UNIVERSITÀ
DEGLI STUDI
DI BRESCIA

DOTTORATO DI RICERCA IN

GENETICA MOLECOLARE, BIOTECNOLOGIE E MEDICINA SPERIMENTALE

BIO/13 BIOLOGIA APPLICATA

CICLO XXXV

TITOLO TESI

Involvement of *NF1* 3' tertile and its interactors in spinal neurofibromatosis type 1 and role of double mutations in NF1 compound heterozygotes

DOTTORANDA:

DR. BETTINAGLIO PAOLA

SUPERVISORE:

PROF.SSA RIVA PAOLA

INDEX

ABSTRACT

Neurofibromatosis Type 1 (NF1) is an autosomal dominant Mendelian disease with variable expression, caused by NF1 gene mutations. In contrast to the classical NF1, the clinical signs of the spinal form (SNF) occur with late onset, often associated with severe pain, due to multiple spinal neurofibromas. To date, there are no known genotype-phenotype correlations in classical and spinal neurofibromatosis and, despite having clearly distinguishable phenotypes, the mechanism that determines one or the other form is not known. This phenotypic variability, together with the heterogeneity of *NF1* mutations, makes it difficult to establish genotype-phenotype correlations. The objective of this study was 1) to verify if the mutational spectrum of NF1 is different in the two forms of neurofibromatosis, 2) evaluate the presence of variants in genes coding for neurofibromin interactors or in genes of the RAS pathway that could constitute, with mutations in the NF1 gene, the genetic basis that distinguishes the two forms and 3) verify the contribution of some variants in the development of a more severe phenotype. The gDNAs of 106 classical and of 74 SNF patients were sequenced by Targeted Resequencing. The mutational analysis of *NF1* confirmed the prevalence of missense mutations in SNF, as already reported in small casuistries, and demonstrated that mutations in the 3' tertile of the NF1 gene, are more frequent in spinal than in classical patients. These results were confirmed in a combined statistical analysis by adding 25 SNF patients, reported in the literature, to our 74 SNF patients. To verify a functional significance of the prevalence of 3' tertile *NF1* mutations in SNF, we searched for rare variants in the interactors of the 3' *NF1* tertile and found six variants in the genes encoding syndecans (*SDC1, SDC2, SDC3* and *SDC4*), in 5 SNF and 1 classical NF1 patient. We investigated the expression of syndecans genes in SNF and NF1 patients, by quantitative real time PCR: *SDC2* and *SDC3* were significantly hyper-expressed in spinal and classical patients compared to a group of controls. Moreover, *SDC2, SDC3* and *SDC4* were significantly hyper-expressed in patients with NF1 mutations in the 3' tertile as compared to controls. Furthermore, we found 5 SNF patients carrying two variants in *NF1*. To verify whether the additional variant could contribute to the phenotype with the main causative *NF1* mutation, we established, where possible, if the two variants were *in cis* or *in trans* and found two compound heterozygotes with a severe SNF form, in two unrelated families. To establish the role of the two *NF1* mutated alleles, we performed expression analysis in the SNF compound heterozygotes proband of family 1, by digital PCR, and on all family members, establishing the comparable expression of both mutated alleles in all the mutations' carriers and the hyper-expression of the mutated alleles as compared to wild-type alleles of healthy controls. Our data confirm the presence of different mutational

spectra characterizing classical and spinal neurofibromatosis. SNF is characterized by *NF1* missense mutations leading to a possible gain-of-function neurofibromin, by *NF1* mutations in the 3' tertile, supporting the role of mutations targeting the C-terminal neurofibromin domains, and by the presence of a "second" *NF1* subclinical variant, possibly contributing to the phenotype. The syndecans hyper-expression in SNF and classical NF1, in patients with 3' tertile *NF1* mutations, suggest their involvement in the activity and correct functionality of the neurofibromin. These results, in addition to providing new knowledge on the genetic basis of the disease, could favor further research aimed at establishing a possible prognostic significance of syndecans and subclinical *NF1* variants, facilitating personalized management of patients.

RIASSUNTO

La neurofibromatosi di tipo 1 (NF1) è una malattia mendeliana autosomica dominante con espressione variabile, causata da mutazioni del gene NF1. Contrariamente alla forma classica, i segni clinici della forma spinale (SNF) si manifestano con esordio tardivo, spesso associato a grave dolore, dovuto a molteplici neurofibromi spinali. Ad oggi non sono note correlazioni genotipo-fenotipo nella neurofibromatosi classica e spinale e, pur avendo fenotipi chiaramente distinguibili, non è noto il meccanismo che determina l'una o l'altra forma. Questa variabilità fenotipica, unita all'eterogeneità delle mutazioni di *NF1*, rende difficile stabilire correlazioni genotipo-fenotipo. L'obiettivo di questo studio è stato 1) verificare se lo spettro mutazionale di *NF1* è diverso nelle due forme di neurofibromatosi, 2) valutare la presenza di varianti nei geni che codificano per gli interattori della neurofibromina o nei geni del pathway di RAS che potrebbero costituire, con mutazioni nel gene *NF1*, la base genetica che distingue le due forme e 3) verificare il contributo di alcune varianti allo sviluppo di un fenotipo più grave. I DNA genomici di 106 pazienti con NF1 classica e di 74 pazienti con SNF sono stati sequenziati mediante Targeted Resequencing. L'analisi mutazionale di *NF1* ha confermato la prevalenza di mutazioni missenso nella SNF, come già riportato in piccole casistiche, e ha dimostrato che le mutazioni nel terzile 3' del gene *NF1*, sono più frequenti nei pazienti spinali rispetto ai classici. Questi risultati sono stati confermati in un'analisi statistica combinata aggiungendo 25 pazienti SNF, già riportati in letteratura, ai nostri 74. Per verificare un significato funzionale della mutazioni nel 3' terzile di *NF1* abbiamo cercato varianti rare negli interattori del 3' terzile e abbiamo trovato sei varianti nei geni che codificano i sindecani (*SDC1, SDC2, SDC3* e *SDC4*), in 5 pazienti SNF e in un paziente classico. Abbiamo poi studiato l'espressione dei sindecani mediante qPCR: *SDC2* e *SDC3* erano significativamente iper-espressi nei pazienti spinali e classici rispetto a un gruppo di controlli. Inoltre, *SDC2, SDC3* ed *SDC4* erano significativamente iper-espressi nei pazienti con mutazioni NF1 nel terzile 3', rispetto ai controlli. Inoltre, abbiamo trovato 5 pazienti spinali portatori di due varianti NF1. Per verificare se la variante aggiuntiva poteva contribuire al fenotipo con la principale mutazione di *NF1*, abbiamo stabilito, dove possibile, se le due varianti erano in cis o in trans e abbiamo trovato due eterozigoti composti con una forma severa di SNF, in due famiglie non imparentate. Per stabilire il ruolo dei due alleli mutati, abbiamo eseguito un'analisi di espressione nel probando eterozigote composto e su tutti i membri della famiglia 1, mediante digital PCR, stabilendo l'espressione comparabile di entrambi gli alleli mutati in tutti i portatori di mutazioni e l'iper-espressione degli alleli mutati rispetto agli

alleli wild-type di soggetti controlli. I nostri dati confermano la presenza di diversi spettri mutazionali caratterizzanti la neurofibromatosi classica e spinale. La SNF è caratterizzata da mutazioni *NF1* missenso che portano a un possibile guadagno di funzione della neurofibromina, da mutazioni nel 3' terzile, a supporto del ruolo di mutazioni che colpiscono i domini C-terminali della neurofibromina, e dalla presenza di una "seconda" variante *NF1* subclinica, che potrebbe contribuire al fenotipo. L'iper-espressione dei sindecani nella neurofibromatosi classica e spinale, supporta il loro coinvolgimento nell'attività e nella corretta funzionalità della neurofibromina. Questi risultati, se confermati in studi funzionali, oltre a fornire nuove conoscenze sulle basi genetiche della malattia, potrebbero favorire ulteriori ricerche volte a stabilire un eventuale significato prognostico dei sindecani e delle varianti subcliniche di *NF1*, facilitando una gestione clinica personalizzata dei pazienti.

1. INTRODUCTION

1.1. Classification of Neurofibromatosis

Neurofibromatosis is a group of genetic disorders with autosomal dominant transmission that predispose to the onset of tumors in the central and peripheral nervous system. There are three forms: neurofibromatosis type 1 (NF1), neurofibromatosis type 2 (NF2) and Schwannomatosis (Figure 1).[1]

The neurofibromatoses

Disorder	Gene	Nonneoplastic clinical phenotype	Cancer phenotypes
NF1	*NF1*	Learning, behavioral, and attention deficits Long-bone dysplasia Pigmentary abnormalities (CALMs, freckling, Lisch nodules)	Optic pathway glioma Brainstem glioma Cutaneous neurofibroma Plexiform neurofibroma Malignant peripheral nerve sheath tumor
NF2	*NF2*	Juvenile cataracts Retinal hamartomas, epiretinal membranes Large CALMs (fewer than in NF1)	Vestibular schwannomas Peripheral and cranial nerve schwannomas Meningioma Spinal tumor glioma
Schwannomatosis	*LZTR1* *SMARCB1*	None	Schwannomas

CALMs, café-au-lait macules; NF1, neurofibromatosis type 1.

Figure 1. The neurofibromatosis. Patrick J. Cimino, Handbook of Clinical Neurology, 148 (51), 2018.

Neurofibromatosis type 1 (NF1; MIM #162200), or von Recklinghausen disease, is the most common form among neurofibromatosis with a prevalence in the population of 1 in 3000-3500 live births. It is an autosomal dominant neurocutaneous genetic disorder with pediatric onset and shows complete penetrance [2] within six years of age.

Neurofibromatosis type 1: history

The earliest examples of type 1 neurofibromatosis probably date back to information found in an Ancient Egyptian parchment, a Hellenistic statuette, and drawings by thirteenth-century Cistercian monks; though the first to clinically describe and observe the hereditary nature of the disease was Mark Akenside in 1768.[3]

The first systematic review of neurofibromatosis was published in 1849 by professor of surgery Robert William Smith, who in his monograph "A Treatise on the Pathology, Diagnosis and

Treatment of Neuroma" describes several cases of patients with "neuromatoustumors" and hypothesizes that the origin of the latter was the connective tissue surrounding small nerves.[4]

Another relevant figure in the history of NF1 was that of the pathologist Rudolf Ludwig Karl Virchow, who described the clinical and neuropathological features of the disease in a series of reports between 1847 and 1863. In the first of these articles Virchow states that fibroids, or tumors of the connective tissue that surround the nerves, should be placed in a separate category: he classified the tumor, today called neurofibroma, as *fibroma molluscum* or *elephantiasis molluscum*. In 1857 he introduced the concept of true and false neuromas based on their histological appearance and according to Virchow a true neuroma consists of nerve fibers and the nerve sheath, while a false neuroma includes only the nerve sheath. In his "Die krankhaftenGeschwulste" published in 1863, Virchow analyzes malignant tumors and includes the study of two cases with multiple neuromas, which fit the description of neurofibromatosis: tumors of many sizes that create lumps in the skin with a characteristic soft texture and familial transmission of the disease. However, Virchow thought that the malignancy was caused by connective tissue metaplasia and this error paved the way for his student, Friedrich Daniel von Recklinghausen, who established the link between connective tissue tumors and nerve tumors.[5, 6]

On January 24, 1879, von Recklinghausen performed a thorough autopsy on Marie Kientz's body and, unlike his predecessors, was the first to recognize that skin pigmentation was a symptom of the disease and to notice inguinal and axillary freckles, which are one of the markers used today in the diagnosis of NF1. At the end of his analysis, he assumed a mixture of connective tissue and nerve fibers in the patient's tumors, which was confirmed by the study of the composition of the tumor samples of the patient Michel Bur.

The results of his studies first appeared in the treatise entitled "Über die multiplenFibrome der Haut und ihreBeziehungzu den multiplenNeuromen" (1882), where von Recklinghausen began to use the term neurofibroma to indicate the simultaneous existence of fibroids and neuromas intimately intertwined with each other. It is based on this article that the eponym "von Recklinghausen disease" was applied to neurofibromatosis type 1.[7]

Diagnostic criteria and clinical manifestations of NF1

To date, the diagnosis of neurofibromatosis type 1 is made following the clinical criteria originally established at the conference of the National Institutes of Health Consensus Development, held in 1987. The diagnosis of NF1 requires the presence of two or more of these clinical signs:[8]

1. Six or more coffee latte spots greater than 5 mm in diameter in pre-pubertal age and greater than 15 mm in post-pubertal age
2. Two or more neurofibromas of any kind or one plexiform neurofibroma
3. Axillary or inguinal freckles
4. Optic glioma
5. Two or more Lischnodules
6. A characteristic bone lesion (dysplasia of the sphenoid wing and dysplasia of the long bones with or without tibial pseudoarthrosis)
7. A first-degree relative with NF1

Neurofibromatosis type 1 is a disorder that occurs in a heterogeneous way and even within the same family there may be different phenotypes. Non-neurological clinical manifestations and symptoms involving the nervous system are recognized. To the first group belong:

→ Pigmental manifestations, including milk coffee spots (Cafe-au-lait macules (CALMs)), axillary and inguinal freckles and Lisch nodules. CALMs are benign hyperpigmented skin lesions that usually appear in the first two years of life, tend to darken with sun exposure and fade with advancing age, while axillary and inguinal freckles appear around 5-8 years. Lisch nodules are benign melanocytic hamartomas of the iris and are present in most adults with NF1.

