) [39.2-72.6]	6/18 (33.3) [16.3-56.3]
Symptomatic spinal neurofibromas [A]	0/11 (0) [0-25.9]	1/6 (16.7) [3-56.4]	0/2 (0) [0-65.8]	3/40 (7.5) [2.6-19.9]	8/20 (40) [21.9-61.3]
Spinal neurofibromas by MRI [A, C]	0/5 (0) [0-43.5]	2/3 (66.7) [20.8-93.9]	0/1 (0) [0-79.4]	5/22 (22.7) [10.1-43.4]	12/14 (85.7) [60.1-96]
Symptomatic OPGs, age ≥5 years [D]	1/21 (4.8) [0.9-22.7]	0/7 (0) [0-35.4]	2/5 (40) [11.8-76.9]	5/47 (10.6) [4.6-22.6]	3/24 (12.5) [4.3-31]
Asymptomatic OPGs, age ≥5 years [E]	4/10 (40) [16.8-68.7]	0/4 (0) [0-49]	0/3 (0) [0-56.2]	11/25 (44) [26.7-62.9]	1/10 (10) [1.8-40.4]
Other neoplasms [F]	2/23 (8.7) [2.4-26.8]	0/7 (0) [0-35.4]	0/6 (0) [0-39]	15/72 (20.8) [13.1-31.6]	4/31 (12.9) [5.1-28.9]
Bone abnormalities	9/25 (36) [20.3-55.5]	5/10 (50) [23.7-76.3]	0/8 (0) [0-32.4]	23/67 (34.3) [24.1-46.3]	11/34 (32.4) [19.1-49.2]
Noonan syndrome features	2/22 (9.1) [2.5-27.8]	1/7 (14.3) [2.6-51.3]	0/7 (0) [0-35.4]	6/66 (9.1) [4.2-18.5]	1/32 (3.1) [0.6-15.8]
Pulmonic stenosis	1/19 (5.3) [0.9-24.6]	1/5 (20) [3.6-62.5]	0/8 (0) [0-32.4]	0/56 (0) [0-6.4]	0/25 (0) [0-13.3]
Short stature [G]	1/13 (7.7) [13.7-33.3]	1/5 (20) [3.6-62.5]	1/4 (25) [4.6-69.9]	7/42 (16.7) [8.3-30.6]	5/27 (18.5) [8.2-36.7]
Macrocephaly	4/17 (23.5) [9.6-47.3]	1/4 (25) [4.6-69.9]	2/4 (50) [15-85]	20/49 (40.8) [28.2-54.8]	9/24 (37.5) [21.2-57.3]
Cognitive impairment and/or learning disabilities	6/21 (28.6) [13.8-50]	4/10 (40) [16.8-68.7]	4/9 (44.4) [18.9-73.3]	29/65 (44.6) [33.2-56.7]	13/33 (39.4) [24.7-56.3]
Severe phenotype, age ≥19 years [H]	7/9 (77.8) [45.3-93.7]	4/6 (66.7) [30-90.3]	1/6 (16.7) [3-56.4] [I]	32/36 (88.9) [74.7-95.6]	12/18 (66.7) [43.8-83.7]

[A] In individuals ≥9 years. [B] In individuals ≥19 years. [C] The frequency of both symptomatic and asymptomatic spinal neurofibromas in individuals who had done MRI examination. [D] The presence or absence of symptomatic OPGs was determined by ophthalmological examination and confirmed by MRI. [E] Including only individuals without signs of symptomatic OPGs who underwent MRI examination. [F] Including benign and malignant neoplasms, except for OPG and neurofibromas. [G] As no specific growth curves are available for the Hispanic and Asian populations, Hispanic and Asian individuals were excluded as having short or normal stature. [H] Individual was classified as having a severe phenotype if at least one of the following features was observed: plexiform and/or symptomatic spinal neurofibroma, symptomatic OPG, malignant neoplasm or osseous lesions. [I] Among individuals with a missense mutation affecting codon 846, the status of plexiform and spinal neurofibromas was known only for 2/6 individuals (UG-R0781-S and UG-R665-F), thus a severe phenotype cannot be excluded in the remaining four individuals with missing data.