→ Skeletal deformities, including scoliosis, dysplasia of the long bones and wing of the sphenoid and pseudoarthrosis; in addition, a decrease in bone mineral density has been reported in patients with NF1.[9]

→ Hypertension and other cardiovascular abnormalities, which include renal and cerebral artery stenosis, aortic coarctation and arteriovenous malformations.[10]

→ Increased risk of tumors. The risk of breast cancer is five times greater in women with NF1 under the age of 50 years.[11] In addition, mortality rates in women with neurofibromatosis type 1 and breast cancer are higher than those in women with breast cancer in the general population. Children with neurofibromatosis type 1 have at least a seven-fold greater risk of developing myeloid leukemia than children in the general population. Other tumors not associated with the nervous system are: pheochromocytoma, gastrointestinal carcinoid tumors, malignant melanoma, rhabdomyosarcoma, and gastrointestinal stromal tumors (GISTs).[12,13,14,15,16]

Among the clinical manifestations associated with the nervous system, we find the following:

→ Glioblastoma.

→ Optic pathway glioma (OPG) is the most common brain tumor, observed in 15-20% of children with NF1 and usually appears before the age of seven. Instead, in the first decade of life, gliomas of the brainstem may appear, which are usually pilocytic astrocytomas, like most gliomas of the optic pathways.[17]

→ Neurocognitive deficits and learning difficulties.

→ Epilepsy is present in 4-13% of people with NF1.[18]

→ Neurofibromas are benign tumors of the sheaths of peripheral nerves, composed of neoplastic Schwann cells, fibroblasts, blood vessels, and mast cells. They can develop into different positions and can be divided into:[19]

- Dermic
- subcutaneous
- spinal
- plexiform

→ Malignant peripheral nerve sheath tumours (MPNST) are soft tissue sarcomas and occur at various anatomical sites; over a lifetime, individuals with NF1 have an 8-13% risk of developing MPNST and often these neoplasms develop from plexiform neurofibromas.[20]

1.2 The different forms of NF1

In the field of neurofibromatosis type 1, three peculiar conditions are recognized, caused by mutations in the *NF1* gene, but which differ from the classical form for their lower variability of symptomatology:

- microdeletion syndrome, present in 4-11% of NF1 patients, is caused by a deletion of *NF1* and other genes in the 17q11.2 region and is associated with a more severe phenotype than classical neurofibromatosis;
- segmental (or mosaic) neurofibromatosis, is characterized by the same clinical manifestations and genetic alterations as classical NF1; however, these are limited to a certain body region. This condition is determined by the presence of a somatic mosaicism originating from a post-zygotic mutation in the *NF1*gene;

- Spinal neurofibromatosis is characterized by the presence of CALMs and neurofibromas that develop along all the roots of the spinal nerves.

Differences between Classical and Spinal NF1

Spinal neurofibromatosis (SNF) is an alternative form of NF1, characterized by bilateral neurofibromas (histologically analyzed) at the level of all spinal roots, with or without other manifestations typical of classical NF1.

The first to introduce the concept of SNF were Norman E. Leeds and Harold G. Jacobson in 1976. Subsequently, this form of NF1 was reported in the literature under the names "familial spinal neurofibromatosis" and "hereditary spinal neurofibromatosis", respectively, by Pulst et al. in 1991 and by Poyhonen et al. [21,22,23] in 1997. For the first time, both studies observed that, regardless of the complete or partial involvement of the spinal roots, the phenotypes of the affected individuals were:

- other than NF1 and NF2;
- distinct, as the predominant feature is multiple spinal neurofibromas;
- more limited in their manifestation with the absence of Lisch nodules and non-tumour characteristics typical of classical NF1

Further studies recorded cases, both familial and sporadic, of spinal neurofibromatosis, observed a phenotypic variability that included features typically associated with classical neurofibromatosis, particularly Ruggeri et al. noted that:[24]

- CALMs, in the context of SNF, were lighter in color, larger in size and with less regular margins than typical spots in NF1 patients and were somewhat closer in size, shape, and color to the spots encountered in NF2;
- Cutaneous neurofibromas and Lisch nodules in SNF patients occur in a lower percentage than in classical NF1 patients;
- Neuropsychiatric manifestations have rarely been reported in SNF, while in NF1 patients they are usually greater than 50%.

The main feature of individuals suffering from spinal neurofibromatosis is bilateral neurofibromas on all nerve roots. The involvement of all the roots of the spinal nerves clearly defines the SNF phenotype, distinguishing it from the classical one, which can present spinal neurofibromas, but only in certain segments of the spine. Nerve compression at the spinal level, due to the increase in neurofibromas size, worsens the quality of life of patients, who often do

not present symptoms during childhood, making it difficult to make an early diagnosis. Spinal tumours can remain asymptomatic for years; however, once severe neurological deficits and myelopathy have developed, the likely success of any surgical treatment is greatly reduced.[24]

Some patients, when coming to clinical observation, have spinal neurofibromas that do not involve all nerve roots; for this reason, Ruggeri et al. created the term MNFSR (multiple not few spinal roots), which identifies individuals with spinal neurofibromas in multiple but not all spinal roots. [24]

Unlike classical neurofibromatosis, the absence of specific clinical symptoms makes it difficult to diagnose spinal neurofibromatosis, which is confirmed only by magnetic resonance imaging.

1.3 Rasopaties and RAS pathway

Rasopathies are a clinically defined group of syndromes caused by a deregulation of the RAS pathway due to the presence of mutations in genes encoding components or regulators of the Ras/mitogen-activated protein kinase (MAPK) pathway. This pathway is activated in response to a wide range of extracellular stimuli (e.g. growth factors, hormones, cell/cell interactions) and plays a key role in the regulation of the cell cycle, growth, differentiation, senescence and cellular apoptosis. Each Rasopathy shows a unique phenotype, but they also share many clinical manifestations, since mutations are present in genes that participate in the same transduction pathway[25] (figure 2).

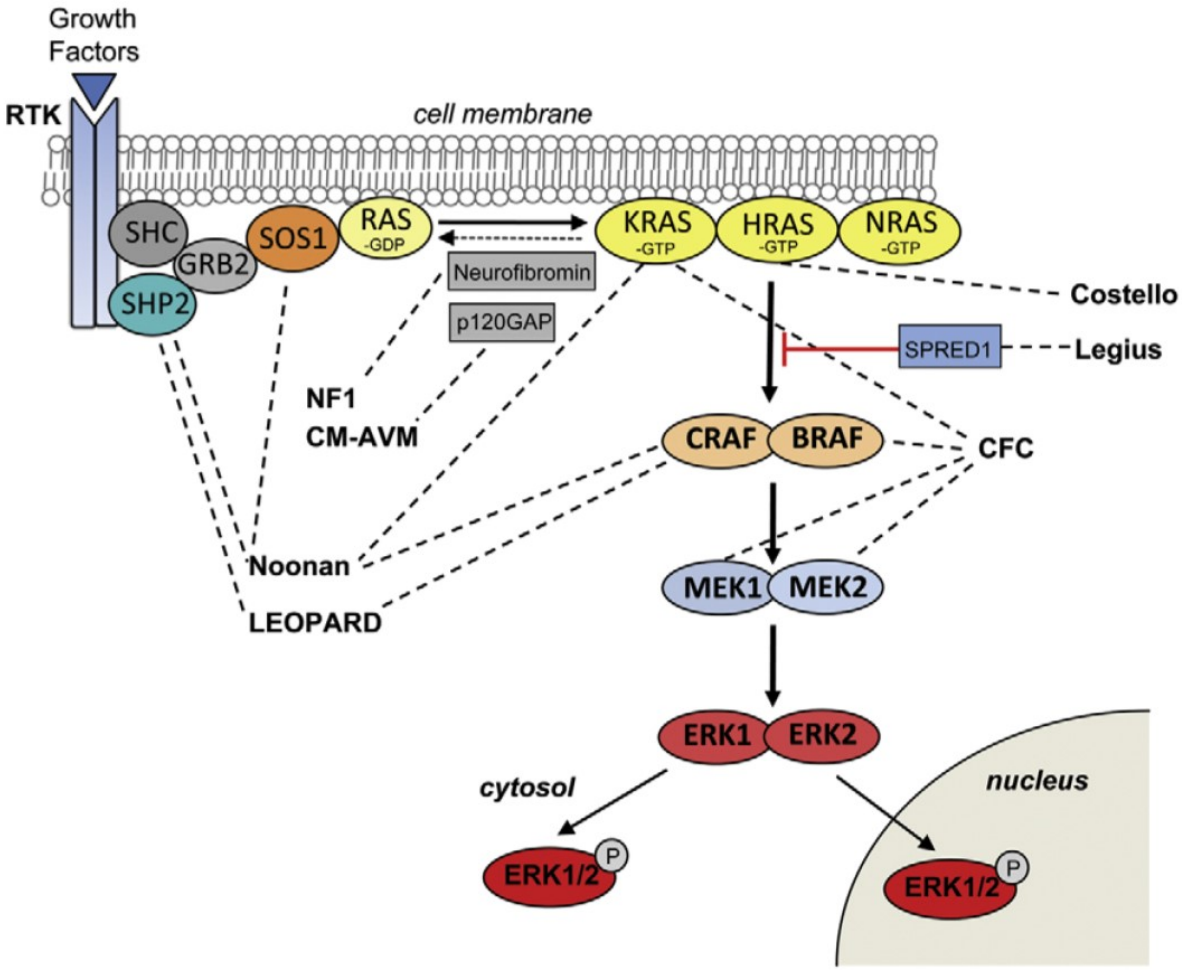


Figure 2. RAS pathway and associated developmental syndromes (indicated by dashed lines).Tidyman, W. E. &Rauen, K. A. RASopathies: Developmental syndromes of Ras/MAPK pathway dysregulation. Current Opinion in Genetics and Development (2009)

These disorders are Neurofibromatosis Type 1 (NF1), Noonan Syndrome (NS), Noonan Syndrome with Multiple Freckles (NSML), Capillary-Arteriovenous Malformation Syndrome (CM-AVM), Costello Syndrome (CS), Cardio-Faciocutaneous Syndrome (CFC) and Legius Syndrome (figure 3).

Syndrome	Gene	Chromosome location	Protein function	Clinical phenotype
Autosomal dominant intellectual disability, type 5	*SYNGAP1*	6p21.3	RasGAP	Typically nondysmorphic to mild dysmorphic craniofacial features, moderate to severe intellectual disability, global developmental delay with behavioral issues, autism spectrum disorder, ophthalmologic findings, hypotonia, seizures.
Capillary malformation-AV malformation	*RASA1*	5q14.3	RasGAP	Nondysmorphic craniofacial features, multifocal capillary malformations which may be associated with arteriovenous malformations and fistulae.
Cardio-facio-cutaneous	*BRAF*	7q34	Kinase	Dysmorphic craniofacial features, congenital heart defects, failure to thrive with short stature, ophthalmologic abnormalities, multiple skin manifestations including progressive formation of nevi; variable neurocognitive delay; hypotonia, may be predisposed to cancer
	MAP2K1	15q22.31	Kinase	
	MAP2K2	19p13.3	Kinase	
	KRAS	12p12.1	GTPase	
Costello	*HRAS*	11p15.5	GTPase	Dysmorphic craniofacial features, congenital heart defects, failure to thrive with short stature, ophthalmologic abnormalities, multiple skin manifestations including papilloma; variable neurocognitive delay; hypotonia; predisposition to cancer.
Legius	*SPRED1*	15q14	Negative Regulator	Café-au-lait maculae, intertriginous freckling, normal to mild neurocognitive impairment, macrocephaly; unclear predisposition to cancer.
Noonan	*PTPN11*	12q24.1	Phosphatase	Craniofacial dysmorphic features, congenital heart defects, short stature, undescended testicles, ophthalmologic abnormalities, bleeding disorders, normal to mild neurocognitive delay; predisposition to cancer.
	SOS1	2p22.1	RasGEF	
	RAF1	3p25.1	Kinase	
	KRAS	12p12.1	GTPase	
	NRAS	1p13.2	GTPase	
	SHOC2	10q25.2	Scaffolding	
	CBL	11q23.3	Ubiquitin ligase	
	RRAS	19q13.33	GTPase	
	RIT1	1q22	GTPase	
	RASA2	3q23	RasGAP	
	SOS2	14q21.3	RasGEF	
	MAP3K8	10p11.23	Kinase	
	SPRY1	4q28.1	Inhibitor	
	MYST4	10q22.2	Acetyltransferase	
	LZTR1	22q11.21	Adaptor	
	A2ML1	12p13.31	Protease inhibitor	
Noonan with multiple lentigines	*PTPN11*	12q24.1	Phosphatase	Same as Noonan syndrome but may develop multiple skin lentigines as individuals gets older; unclear predisposition to cancer.
	RAF1	3p25.1	Kinase	
Neurofibromatosis 1	*NF1*	17q11.2	RasGAP	Café-au-lait maculae, intertriginous freckling, neurofibromas and plexiform neurofibromas, iris Lisch nodules, osseous dysplasia, optic pathway glioma, normal to mild neurocognitive delay; predisposition to cancer

Figure 3. RASopathies characteristics. Tidyman, W. E. &Rauen, K. A. Pathogenetics of the RASopathies. Human Molecular Genetics (2016).