Table 3. Comparison of clinical features of the studied group with the *NF1* Arg1809 cohort, the *NF1* Met992del cohort and large-scale cohorts of individuals with "classic" NF1.

NF1 feature	Number of individuals (%)				p value (2-tailed Fisher's exact test) *		
	AA844-848	Arg1809 [A]	Met992del [B]	Previous NF1 cohorts [C]	AA844-848 vs. Arg1809	AA844-848 vs. Met992del	AA844-848 vs. "classic" NF1
>5 CALMs	130/157 (82.8)	157/169 (92.9)	46/47 (97.9)	1537/1728 (89) [e]	**0.0060** ↘	**0.0067** ↘	0.0263 ↘
Skinfold freckling	104/144 (72.2)	95/161 (59)	32/47 (68.1)	1403/1667 (84.2) [e]	0.0164 ↗		**0.0007** ↘
Lisch nodules	42/98 (42.9)	12/120 (10)	3/38 (7.9)	729/1237 (58.9) [e]	**<0.0001** ↗	**<0.0001** ↗	**0.0028** ↘
Major external plexiform neurofibromas [D]	36/92 (39.1)	0/105 (0)	0/41 (0)	120/648 (18.5) [a,g]	**<0.0001** ↗	**<0.0001** ↗	**<0.0001** ↗
Cutaneous neurofibromas [E]	47/69 (68.1)	0/57 (0)	0/18 (0)	656/723 (90.7) [b,g,k,l]	**<0.0001** ↗	**<0.0001** ↗	**<0.0001** ↘
Subcutaneous neurofibromas [E]	33/65 (50.8)	0-5/57 (0-8.8) [I]	ND	297/515 (57.7) [g,k,l]	**<0.0001** ↗		
Symptomatic spinal neurofibromas [D, F]	12/79 (15.2)	0/40 (0)	1/41 (2.4)	2/119 (1.7) [a]	**0.0080** ↗	0.0341 ↗	**0.0004** ↗
	13/127 (10.2)	0/76 (0)	1/47 (2.1)	36/2058 (1.8) [a,g,h]	**0.0022** ↗		**<0.0001** ↗
Symptomatic OPGs, age ≥5 years [F, G]	11/104 (10.6)	0/114 (0)	0/46 (0)	7/180 (3.9) [a,d]	**0.0002** ↗	0.0186 ↗	0.0404 ↗
	12/136 (8.8)	0/139 (0)	0/47 (0)	64/1650 (3.9) [e]	**0.0002** ↗	0.0384 ↗	**0.0125** ↗
Asymptomatic OPGs, age ≥5 years [F, H]	16/52 (30.8)	0/35 (0)	ND	2/45 (4.4) [d]	**0.0001** ↗		**0.0012** ↗
	18/63 (28.6)	0/38 (0)		70/519 (13.5) [c,j,m]	**<0.0001** ↗		**0.0043** ↗
Other malignant neoplasms [J]	13/139 (9.4)	2/155 (1.3) [K]	0/47 (0)	18/523 (3.4) [g]	**0.0023** ↗	0.0409 ↗	**0.0061** ↗
Bone abnormalities [D, F]	38/91 (41.8)	14/72 (19.4)	8/41 (19.5)	14/96 (14.6) [a]	**0.0025** ↗	0.0174 ↗	**<0.0001** ↗
	48/144 (33.3)	21/126 (16.7)	9/47 (19.2)	144/948 (15.2) [a,f,g,l]	**0.0020** ↗		**<0.0001** ↗
Scoliosis [E]	20/64 (31.3)	6/48 (12.5)	2/18 (11.1)	51/236 (21.6) [b,l]	0.0241 ↗		
Noonan syndrome features	10/134 (7.5)	46/148 (31.1)	4 (all from 1 family)	57/1683 (3.4) [e]	**<0.0001** ↘		0.0276 ↗
Pulmonic stenosis	2/113 (1.8)	14/132 (10.6)	4/47 (8.5)	25/2322 (1.1) [i]	**0.0076** ↘		
Short stature	15/91 (16.5)	32/111 (28.8)	5/47 (10.6)	109/684 (15.9) [a,k]	0.0451 ↘		
Macrocephaly	36/98 (36.7)	31/107 (29)	4/45 (8.9)	239/704 (33.9) [a,k]		**0.0005** ↗	
Cognitive impairment and/or learning disabilities	56/138 (40.6)	80/159 (50.3)	8/47 (17)	190/424 (44.8) [a,g]		**0.0042** ↗	