The *RAS* gene belongs to a multigenic family that includes *HRAS*, *NRAS* and *KRAS*. Ras activation is a complex and tightly regulated mechanism, which will be presented here according to a canonical, but not exclusive, model in the case of tyrosine kinase receptors (figure 2). Ras proteins are small GTPases activated because of the binding of growth factors and tyrosine kinase receptors (RTKs), which dimerize and self-phosphorylate on different tyrosine residues, thus creating binding sites for the intracellular protein Grb2 (Growth Factor Receptor-Bound Protein 2). This allows to recruit at the plasma membrane level the SOS1 protein (Son of Sevenless 1) which induces through its GEF domain (Guanine Exchange Factor) the activation of Ras, which

passes from an inactive conformation, Ras-GDP, to an active one, Ras-GTP. Ras-GTP activates a series of downstream kinases: the first is RAF, which phosphorylates and activates MEK 1/2 kinases, which in turn activate ERK 1/2 (Extracellular signal-Regulated Kinases). ERK 1/2 is a protein that performs multiple functions, at the level of the phosphorylation of different nuclear transcription factors that induce cell proliferation, while at the cytoplasmic level it phosphorylates membrane proteins and protein kinases, promoting differentiation.[26,27]

Given the role played by this pathway in cell proliferation and survival, and given the high number of proteins involved, a *de novo* or inherited mutation may increase the risk of developing tumours or serious pathological phenotypes. Since Ras is a protooncogene, if a mutation involve, for example, the constitutive activation of Ras-GTP, it would result in a proliferation and uncontrolled growth of the cell, typical of cancer cells, such as the activation of the final MEK effector through mutations in genes downstream of Ras.

1.4 NF1 gene and neurofibromin

The *NF1* gene (figure 4) is in the 17q11.2 region, has a length of 350 kb and contains 60 exons. Its transcript measures 11-15 kb with an open-reading frame of 8454 bp and a 3' UTR of 3.5 kb. An interesting aspect of this gene is the presence at the level of intron 35 of three genes, transcribed on the opposite strand: *EVI2A* (ectotrophic viral integration site), *EVI2B*, and *OMG* (oligodendrocyte myelin glycoprotein). None of these genes are mutated in NF1 patients. Given the presence of tumors in patients with neurofibromatosis type 1 and the identification of somatic mutations of *NF1* in sporadic tumors independent of the disease, *NF1* has been identified as a tumor suppressor.[28]

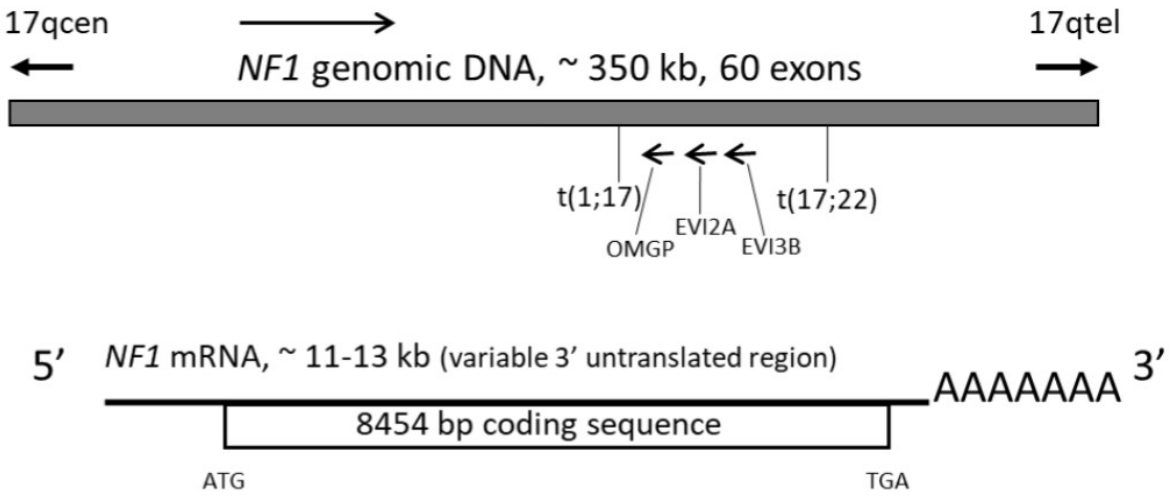


Figure 4. Schematic representation of the NF1 gene and its transcribed mRNA. Bergoug, M. et al. Neurofibromin Structure, Functions, and Regulation. Cells (2020). 17qtel=17q thelomer, 17qcen=17q centroomer

The NF1 gene encodes a protein called neurofibromin,[29] which has a molecular weight of 250-280 kDa and is expressed ubiquitously but is present in higher concentrations within the nervous system, particularly in nonmyelinated neurons, Schwann cells and oligodendrocytes, as well as in the adrenal medulla, leukocytes, and testes.[30] There are several tissue-specific isoforms, resulting from the alternative splicing of pre-mRNA, but the most expressed form is a neurofibromin of 2818 amino acids, translated of an mRNA containing 57 exons. The most studied alternative splicings, which allow the expression of different *NF1* isoforms, involve the exons: 9a, 10a-2, 23a, 43 and 48a (figure 5). Exon 23a is within the GRD (GAP-related domain) domain and its alternative splicing generates two transcripts: one encoding the neurofibromin I (or type 1) isoform consisting of 2818 aa and mostly expressed in neurons; the other coding for the neurofibromin II isoform (or type 2) formed by 2839 aa and predominant in glial cells. Neurofibromin type 3, resulting from the alternative splicing of exon 48a, is expressed exclusively in cardiac and muscle cells; Neurofibromin type 4 contains both exon 23a and 48a.The delta E43 NF1 isoform, in which exon 43 is deleted, is highly expressed in lungs, liver, placenta, kidneys, and skeletal muscle, and poorly expressed in neurons compared to the isoform containing exon 43; exon 43 contains the signal of nuclear localization and compared to other tissues, suggesting the important nuclear function of neurofibromin in the nervous system.

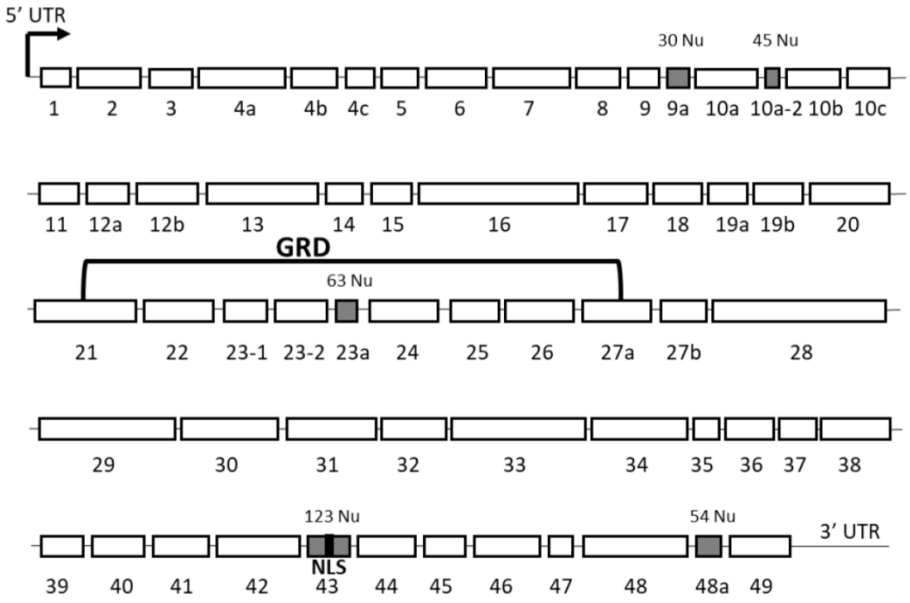


Figure 5. Schematic representation of the full-length transcript of *NF1*. The GAP-related domain (GRD) is indicated; in gray are highlighted the exons that undergo alternative splicing with the respective number of nucleotides (Nu) and the nuclear localization signal (NLS) is in black. Bergoug, M. et al. Neurofibromin Structure, Functions and Regulation. Cells (2020)

Neurofibromin has several domains (figure 6 and 8):

- GRD (GAP-related domain) promotes hydrolysis from Ras-GTP to Ras-GDP. Upstream of this domain there is a region, called TBD (tubulin-binding domain), which allows the interaction between neurofibromin and cytoskeletal tubulins, resulting in a decrease in the GAP activity of neurofibromin.
- CSDR (cysteine-serine-rich domain) in the N-terminal position is phosphorylated by both protein kinase A (PKA) and protein kinase C (PKC); its PKC-dependent phosphorylation increases the Ras-GAP activity of neurofibromin.
- Sec14-PH is a lipid-binding domain.
- CTD (C-terminal domain) plays an important role in regulating the transition from metaphase to anaphase during the cell cycle and contains a nuclear localization signal (NLS), phosphorylation on Serine 2808 by PKC-ε is important for the nuclear transfer of neurofibromin. In addition, this domain interacts with different molecules, such as CRMP2 (collapsin response mediator protein 2), FAK (focal adhesion kinase), CASK (calcium/calmodulin-dependent serine protein kinase) [31,32] and syndecanes, to which a domain called SBR (syndecan-binding region, also called SBD, Figure 8) has been reserved.

- HLR (HEAT-like repeat) is a domain involved in protein-protein interactions and often forms a complex solenoid structure that incorporates a series of α-helices.

Some studies, aimed at assessing the risk of OPG in patients suffering from neurofibromatosis, present a subdivision of the NF1 gene into three homogeneous regions: the 5' tertile, which includes exons from 1 to 21, corresponds in part to the CSRD domain of neurofibromin, the intermediate tertile, includes exons from 22 to 38, coding for TBD domains, GRD and SecPH of the protein, and finally the tertile 3' contains exons from 39 to 57, which at the protein level will form the HLR, NLS and SBR domains.[33]

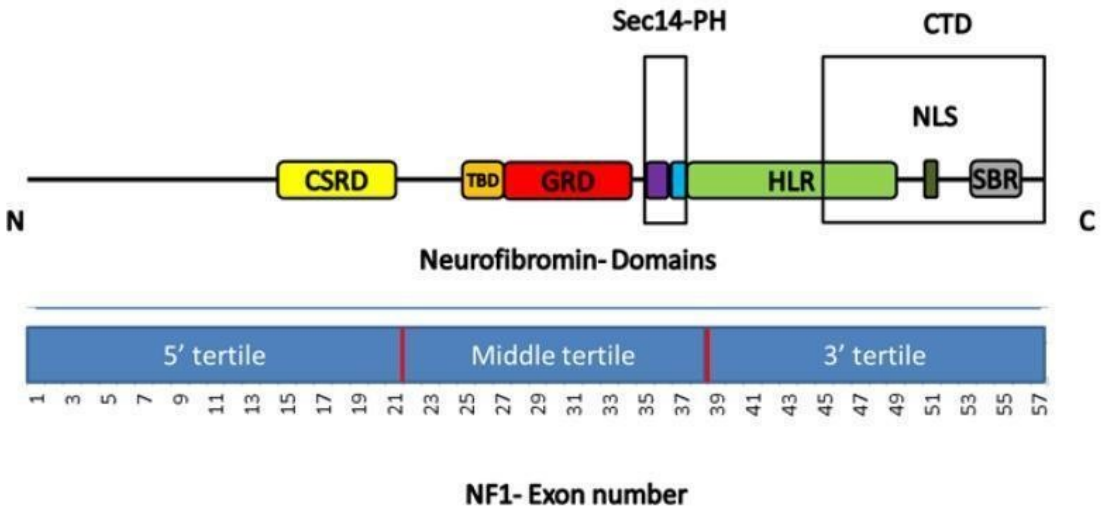


Figure 6. The NF1 gene subdivision in tertiles and representation of the neurofibromin domains. Melloni, G. et al. Risk of optic pathway glioma in neurofibromatosis type 1: No evidence of genotype–phenotype correlations in a large independent cohort. Cancers (2019).

Neurofibromin is present in both the nucleus and cytoplasm of cells and performs various functions:

- Due to the interaction with Spred1 or the binding with Cav-1, neurofibromin translocates into the plasma membrane, where it negatively regulates the Ras pathway, promoting the intrinsic GTPasic activity of Ras, which passes from an active form (Ras-GTP) to an inactive form (Ras-GDP). Consequently, the decreased or absent expression of neurofibromin leads to the constitutive activation of the Ras pathway with negative consequences for the cells;
- is a positive regulator of cAMP levels by activating two distinct pathways involving adenylate cyclase;

- is a positive regulator of dopamine homeostasis, in particular a dose-dependent relationship between neurofibromin levels, dopamine signaling pathway and cognitive defects present in NF1 patients has been identified, although the molecular mechanism is still unknown[34];
- inactivates the Ras/PI3K pathway (figure 7), which regulates upstream the activation of mTOR, a protein kinase that drives cell proliferation and survival. In the absence of neurofibromin, Ras hyperactivation increases the action that PI3K performs on PIP2, which, converted to PIP3, phosphorylates PDK1 and activates the transductional cascade mediated by Akt. Finally, an increase in mTOR functions is obtained with the possible development of neoplasms;

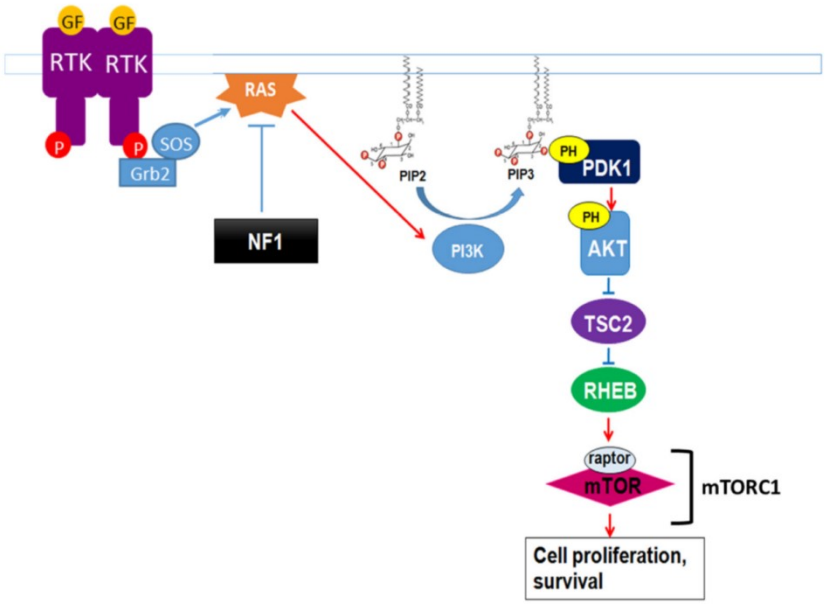


Figure 7. Role of neurofibromin in the RAS / PI3K pathway. Bergoug, M. et al. Neurofibromin Structure, Functions and Regulation. Cells (2020)

- controls the organization and turnover of actin cytoskeletal filaments by negative regulation of two pathways;
- interacts with proteins involved in microtubular transport in melanocytes, neurons, and Schwann cells, for example, neurofibromin has been found in complex with kinesin 1, a protein involved in anterograde transport along microtubules[35];
- De Sheppers et al. [36] demonstrated a direct interaction between the neurofibromin GRD domain and the APP protein (amyloid precursor protein) in melanosomes (Figure 8)

- regulates the transition from metaphase to anaphase during the cell cycle, contributes to the formation of the mitotic spindle, and the correct alignment of chromosomes on the metaphasic plate. The absence of neurofibromin causes chromosomal instability and aneuploidy[32];

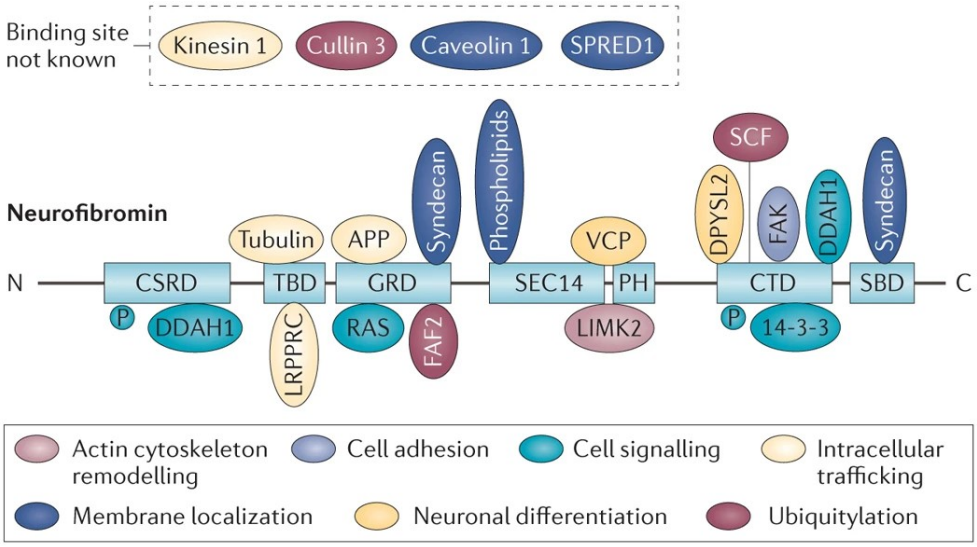


Figure 8. Neurofibromin domains and its interactors. Ratner, N. & Miller, S. J. A RASopathy gene commonly mutated in cancer: the neurofibromatosis type 1 tumour suppressor. *Nature Reviews Cancer* (2015).

1.5 The family of syndecans and its interaction with neurofibromin

Syndecans are a family of transmembrane heparan sulfate proteoglycans (HSPGs) involved in numerous functions, including the binding of extracellular matrix proteins (such as, for example, laminin and fibronectin), adhesion to the cell matrix, cell movement, and tissue morphogenesis. In mammals there are 4 of them, each encoded by a different gene [37]:

- Syndecan 1 (*SDC1*): it is expressed very early in development and is found in epithelial and mesenchymatic cells, associated with tissue morphogenesis. It is significant for cell-to-cell and cell-to-matrix interactions.
- Syndecan 2 (*SDC2*): Formerly called fibroglycan, it is found particularly in mesenchymal tissue, neurons, and the liver.
- Syndecan 3 (*SDC3*): it is widely present at the neural level, but it is also expressed in some musculoskeletal tissues.
- Syndecan 4 (*SDC4*): it is present in many different cell types, including glial cells.

All members of the syndecan family have common structural features: the transmembrane (TM) and cytoplasmic (CM) domains are very similar (60-80% of amino acid identity) in all syndecans, while the extracellular domain (ED) is specific to each syndecan, sharing only a minimal homology [38] .

Both the transmembrane and cytoplasmic domains are highly conserved. Of the second, the 2 most preserved regions are the following [39]:

1. EFYA (C-terminal)
2. RMRKKDEGSY (membrane-proximal segment): is the segment involved in the interaction with neurofibromin, more specifically, the RKKD sequence.
3. DLGERK and YQKAPT sequences have also been shown to contribute to interaction with neurofibromin, however, they are considered less important than RKKD.

Following several studies carried out by Yi-Ping Hsueh et. al in 2001 [39], it was observed that the transmembrane and cytoplasmic domains of syndecans are necessary to bind neurofibromin. The latter is widely distributed in neurons, both in axons, in dendrites and in cell bodies, where it is possible that it interacts and binds to all 4 syndecans, through part of the GRD domain (amino acids from 1356 to 1564) and the entire SBR domain (amino acids from 2616 to 2719), localized in the 3' tertile.

All four syndecans also interact in turn with CASK, a membrane-associated guanylate cyclase expressed in the brain in the embryonic and postnatal phase, determining, following the link with neurofibromin, the formation of the neurofibromin-syndecan-CASK ternary protein complex, whose role has not yet been identified. However, the colocalization of neurofibromin near the pre or post-synaptic membrane could promote the GTPase activity of RAS or give NF1 new features not currently known (Figure 9).

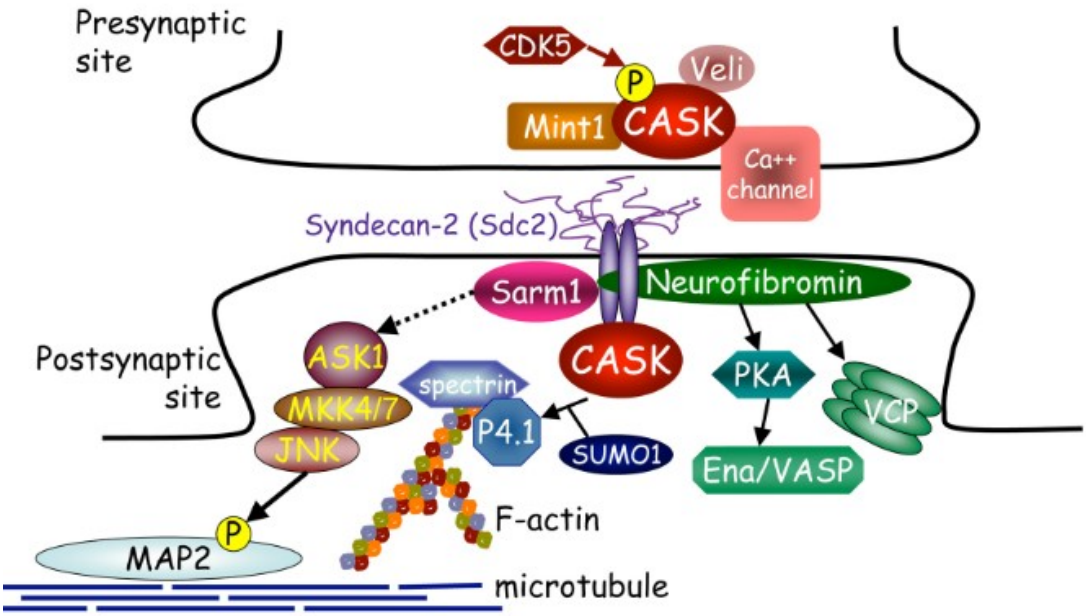


Figure 9. Neurofibromin-syndecan-CASK ternary protein complex near the pre or post-synaptic membrane (http://www.imb.sinica.edu.tw/~hsueh)

1.6 Mutations in the NF1 gene in classical and spinal neurofibromatosis

Molecular analysis of the NF1 genefor mutations is quite complicated for several reasons:

- the large size of the gene;
- the presence, throughout the human genome, of pseudogenes that have a sequence homology with the *NF1* gene [40]
- the lack of well-defined mutational hot spots, since mutations arise rather randomly throughout the gene;
- the wide allelic heterogeneity, with wide variability in the type and size of mutations;
- the absence of clear genotype-phenotype correlations, apart from some exceptions currently recognized.

To date, 7963 variants have been reported in the LOVD database (Leiden Open Variation Database [41]; data updated to October 2022), of which 36% are missense, 23% are frameshift, 18% are stop gain, and 9% silent mutations.