*All bold and underlined p-values represent statistically significant p-values with false discovery rates of 0.05 (only bold **p-values**) and 0.01 (bold and underlined **p-values**), respectively after correction for multiple testing using Benjamini-Hochberg procedure (see details in Table S10). After applying the Benjamini-Hochberg correction p-values ≤0.0125 remained statistically significant at FDR of 0.05, while p-values ≤0.0012 were still be considered as significantly different at FDR of 0.01. The black arrows indicate the statistically significant differences of the NF1 clinical features prevalence between the studied group and the cohort(s) used for the comparison with the up and down arrows representing an increase and a decrease of the prevalence in the studied group, respectively.

[A] Based on data from Nyström et al. (2009) [26], Ekvall et al. (2014) [27], Pinna et al. (2015) [14], Rojnueangnit et al. (2015) [15] and Santoro et al. (2015) [28]. [B] Based on data from Upadhyaya et al. (2007) [13]. [C] Previous NF1 cohorts used for comparison: a/ Huson et al. (1988) [8]; b/ Huson et al. (1989a) [29] and Huson et al. (1989b) [30]; c/ Listernick et al. (1994) [31]; d/ Van Es et al. (1996) [32]; e/ Friedman and Birch (1997) [33]; f/ Cnossen et al. (1998) [34]; g/ McGaughran et al. (1999) [35]; h/ Thakkar et al. (1999) [36]; i/ Lin et al. (2000) [37]; j/ Blazo et al. (2004) [38]; k/ Khosrotehrani et al. (2005) [39]; l/ Plotkin et al. (2012) [40]; m/ Blanchard et al. (2016) [41]. [D] In individuals ≥9 years in this study and Arg1809, ≥10 years in Met992del and other studies. [E] In individuals ≥19 years in this study and Arg1809, ≥20 years in Met992del and other studies. [F] Second value is the frequency of a particular feature regardless of the individuals' age. [G] The presence or absence of symptomatic OPGs was determined by ophthalmological examination and confirmed by MRI. [H] Including only individuals without signs of symptomatic OPGs who underwent MRI examination. [I] Five individuals with few (1-6) small, subcutaneous "possible" neurofibromas, none were biopsied and therefore none have been histologically confirmed (Rojnueangnit et al., 2015) [15]. [J] Only malignant neoplasms, hence excluding neurofibromas and OPGs, have been taken into account. [K] Breast cancer (n=1) and Ewing sarcoma (n=1) were found in the *NF1* Arg1809 cohort (Rojnueangnit et al., 2015) [15]; no follow-up information on these individuals was available. **ND**: **no** **d**ata.

Figure 1. Spectrum of missense mutations affecting *NF1* codons 844-848 in the cohort of 129 probands (A) and 33 relatives (B).

Each number in circle corresponds with the total number of individuals heterozygous for a specific mutation. The black dotted lines on the panels present the regions 844-848. The figure was prepared using the ProteinPaint application [90].