Despite numerous studie s aimed at identifying a possible correlation between the specific mutation in the NF1geneand the associated clinical picture, few genotype-phenotype correlations have currently been identified. In the case of microdeletions, involving the entire NF1 geneand flanking sequences, the presence of a more severe clinical picture has been observed. This condition, known as microdeletion syndrome, is found in approximately 5% of patients with NF1

and is characterized by an increased frequency of craniofacial dysmorphism, intellectual deficit, early onset of a high number of cutaneous and plexiform neurofibromas and increased risk of developing MPNST later in life. Also, in the group of patients with microdeletion syndrome there is a phenotypic variability that could be related to the different extent of the deletion and the possible contribution of contiguous genes.[42]

The second genotype-phenotype correlation described is the in-frame deletion of 3 bp in exon 17 of the NF1 gene (c.2970_2972delAAT) resulting in the loss of one or two methionine residues (Met991) in a highly conserved region of neurofibromin. The mutation is normally associated with a milder clinical phenotype, characterized by the absence of cutaneous, subcutaneous, and external plexiform neurofibromas.[43]

In 2015 Pinna et al. reported a missense c.5425C>T substitution of c.5425C> T*(p.Arg1809Cys)* in exon 29 of the NF1 gene associated with a mild phenotype, characterized by limited involvement of the skin district (coffee-milk spots and freckles), absence of neurofibromas, and lower incidence of Lisch nodules compared to the general population of NF1. The biological mechanism underlying the absence of neurofibromas in patients carrying these mutations is not currently known. [44]

Similarly, the missense variant *p.Arg1038Gly* has been associated with a mild phenotype characterized by CALMs, without neurofibromas or other complications typical of NF1, and characteristics similar to Noonan syndrome have been observed in many patients.[45]

Missense mutations involving codons from 844 to 848 localized in the CSDR domain, are related to a more severe phenotype with an increased prevalence of plexiform and/or spinal neurofibromas, a high risk of symptomatic or asymptomatic OPGs, malignant neoplasms, and skeletal deformities. In particular, individuals affected by the c.2542G>C missense mutation *(p.Gly848Arg)* have a small number of symptomatic spinal neurofibromas and a clear decrease in CALMs, freckles and cutaneous neurofibromas[46]; however, functional studies in mice with this mutation do not recapitulate the phenotype found in these patients. Recently, Koczkowska et al. have identified variants affecting three specific aminoacid of NF1 *(p.Met1149, p.Arg1276, p.Lys1423)* with a possible genotype-phenotype correlation. Patients with the *p.Met1149* mutation have a mild phenotype with several CALMs and freckles, without symptomatic plexiform or spinal neurofibromas. Individuals positive for the *p.Arg1276* mutation show a lower prevalence of cutaneous neurofibromas and an increased incidence of spinal neurofibromas;

while the variant *p.Lys1423* predisposes to a greater development of plexiform neurofibromas. Additionally, the latter two mutations are associated with cardiovascular abnormalities, including pulmonary stenosis.[47]

Despite the limited genotype-phenotype correlations, molecular analysis of the NF1 gene plays an important role in the diagnosis of this condition at an early age (before the clinical signs necessary for confirmation manifest themselves clinically, particularly in sporadic forms) and in doubtful clinical cases. The identification of the causative mutation also allows to extend the research to family members at risk and allows the early prenatal diagnosis of this condition.

Spinal neurofibromatosis is also characterized by high intrafamilial phenotypic variability: SNF patients may belong to families where other family members have a diagnosis of classical NF1 or MNFSR.

From a genetic-molecular point of view, the causes that lead to the different manifestation of classical NF1 compared to SNF have not been identified: At present only 98 patients with complete or partial SNF have been described, of which 90% are carriers of a mutation in the *NF1* gene and with a significant increase in missense mutations in the same gene in patients with SNF compared to those with classical NF1, as reported by Ruggieri et al.[24] in a restricted series of 49 patients.

The gaps in the knowledge of the etiopathogenesis of this disease concern the identification of genes and / or pathogenetic mechanisms that, in addition to the mutation in the NF1 gene, lead to the worsening of the phenotype, prognosis and management of the patient. It is possible that in addition to the NF*1* gene, other components of the Ras/MAPK pathway are also involved, thus contributing to the alteration of the phosphorylation of the ERK 1/2 effector and consequently also of the signal transduction in specific spatio-temporal conditions. The neurofibromin resulting from the *NF1* mutations observed in patients, seems to maintain the function of its domains, this could suggest that neurofibromin interactors, even if not belonging to the Ras/MAPK pathway, may also play a role in modulating the effect of *NF1* mutations.

The presence of variants in these interactors could result in a loss or gain of functions, leading, together with mutations in *NF1*, to modulation or determination of the spinal phenotype.

Despite several years of studies, today it is still difficult to identify predictors of disease severity. Other mechanisms and strategies should be hypothesized to provide clinicians with adequate tools for effective diagnosis, prognosis, and genetic counselling.

1.7 Role of double *NF1* mutations in the NF1

99% of the cases genetically analyzed, present the mutation in one of the two copies of the *NF1* gene, consistently with studies conducted on animal models, where double KO *NF1* are lethal. Nevertheless, in the literature there is only one case described in which both the copies are affected. Fauth et al. [48] described a sporadic SNF patient with double *in trans NF1* mutations: the missense c3046T>C (*p.Cys1016Arg*) in exon 18 and a 3-bp deletion c8131-8133delGTT (*p.2711delVal*) in exon 48 of *NF1*. The authors assumed that the *p.2711delVal* is a most likely benign unclassified variant and the *p.Cys1016Arg* represents the pathogenic mutation responsible for a severe phenotype of patient, characterized by a mild dermal feature, paraparesis, spinal neurofibromas and MPNSTs at spinal level, confirming that the presence of a second subclinical variant in the other copy of *NF1* is consistent with the lethal condition of loss of function mutations in both *NF1* copies. Other two papers reported familial NF1 cases in which two mutations *in cis* have been described, therefore one of the two *NF1* copies remains functional. E. Hernández-Imaz[49] reported a case of a large NF1 family with two truncating mutations in exon 37 of *NF1*, the recurrent c.6792C>A and the novel c.6799C>T change, that occur in *cis* and segregate with *NF1*. The double mutation induces defective splicing of exon 37 and by expression analyses the authors demonstrated the skipping of exon 37 was greater and there were fewer mutant full-length transcripts in samples with the double mutation than in those carrying single mutations. Thus, the combination of the c.6792C>A and c.6799C>T mutations augmented exon 37 skipping. Terzi et al.[50] described two pathogenic *NF1* gene mutations identified in DNA from a child with mild phenotype inherited one from the NF1 father and the second absent from the parents (de novo).

1.8 Evolution of *NF1* mutational analysis techniques up to Targeted NGS

The *NF1* gene was mapped to the long arm of chromosome 17 via linkage studies with the Nerve Growth Factor Receptor (NGFR) gene and located at position 17q22. Confirmation of the exact location was obtained by studying two chromosomal abnormalities found in two different patients with neurofibromatosis type 1, namely the two independent translocations t(1;17)(p34.3;q11.2) and t(17;22)(q11.2;q11.2). The translocated chromosomes had the breaking point at chromosome 17 in the 17q11.2 region where the NF1 gene had been mapped, which was defined as the causative gene of the disease.

To date, the most used method for genomic analysis is Next Generation Sequencing, which based on the size of the portion of the genome analyzed, is divided into:

- Whole Exome Sequencing (WES)
- Whole Genome Sequencing (WGS)
- Targeted NGS testing (Panel NGS testing)

The analysis of highly heterogeneous pathologies, of which most of the causative genes are known, can be carried out by means of "pre-sequencing" panels (Targeted Resequencing), which allow the enrichment of specific genomic regions before sequencing so as to simultaneously analyze a certain number of patients and reduce costs, or through *"in silico"* panels (Target Data Analysis), which consist of a WES followed by the analysis of only those genes already know nto be associated with the pathology under consideration. These high-throughput systems identify a considerable number of variants, the interpretation of which must be standardized; for this reason, in 2013, the ACGM (American College of Medical Genetics and Genomics) convened a group of experts to define new criteria for the classification of variants, which were divided into:

- Pathogenic (causative; class 5)
- Likely pathogenic (class 4)
- Of uncertain clinical significance ("VUS"; class 3)
- Likely benign (class 2)
- Benign (class 1)

The probability threshold for defining a probably pathogenic and probably benign variant has been arbitrarily established (e.g. 90% for ACMGG-AMP), however new scientific evidence could change this classification.

2. PURPOSE

Neurofibromatosis type 1 is a genetic disorder with autosomal dominant transmission, which predisposes to the development of neoplasms at the level of the central and peripheral nervous system. There are several forms of neurofibromatosis: classical, spinal (SNF) and microdeletionsyndrome; the latter is caused by a deletion of *NF1* and other genes in the 17q11.2 regionand is the only form associated with a more severe phenotype. To date, there are no known genotype-phenotype correlations in classical and spinal neurofibromatosis and, despite having more clearly distinguishable phenotypes, the mechanism that determines one or the other form is not known; in fact, some patients show a severe phenotype, while others, with the same mutation, develop milder symptoms. This phenotypic variability, together with the enormous heterogeneity of pathogenicmutationsin the NF1 gene, make it difficult to establish genotype-phenotype correlations. We hypothesized that the variable expressiveness of the phenotype between the spinal and classical forms, but also at the level of each form, may be associated with the co-presence of one or more variants, as well as in the *NF1* gene, also in genes involved in the RAS pathway and interactors of neurofibromin. During my internship, supervised by Prof. Riva's group and in collaboration with the C. Besta Neurological Institute, the IRCCS Ca' Granda Ospedale Maggiore Policlinico Foundation and the Council of National Research in Milan, the genomic DNA of a group of patients suffering from classical NF1 and one of patients with SNF, with the aim of:

- to verify if the mutational spectrum of *NF1* is different in the two forms of neurofibromatosis, classical and spinal NF1;
- verify the presence of variants in genes coding for neurofibromin interactors or in genes of the RAS pathway that could explain, together with mutations in the *NF1* gene, the genetic basis that distinguish the two forms;
- verify the presence of variants associated with a more severe phenotype.

The results of this study, in addition to providing new knowledge on the genetic basis of the disease, could promote further research aimed at studying variants in these genes, which could take on a prognostic significance for one or the other form of neurofibromatosis favoring a personalized clinical management of patients.

3. MATERIALS AND METHODS

3.1 Study subjects and samples collection

A total of 74 individuals with diagnosis of SNF according to Ruggieri's[24] criteria were identified by means of "Fondazione IRCCS Istituto Neurologico Carlo Besta", "Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milano", and "Azienda Ospedaliera Universitaria dell'Università degli studi della Campania Luigi Vanvitelli" electronic databases. 41 had sporadic NF1, while 33 had a familiar form of the disease. All other affected members that were alive and viable were clinically evaluated, and a cohort of 28 affected relatives (12 MNSFR, 7 SNF, and 9 classical NF1) was added to the original patient set, making it possible to identify 19 NF1 families with at least one member affected by SNF. Two families had four affected members, five three affected relatives and twelve families two.

To compare phenotype and genotype an additional cohort of 106 classical NF1 patients without spinal tumours were included in the study. The oldest patients in our database were chosen because the probability of developing spinal NF was very low.

All medical records were surveyed. Data were collected at the time of mutation analysis and reverified for accuracy at the time of this study. We focus on symptoms and signs possible related to spinal neurofibromas, such as pain, neurological symptoms including weakness, sensory deficit, and changes in tone and reflexes, in addition to the already well-known NF1 characteristics: CALm, skinfold freckling, cutaneous, subcutaneous and plexiform neurofibromas, Lisch nodules visual impairment, pain, epilepsy, cognitive impairment, optic nerve glioma (OPG), other neoplasms of central nervous system and of other organs, skeletal and vascular abnormalities. Data of birth, sex, age at the time of last visit, and mode of inheritance were also recorded. All patients underwent gadolinium-powered brain and spinal MRI.

Cases with missing data for a particular sign and/or symptom were classified as 'unknown' and therefore excluded from this part of the genotype-phenotype analysis.

This study was approved by the Fondazione IRCCS Istituto Neurologico Carlo Besta Ethics Committee and Scientific Board (N°50- 19/3/2018).

The series of patients was selected based on the presence of distinctive clinical signs of the classic or spinal form, as reported in Table 1, and then subjected to NGS-targeted resequencing

Table 1. *Clinical signs used for the diagnosis of classical NF1 and SNF*
Clinical features characteristic of the spinal phenotype
spinal neurofibromas (symptomatic, asymptomatic, absent)
internal neurofibromas (yes, no, unknown)
enlarged spinal nerve roots (yes, no, unknown)
Clinical CutaneousSigns
milk coffee stains, HP: 0000957 (n = 0, 1-5, 6-10, 11-100,> 100)
freckles, HP: 0001480 (groin, axillary, inframammary, absent)
Ocular Clinical Signs
choroid nodules (yes, no, unknown)
Lisch nodules, HP: 0009737 (yes, no, unknown)
Characteristics of neurofibromas
cutaneous (n = 0, 1-10, 11-100,> 100)
subcutaneous, HP: 0100698 (n = 0, 1-10, 11-100,> 100)
plexiform, HP: 0009732 (yes, no, unknown)
Skeletal clinical signs
dural sac dysplasia (yes, no, unknown)
bone dysplasia (yes, no, unknown)
scoliosis, HP: 0002650 (yes, no, unknown)
Tumors
optic glioma, HP: 0009734 (yes, no, unknown)
glioma, HP: 0009592 (yes, no, unknown)
pheochromocytoma, HP: 0002666 (yes, no, unknown)
breast cancer (yes, no, unknown)
MPNST (yes, no, unknown)
leukemia (yes, no, unknown)
other cancers (yes, no, unknown)
Cognitive and Behaviouralprofile
neurodevelopmental delay, HP: 0012758 (yes, no, unknown)

ADHD, HP: 0007018 (yes, no, unknown)
specific learning disorder (yes, no, unknown)
epilepsy (yes, no, unknown)
Development
microcephaly, HP: 0000252 (yes, no, unknown)
macrocephaly, HP: 0000256 (yes, no, unknown)
overgrowth, HP: 0001548 (yes, no, unknown)
short stature, HP: 0004322 (yes, no, unknown)
facial dysmorphism, HP: 0001999 (yes, no, unknown)
Other signs of the nervous system
UBO
Arnold Chiari type I, HP: 0002308 (yes, no, unknown)
hydrocephalus, HP: 0000238 (yes, no, unknown)
syringomyelia (yes, no, unknown)
neuropathy (yes, no, unknown)
Vascular Clinical Signs
heart malformation (yes, no, unknown)
malformation of the vessels (yes, no, unknown)
hypertension, HP: 0000822 (yes, no, unknown)

HP: disease code according to the Human Phenotype Ontology

3.2 Mutational analysis of *NF1* and other genes by Targeted NGS

3.2.1 Design of the NGS custom gene panels

To identify the pathogenic *NF1* variants of the patients' populations, we used two different custom targeted resequencing panels (NGStr2, Table 2; NGStr3, Table 3), produced by Agilent Technologies (SureSelect XT panel). The NGStr2 [51] and NGStr3 panels include the coding regions (10 bases from the 3 'end and 10 bases from the 5' end) and the 5 'UTR and 3' UTR regions of 416 genes, of which 19 associated with RASopathies (including *NF1*), 113 belonging to the Ras pathway, 132 coding for neurofibromin interactors, 143 belonging to the deletion range 17q11.2 and 9 associated with tumours.

Table 2. List of NGStr2 panel genes.

APP	*ADAP2*	*GOSR1*	*TMEM199*	*A2ML1*	*RAC1*
CALM1	*ATAD5*	*IFT20*	*TMEM97*	*BRAF*	*RAF1*
CASK	*COPRS*	*KIAA0100*	*TMIGD1*	*CBL*	*RASA1*
CDC5L	*CRLF3*	*MYO18A*	*TNFAIP1*	*GAB1*	*RASA2*
DYNC1H1	*EVI2A*	*NEK8*	*TP53I13*	*GAB2*	*RASA3*
GRIN1	*EVI2B*	*NSRP1*	*TRAF4*	*GRB2*	*RASA4*
GRIN2B	*LRRC37B*	*NUFIP2*	*UNC119*	*HRAS*	*RASAL1*
MAPK3	*NF1*	*PHF12*	*VNT*	*KRAS*	*RASAL2*
PML	*OMG*	*PIGS*	*ASIC2*	*LIMK2*	*RASAL3*
SDC1	*RAB11FIP4*	*PIPOX*	*C17orf75*	*LRP1*	*RASGRP1*
SDC2	*RNF135*	*POLDIP2*	*CDK5R1*	*LZTR1*	*RASGRP2*
SDC3	*SUZ12*	*PROCA1*	*MYO1D*	*MAP2K1*	*RASGRP3*
SDC4	*TEFM*	*RAB34*	*PSMD11*	*MAP2K2*	*RASGRP4*
SUMO1	*ABHD15*	*RPL23A*	*RHBDL3*	*MAP3K1*	*RIT1*
YWHAB	*ALDOC*	*SARM1*	*RHOT1*	*MAPK1*	*RRAS*
YWHAZ	*ANKRD13B*	*SDF2*	*SPACA3*	*MRAS*	*RRAS2*
	BLMH	*SEZ6*	*TMEM98*	*NRAS*	*SHC1*
	CORO6	*SGK494*	*ZNF207*	*PAK1*	*SHC2*
	CPD	*SLC13A2*		*PAK2*	*SHC3*
	CRYBA1	*SLC46A1*		*PAK3*	*SHC4*
	DHRS13	*SLC6A4*		*PAK4*	*SHOC2*
	EFCAB5	*SPAG5*		*PAK6*	*SOS1*
	ERAL1	*SSH2*		*PAK7*	*SOS2*
	FAM222B	*SUPT6H*		*PTPN11*	*SPRED1*
	FLOT2	*TAOK1*			*SYNGAP*
	FOXN1	*TBC1D29*			
	GIT1	*TIAF1*			
		TLCD1			

In green, neurofibromin interactors; in blue, genes belonging to the 17q11 microdeletion interval; in red, genes belonging to the Ras pathway.

Table 3. List of NGStr3 panel genes

A2ML1	CLK1	NOSIP	TNFSF13B	KIAA0100	TNFAIP1	RALGDS
BRAF	DCLK1	NSFL1C	TOP1	LRRC37B	TP53I13	RAP1A
CBL	DENND1A	NXF1	TOP2A	MYO18A	TRAF4	RAPGEF2
GRB2	DENND4A	OSBPL6	TOP3A	MYO1D	UNC119	RAPGEF5
LRP1	DNAJC7	P4HA3	TRAF6	NEK8	UTP6	RASA1
LZTR1	DYNC1H1	PDE4DIP	USP21	NSRP1	VTN	RASA2
MAP2K1	EFNB2	PHLDB2	VCP	NUFIP2	ZNF207	RASA3
MAP2K2	EIF4E2	PLEKHA7	VSIG1	OMG	ABL1	RASA4
MAPK1	EPHA1	PML	VSIG4	PHF12	AKT1	RASAL1
PTPN11	EPHA2	POLR2A	YWHAB	PIGS	ARF6	RASAL2
RAC1	ESR2	POU2F1	YWHAE	PIPOX	BAD	RASAL3
RAF1	FAF2	PRKACA	YWHAH	POLDIP2	BCL2L1	RASGRP1
RIT1	FAM110B	PTEN	YWHAZ	PROCA1	CDC42	RASGRP2
SHOC2	FAM174A	PTPN13	ZBTB21	PSMD11	CHUK	RASGRP3
SOS1	FAM53C	PTPN14	ZNF638	RAB11FIP4	ELK1	RASGRP4
SOS2	GIGYF1	RALGPS2	ZUFSP	RAB34	ETS1	RELA
SPRED1	GIGYF2	RTKN	ABHD15	RHBDL3	FASL	RGL1
SPRY1	GRB7	SCN3B	ADAP2	RHOT1	FOXO4	RGL2
NF1	GRIN1	SDC2	ALDOC	RNF135	GAB1	RHOA
ACTB	GRIN2B	SDC3	ANKRD13B	RPL23A	GAB2	RIN1
ADCY8	HCN1	SDC4	ASIC2	SGK494	HRAS	RIN2
MLLT4	HDAC4	SH3PXD2A	ATAD5	SARM1	KRAS	RIN3
AGAP2	HLA-DPA1	SH3RF3	BLMH	SDF2	LIMK2	RRAS
ANKRD34A	HTR6	SHANK3	C17orf75	SEZ6	MAP3K1	RRAS2
APP	INPP5E	SIGLECL1	CDK5R1	SLC13A2	MRAS	SDC1
ATF2	KCTD3	SIPA1L1	COPRS	SLC46A1	NFKB1	SHC1
BRCA1	KIF13B	SIRT7	CORO6	SLC6A4	NRAS	SHC2
CA14	KIF1C	SLAMF1	CPD	SPACA3	PAK1	SHC3
CALM1	KSR1	SMARCA4	CRLF3	SPAG5	PAK2	SHC4
CAMSAP2	LIMA1	SMARCD1	CRYBA1	SSH2	PAK3	SYNGAP1
CASK	LPIN3	SOX4	DHRS13	SUPT6H	PAK4	CDKN2A
CAV1	LRFN1	SRGAP2	EFCAB5	SUZ12	PAK6	EGFR
CBY1	MAGI1	SRSF12	ERAL1	TAOK1	PAK7	IKBKG
CCDC8	MAP2K3	STARD13	EVI2A	TBC1D29	PIK3CA	PLA1A
CD274	MLK4	SUMO1	EVI2B	TEFM	PIK3CB	RAPH1
CD79B	MAPK3	SYDE1	FAM222B	TIAF1	PIK3CD	TIAM1
CDC25B	MAPKAP1	TANC2	FLOT2	TLCD1	PIK3R1	TRAP1
CDC25C	MAST3	TBPL1	FOXN1	TMEM199	PLD1	VDR
CDC5L	MYC	TESK2	GIT1	TMEM97	RALA	DDAH1
CDK16	NADK	TGOLN2	GOSR1	TMEM98	RALB	
CGN	NAV1	TIRAP	IFT20	TMIGD1	RALBP1	

In black, the genes associated with RASopathies; in green, the neurofibromin interactors; in blue, the genes belonging to the 17q11 microdeletion interval; in red; the genes belonging to the Ras pathway; in grey, the genes associated with tumours

3.2.2 Target Resequencing

In this study, the operational flow consists of four main stages.

- Preparing Sequencing Libraries
- Selection of fragments of interest by hybridization
- High-resolution sequencing
- Bioinformatics analysis of data

Preparation of sequencing libraries

Three to five microliters of whole blood were drawn from each subject at the Fondazione IRCCS Istituto Neurologico Carlo Besta, and DNA was extracted from 3 ml of whole blood samples using the Gentra Puregene Blood Kit (Qiagen).

For each sample, 3 µg of gDNA, resuspended in Tris-EDTA, were subjected to random mechanical fragmentation, using a sonicator (Covaris), in order to obtain fragments on average of 150-200 bp. Next, the samples were purified using magnetic beads (AMPure XP, Beads-Beckman Coulter) and qualitatively evaluated using the 2200 TapeStation capillary electrophoresis platform (D1000 kit, Agilent Technologies).

In order to make the ends of the fragments flat (i.e. free of extensions), the reaction called DNA fragment repair was performed by adding 52 µl of End Repair Master Mix to the sample (Table 4). Then, 20 µl of dA-Tailing Master Mix (Table 5) was added to an aliquot of the sample thus obtained, to perform adenylation of the 3 'ends of the DNA fragments. Finally, the ligation of the adapters was obtained by adding 37 µl of Ligation Master Mix to an aliquot of the adenylated sample (Table 6). The protocol set on the thermal cycler, for the last three reactions, is described in Table 4, Table 5 and Table 6, respectively.

Table 4. Mix and thermal cycler conditions used for the DNA fragment repair reaction

Mix		**Termocycling program**	
Reagents	**Volume**	**Time**	**Temperature**
10X End Repair Buffer	10 µl	30’	20 °C
dNTP mix (100 mM)	1,6 µl	*Hold*	4 °C
T4 Polymerase	1 µl		
Klenow DNA Polymerase	2 µl		
T4 Polynucleotide Kinase	2,2 µl		
DNA sample	x µl		
H_2O	35,2 µl		

Table 5. Mix and conditions at the thermocycler used for the adenylation reaction of the 3 'ends of the DNA fragments

Mix		**Termocycling program**	
Reagents	**Volume**	**Time**	**Temperature**
10X Klenow Polymerase Buffer	5 µl	30’	37 °C
dATP	1 µl	*Hold*	4 °C
Exo(-)Klenow	3 µl		
DNA sample	30 µl		
H_2O	11 µl		

Table 6. Mix and conditions at the thermal cycler used for the ligation reaction of the adapters

Mix		**Termocycling program**	
Reagents	**Volume**	**Time**	**Temperature**
5X T4 DNA Ligase Buffer	10 µl	15’	20 °C
SureSelect Adaptor Oligo Mix	10 µl	*Hold*	4 °C
T4 DNA Ligase	1,5 µl		
DNA sample	13 µl		
H_2O	15,5		

At this point, the amplification of the libraries were linked to the specific adapters, by PCR, and the purification of the same with magnetic beads were performed. For the PCR reaction mix, prepared in a final volume of 50 µl, the SureSelect Pre-Capture PCR Reaction Mix was used. The reagents are shown in Table 7, which also describes the program that has been set on the thermal cycler.

Table 7. Mix and conditions at the thermal cycler used for the pre-capture PCR reaction

***Mix* PCR**		**Termocycling program**		
Reagents	**Volume**	**Time**	**Temperature**	**N° cycles**
5X Herculase II Reaction Buffer	10 µl	2’	98 °C	
dNTP mix (100 mM)	0,5 µl	30’’	98 °C	
Herculase II Fusion DNA Polymerase	1 µl	30’’	65 °C	6
SureSelect Primer	1,25 µl	1’	72 °C	
SureSelect ILM Indexing Pre-Capture PCR reverse primer	1,25 µl	10’	72 °C	
		Hold	4 °C	
DNA sample (library)	15 µl			
H_2O	21 µl			

Finally, the quality of the samples was checked using the 2200 TapeStation platform, in order to evaluate the electrophoretic profile of the libraries and their concentration, expressed in ng / µl.

Selection of fragments of interest by hybridization

For the hybridization reaction, an aliquot containing 750 ng of each DNA library was subjected to denaturation, by adding 5.6 µl of SureSelect Block Mix. For the reaction mix, the protocol shown in Table 8 was followed, which also describes the program that was set up on the thermal cycler.

Table 8. Mix and conditions at the thermal cycler used for the denaturation reaction

Mix		**Termocycling program**	
Reagents	**Volume**	**Time**	**Temperature**
SureSelect Indexing Block 1	2,5 µl	5’	95 °C
SureSelect Block 2	2,5µl	*Hold* (at least 5’)	65 °C
SureSelect ILM Indexing Block 3	0,6 µl		
DNA sample	750 ng		

To capture the regions of interest, the DNA libraries are hybridized to the Capture Library <3Mb, containing the panel probes, by adding 20 µl of Capture Library Hybridization Mix to the sample and incubating for 24 h at 65 ° C (Table 9).

Table 9. Mix and conditions at the thermal cycler used for the hybridization reaction.

Mix		**Termocycling program**	
Reagents	**Volume**	**Time**	**Temperature**
Hybridization Buffer mixture	13 µl	24 h	65 °C
10% RNase Block solution	5µl		
Capture Library <3 Mb	2 µl		
DNA sample	~10 µl		

Washes with Streptavidin-coated Magnetic Beads were performed to remove fragments that did not hybridize. The hybridized DNA, bound to the beads, was recovered for the next steps.

Indexing of samples

To proceed with the sequencing protocol, it was necessary to add to each enriched sample the sequences indexed by means of an amplification reaction.

The sequences of *indexing primer* (8 bp) are reported in "*SureSelect*^{XT} *Target Enrichment System* for *Agilent Technologies' Illumina Paired-End Multiplexed Sequencing Library"* protocol. In Table 10, the Post-Capture PCR reaction Mix and the program that was set on the thermal cycler are shown.

Table 10. Mix and cycler conditions used for the post-capture PCR reaction

***Mix* PCR**		**Termocycling program**		
Reagents	**Volume**	**Time**	**Temperature**	**N° cycles**
5X Herculase II Reaction Buffer	10 µl	2'	98 °C	
dNTP mix (100 mM)	0,5 µl	30''	98 °C	
Herculase II Fusion DNA Polymerase	1 µl	30''	57 °C	
SureSelect ILM Indexing Post-Capture Forward PCR primer		1'	72 °C	
		10'	72 °C	12
SureSelect 8 bp Index (reverse primer)	1 µl	*Hold*	4 °C	
DNA sample	5 µl			
H_2O	14 µl			
	18,5 µl			

High-resolution sequencing: pooling and dilution of samples

The indexed libraries have been grouped in such a way that they are present in equimolar quantities within each pool. Subsequently, the final volume of each pool was corrected to the desired concentration of 4 nM and the sequencing was performed using MySeq (NGStr2, 2x300 bp) and NextSeq550 (NGStr3, 2x150 bp) sequencers (Illumina, San Diego, CA, USA).

3.3 Bioinformatics analysis of data from NGS

The Bioinformatics analysis of data from NGS has been performed in collaboration with Istituto di Biotecnologie Biomediche – Consiglio Nazionale Ricerche. The reads quality assessment and trimming for length at 200 bp were obtained by means of FastQC (v. 0.11.8; https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and Trimmomatic (v. 0.36)[52], respectively. Then, the QC-checked paired end (PE) reads of each sample were mapped to NCBI human reference genome (build GRCh38) using BWA-MEM aligner (0.7.10-r789) [53]. The mapping was done allowing for maximum 3 mismatches and with other default parameters of BWA. Using samtools[54], we then remove the duplicate reads due to PCR amplification during library preparation. For each sample, we retain only high quality (HQ) alignments in sorted BAM files (HQ-BAM) by filtering out unmapped reads and those alignments with mapping quality (MAPQ) less than 15. These high-quality alignments (HQ-BAMs) are then checked for overall mapping statistics (mapping-QC) by an in-house script.

After that, GATK software (v. 3.4) [55] was used to perform quality score recalibration (using the TableRecalibration walker), local realignment around known indels (using the IndelRealigner walker) and variant calling (by the HaplotypeCaller walker) for both single nucleotide variants (SNVs) and insertions/deletions (indels). Poorly confident variants having QUAL < 150, Fisher Strand (FS) strand bias > 60 for SNV and > 200 for indels, or three SNVs within 10 base-windows were flagged for removal in the FILTER field of the VCF file.

3.4 Variant annotation

The functional annotation and impact prediction were performed using ANNOVAR (v.2019Oct24)[56], to identify the type of variation (silent or synonymous mutations, missense, splicing, frameshift, in frame deletions and insertions), their gene position, and the presence of any known polymorphisms in the population, reported in the dbSNPs (dbSNP144), 1000genomes (1000g2015) and ExAC databases. The pathogenetic effect also has been predicted as ANNOVAR includes 20 pathogenicity predictors for SNVs (SIFT, SIFT4G, Polyphen2-HDIV, Polyphen2-HVAR, LRT, MutationTaster2, Mutation Assessor, FATHMM, MetaSVM, MetaLR, PROVEAN, FATHMM-MKL coding, FATHMM-XF coding, fitCons, M-CAP, PrimateAI, DEOGEN2, BayesDel no AF, BayesDel add AF, ClinPred, LIST-S2) and the Damagepredcount, a descriptor with a numerical value from 1 to 20 indicating how many of the 20 predictors evaluate the variant as pathogenic. The software also allows to classify variants according to the ACMG-AMP2015 guidelines in *benign*, *likely benign*, *uncertain*, *pathogenic*, *likely pathogenic*, as it also reports data from the Clinvar database (database that annotates and classifies according to ACMG criteria- AMP 2015 the clinically validated variants in patients) and Intervar[57] (web tool for the clinical interpretation of genetic variants), which automatically generates the interpretations for the 28 different ACMG-AMP2015 criteria, to provide the final interpretation of the variant.

3.5 RNA extraction

To isolate RNA from human whole blood, we used the "TempusTM Spin RNA Isolation" Kit (Applied Biosystems). The procedure involves three steps:

1. Blood collection: peripheral venous blood is collected directly in the TempusTM Blood RNA Tube, containing a reagent that lyses the cells and stabilizes the RNA and is vortexed for 10 seconds.
2. Process: The stabilized blood is transferred into a 50 mL tube and diluted with 1X PBS. Once vortexed for at least 30 seconds, making sure that the lysate is directed towards the top of the tube, and centrifuged at 4 ° C at 3000 x g (rcf) for 30 minutes, the supernatant is discarded and the pellet containing the RNA resuspended.
3. Purification and elution: the resuspended RNA is transferred through tubes containing special filters that purify the RNA through multiple centrifugation steps and from which the RNA is finally eluted. The RNA thus extracted is stored at -80 ° C.

3.6 Reverse transcription

To carry out the reverse transcription reaction, the Maxima H Minus cDNA Synthesis Master Mix kit with dsDNA (Termofisher) was used, which allows you to combine the elimination of genomic DNA (gDNA) and the synthesis of cDNA. This system allows to obtain cDNA that can be used in quantitative PCR experiments (real time PCR).

500 ng of total RNA from each patient was reverse transcribed.

The digestion of the gDNA was carried out by adding to 500 ng of RNA, 1 µL of 10X DNase buffer and 1 µL of dsDNase, bringing the whole to a final volume of 10 µL, by adding H_2O nuclease free.

Each preparation was mixed, centrifuged, and incubated at 37 ° C in a thermostatic block for 2 minutes and then deposited on ice.

To obtain reverse transcription, 4 µL of Maxima cDNA H Minus Synthesis Master Mix (5X) and 6 µL of Nuclease free H_2O were added to each sample.

The samples were then placed in a thermal cycler according to the protocol reported in table 11.

Table 11. Reverse transcription conditions

STEP	TEMPERATURE (°C)	TIME
0. Heating	25	10’
1. Activation	50	30’
2. Inactivation	85	5’
3. Hold	10	Hold

Then, the obtained cDNA was diluted 1: 3, adding 40 µL of nuclease free H2O to the 20 µL of cDNA, to be ready for use in real time PCR.

3.7 Digital PCR on *NF1* double mutations

The samples were initially tested in quantitative real- time PCR (qPCR) to verify the quality and quantity of the cDNA.

For assays test in qPCR 5ng of cDNA was amplified (in triplicate) in a reaction volume of 10 µl containing the following reagents: 5 µl of "TaqMan® Fast Advanced Master Mix, Thermofisher", 0.25 µl of "TaqMan SNP Gene expression assay 40x, Thermofisher" FAM and VIC labelled, Table 12). The specific Taqman probes for the wild type 62T and 528T alleles were VIC-dye labelled and those specific for the mutated 62A and 528A alleles were FAM-dye labelled. The specific Taqman probe for the wild type 2446C was VIC-dye labelled and that specific for the mutated 2446Tallele was FAM-dye labelled.

Table 12. Primers and Taqman probes sequences used to test *NF1*: c.62T>A, *NF1: c.*528T>A and *NF1: c.*2466C>T alleles expression.

Gene symbol	Forward Primer Seq.	Reverse Primer Seq.	Reporter 1 Dye	Reporter 1 Sequence	Reporter 2 Dye	Reporter 2 Sequence
NF1_62TA	CCGTGGTCAGCCGCTT	TGTGTGTTCTGCTGTCCTGTTTT	VIC	ACGAGCAGCTTCCAAT	FAM	CGAGCAGCATCCAAT
NF1_528TA	CTGTTTGTTCAGAAGACAATGTTGATGT	CCTTCAGGAGTCGTTTTAATTTTGCA	VIC	ACTGTAACAATTCTATATCATG	FAM	CTGTAACAATTCTATTTCATG
NF1_2466CT	GCTTTCAGCAGCCTGGC	TGAGTGGAGGAGGATCCATAGATTT	VIC	TAAGAGGCGAATGTCCCA	FAM	TAAGAGGTGAATGTCCCAT

Real-time PCR was carried out on the Quant Studio 12K (Thermofisher), using a pre-PCR step of 20s at 95°C, followed by 40 cycles of 1s at 95°C and 20s at 60°C. One NTC sample was run. 2.5 ng of cDNA was amplified in a reaction volume of 15 µl containing the following reagents: 7.5 µl of "QuantStudio™ 3D Digital PCR Master Mix v2, Thermofisher", 0.375 µl of "TaqMan SNP Gene expression assay 40x, Thermofisher" FAM and VIC labeled. The mix was loaded on the chip using the QuantStudio 3D Digital PCR Chip Loader.

The chips were then loaded on the Proflex PCR System (Thermofisher), and the PCR was carried out using a pre-PCR step of 10min at 96°C, followed by 39 cycles of 2min at 60°C and 30s at 98°C, followed by 2min at 60°C.

Data were analyzed with "QuantStudio 3D Analysis Suite Cloud Software".

3.8 Quantitative real- time PCR

For qPCR assays we selected the genes *SDC2, SDC3* and *SDC4* belonging to the syndecans family, with an expression level in peripheral blood greater than 0.5 TPM (transcripts per million). The *SDC1* gene was excluded from the analysis because its expression level in peripheral blood was <0.5 TPM (GTEx portal source, https://gtexportal.org). Each SYBR Green qPCR assay was performed using GoTaq–qPCR master mix (Promega) and run on a QuantStudio

5 Real-Time PCR Systems (Thermo Fisher Scientific). The SYBR green molecules are non-specifically intercalated with the double-stranded DNA, therefore an accurate design using Primer3 (https://primer3.ut.ee) of the oligonucleotides were performed (table 13) to amplify and detect only the target sequences of interest. Three pairs of oligonucleotides were obtained, in each pair one of the oligonucleotides was designed between an exon-exon junction to reduce the risk of amplifying gDNA in a non-specific way.

Table 13. Sequence of primers for qPCR assays. Ta = annealing temperature

Gene target	**Oligonucleotide name**	**Sequence (5' to 3')**	**Ta**
SDC2	*SDC2*_qPCR_FW	CCTGCTCAGACAAAGTCACC	57°C
	*SDC2*_qPCR_REV	TTGTATCCTCTTCGGCTGGG	57°C
SDC3	*SDC3*_qPCR_FW	TGGCTCAGACCCCAACTCC	57°C
	*SDC3*_qPCR_REV	TCTCTTCTTCTGGCAGCTCG	57°C
SDC4	*SDC4*_qPCR_FW	GATCTGGATGACTTGGAAGAC	57°C
	*SDC4*_qPCR_REV	TTCGGTGGGGACTTGGCTC	57°C

Ta = Annealing temperature

3.9 Statistical Analysis

3.9.1 NF1 variants' distribution

The X2 test or Fisher exact test was used to compare categorical variables. The Benjamini-Hochberg (B-H) method with false discovery rates of 0.05, 0.025 and 0.01 was used to correct p-values for multiple testing. A p value <0.05 was considered as statistically significant. The X2 test or Fisher exact test was performed using the tools available at https://www.socscistatistics.com/tests. The B-H correction was performed using the tool available at https://tools.carbocation.com/FDR.

3.9.2 qPCR analysis

All qPCR experiments were run in triplicate and the average of the threshold cycles (Ct) for each sample was made. To determine the relative gene expression, the $2^{-\Delta Ct}$ method was applied (ΔCt=Ct gene target – Ct housekeeping gene, for each sample). For each gene analyzed, mean, standard deviation, standard error of the mean, and confidence intervals values were calculated in the three groups of samples, which include 16 patients with SNF, 16 patients with classical NF1 and 16 healthy controls. The equal variance Student's *t*-test was applied to compare the means (Table S3 and S4) and the B-H correction for multiple tests with false discovery rate of 0.05, 0.025 and 0.01 was applied. The outliers' values, identified by Tukey test with k=1.5, were excluded from the analysis. The Student's *t*-test and Tukey test were performed using the tools available at https://www.socscistatistics.com/tests.

3.10 NF1 interactors selection

The NF1 interactors were selected by means of the IntAct tool [58] between the interactors with an experimentally proven interaction with *NF1* obtained by socioaffinity inference, two hybrid, anti-tag coip, anti-bait coip, crosslink and two hybrid pooling methods. They were also selected because of the evidence collected in the review by Ratner et al. [59].

4. RESULTS

4.1 Patients' cohorts

4.1.1 Demographics

On 30 June 2020, 768 subjects affected by NF1 and followed by "Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Milano, Italia", "Fondazione IRCCS Istituto Neurologico Carlo Besta, Milano, Italia", and Università della Campania "Luigi Vanvitelli", Dipartimento di Scienze Mediche, Napoli, Italy, had performed spinal MRI. In 220 (28.6%) cases MRI reviewed by a specialist neuroradiologist showed spinal neurofibromas: in 81 (36.8%) bilateral neurofibromas involving all spinal roots were present (SNF), while in 139 (63.2%) single or few isolated spinal neurofibromas were detected.

4.1.2 Clinical characteristics of the SNF cohort

Demographic and clinical characteristics were analyzed in all 81 SNF (26 female and 55 male) patients: 55 from unrelated families and 26 belonging to 19 families with at least one member affected by SNF. Median age was 35 ranging from 15 to 74 years. SNF patients' clinical features are reported in Table 14**.**

Table 14. Clinical features of SNF patients

NF1 features	Number individuals (%) (95%CI)	NF1 features	Number individuals (%) (95%CI)
InternalNeurofibromas	37/81 (45.7) (35-56)	*ADHD*	2/81 (2.5) (0.7 -8)
Nerve Roots Swelling	26/81 (32.1) (23-43)	*Learning specific disorder*	12/78 (15.3) (9-25)
		Headache	8/80 (10) (5.1-18)
Café AuLait spots			
< 6 Cal	15/81 (18.5) (11-28)	*Microcefalia*	0
6-10	49/81 (60.5) (50-70)	*Macrocefalia*	19 (23.5) (15-34)
11-100	26/81 (32.1) (1-31)	*Overgrowth*	1 (1.2) (0.2-6.6)
>100	0	*Short Stature*	6/81 (7.4) (3.4-15)
Freckling	50/81 (61.7) (50-71)	*FacialDysmorphism*	8/81 (9.9) (5-18)
Lischnodules	59/74 (79.7) (69-87)	*UBO*	37/78 (47.4) (36-58)
		Arnold Chiari	1/81 (1.2) (0.2-6.6)
		Hydrocefalus	6/81 (7.4) (3.4-15)
CutaneousNFs		*Siryngomyelia*	2/81 (2.5) (0.7 -8)
0	9/81 (11.1) (6-19)	*Neuropathy*	9/80 (11.2) (6-20)
1 -10	36/81 (44) (34 -55)	*Heart Malformation*	3/78 (3.8) (1.3-11)
11-100	26/81 (32.1) (23-43)	*Vessels Malformation*	10/75 (13.3) (7.4-23)
>100	10/81 (12) (6-21)	*Hypertension*	7/81 (8.6) (4.2-17)
SubcutaneousNFs			
0	23/81 (28.4) (2-39)		
1 -10	38/81 (65.5) (36-58)		
11-100	20/81 (25) (16-35)		
>100	0		
PlexiformNeurofibromas	41/81(50.6) (40-61)		
Dural Sac Dysplasia	2/81 (2.5) (0.7 -8)		
Bone Dysplasia	2/81 (2.5) (0.7-8)		
Scoliosis	34/81 (42) (32 -53)		
Optical Glioma	7/81 (8.6) (4.6-16.7)		
OtherGliomas	9/81 (11.1) (5.9-20)		
Pheochromocytoma	3/81 (3.7) (1.3-10)		
Breast Cancer	0		
MPNST	6/81 (7.4) (3.4-15)		
Leukemia	0		
Othertumors	3/81 (3.7) (1.3-10)		

Spinal neurofibromas were symptomatic in 44 out of 81 (54.3%). In most cases internal neurofibromas (45.7 %) and nerve roots swelling (32.1%) were also found.

15 out of 81 (18.5%) SNF patients had spinal surgery: ten at cervical level, two at cervical and at lumbar level, three at lumbar level. Histopathological diagnosis were neurofibromas in 13 and ganglioneurofibromas in 2 cases.

15/81 (18.5%) cases had less than 6 CALS and 50/81 (47.3%) and no freckling. 9/81 (11.1%) had neither more than 5 CALS nor freckling, they fulfilled the NIH diagnostic criteria for the presence of other clinical signs such as neurofibromas and Lisch nodules. Only 17/81 (21%) individuals had more than 10 CALS.

Cutaneous NFs were present in 88.9 % , in most cases in a low number: less the than ten NFs in 36/81 (44.4%) of patients. Subcutaneous neurofibromas were observed 71.6 % of individuals. Both cutaneous and subcutaneous NF were present in 53/81 cases (65.4%). Plexiform neurofibromas were observed in 41/81 cases (50.6).

For 74 patients, data on the presence or absence of Lisch nodules was available. They were present in 59/74 (79.7%).

Symptomatic and asymptomatic OPGs, were observed in 2.5% (2/81) and 6.2 % (5/81) of subjects respectively. Gliomas other than OPGs were present 9/81 (11.1 %), (4 brainstem gliomas, 1 glioblastomas, 3 pilocytic astrocytomas, 1 subependymal astrocytoma); two cases had both optic glioma and brainstem gliomas. Other malignancies different from central nervous system tumors were observed in 12 individuals.

The most common skeletal abnormalities was scoliosis, present in 34/81, (42%), two patients had bone dysplasia and other two dural ectasia (2.5)%. Hydrocephalus was found in 6/81 (7.4%), syringomyelia in one and Arnold Chiari in another one (1.2%).

Neuropathy was resent in nine out of 80 (11.3 %), headache was referred in 8/80 (10 %) cases, four patients had epilepsy, (4.9%).

Facial dysmorphic features were observed in 8/81 (9.9%) cases, macrocephalia in 19/81 (23.5 %) short stature in 6/81 (7.4%) and overgrowth 1/81 (1.2%).

Neurodevelopment delay was reported in 9/80 patients (11.2%) and learning disability in 11/78 (14.1%).

Hypertension was the common cardiovascular disorders observed in 7 /81 (8.6%); heart and vessels malformation were found 3/78 (3.8%) and 10/75 (1.3%) patients respectively.

4.1.3 Comparisons of clinical characteristics observed in the SNF cohort, in the cohort of classical NF1 phenotype of our Institutions, and in previously described classical NF1 cohorts from literature

The clinical characteristics of our SNF cohort were compared with those observed in a cohort of classical NF1 patients (i.e. patients without spinal tumours) followed by the same institutions. They were 68 females and 38 males, median age 47 (30-75 years). Furthermore, when data were available, clinical features were also correlated with those previously reported in large–scale NF1 classical cohorts, already used in other genotype-phenotype studies. All comparisons are reported in Table 15.

Table 15. Comparisons of clinical characteristics observed in the SNF cohort and in the cohort of classical NF1 phenotype of our Institutions

	SNF (81)	**Classical (106)**	**p value SNF versus classical NF1**
Symptomatic spinal NFs	44/81 (54)	0	/
Internal Neurifibromas	37/81(45.7)	7/106 (6.6)	<0.001**
Nerve Roots Swelling	26/81 (32.1)	3/106 (2.8)	<0.001**
> 5 CALS	66/81(81.5)	99/106 (93.4)	0. 012*
Skin fold freckling	50/81 (61.7)	86/106 (81.1)	0.003*
Lisch nodules	59/74 (79.7)	87/104 (83.7)	0.5
Cutaneous NFs	72/81 (88.9)	99/106 (93.4)	0.27
Subcutaneous NFs	58/81 (71.6)	60/106 (56.6)	0.035
Plexiform Neurofibromas	41/81(50.6)	41/106 (38.7)	0.1
Dural Sac Dysplasia	2 /81(2.5)	5/103 (4.9)	0.4
Skeletal abnormalities without scoliosis	2/81(2.5)	6/106 (5.7)	0.28
Scoliosis	34/81 (42)	35/106 (33)	0.2
Symtomatic OPGs	2/81(2.5)	3/106 (2.8)	0.87
Asymptomatic OPG	5/81 (6.2)	5/106 (4.7)	0.66
Other malignant neoplasms	17/81 (21)	31/106 (29.2)	0.12
Cognitive impairment and/or learning disability	8/80 (10)	13/101 (12.9)	0.71
Epilepsy	4 /81(4.9)	8/106 (7.5)	0.48

Heahache	8/80 (10)	25/105 (23.8)	0.015*
Macrocefalia	19/81 (23.5)	31/106 (29.2)	0.37
Short Stature	6/81 (7.4)	9/106 (8.5)	0.78
Facial Dysmorphism	8 /81 (9.9)	14/106 (13.2)	0.65
Neuropathy	9/80 (11.2)	5/104 (5.1)	0.13
Cardiovascular abnormalities	13/78 (16.7)	22/97(22.7)	0.32
Hypertension	7 /81(8.6)	24/106 (22.6)	0.015*

Statistically significant p value with FDR of 0.05 are indicated by * and p value with FDR 0.01 by **.

The main features of our SNF cases were a higher number of internal neurofibromas (45.7 vs 6.6% p<0.001) as well as nerve-root swelling (32.1 vs 2.8 p<0.001) compared to our classical NF1 cohort. In our classical NF1 cohort patients with spinal NF were deliberately excluded, therefore no symptoms such as low back pain or neurological deficit related to spinal involvement were reported. As concerns pigmentary manifestations, in our SNF cases the number of patients with freckling (61.7% vs 81.1 % and 84.2%), and with >5 CALs (81.5% vs 93.4% vs 89%) was significantly less frequent than those already reported or observed by us in classical NF1.

4.1.4 Phenotypic variability within SNF families

We identified 19 NF1 families with at least one member affected by SNF. In Figure 10, family pedigrees are reported. Two families had four affected members, five three affected relatives and twelve families two. Overall, 26 patients had SNF, 12 MNSFR, 9 a classical form of the disease. We observed a phenotypic variability within the cohort of SNF families. Most families (13/19) had all NF1 individuals affected by SNF or MNSFR

In 3 other families (2 with three affected and one with four NF1 cases) two members had SNF, while the other affected member had MNSFR (family N° 1), classical NF (family N° 9) and one affected MNSFR and the other with classical form of the disease within the same family (family N° 17).

As proposed by Ruggieri we called “Pure SNF families” families in which all affected members had SNF, “partial SNF families” families in which all affected members had SNF or MNSFR and “multiple phenotype families” families in which at least one member had SNF and the others affected members MNSFR or classical NF1.

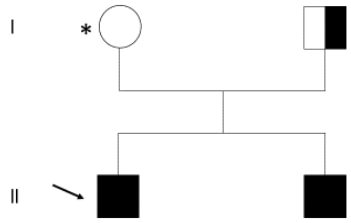
2 families, (N° 10, 12) were “pure” SNF families.

We identified nine partial SNF families (N° 1, 4, 6, 7, 15, 16, 19)

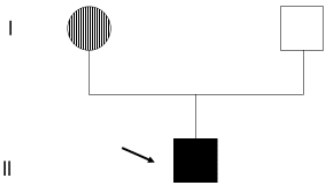
8 families were multiple phenotype families (N° 2, 3, 5, 8, 9, 11, 13,14, 17, 18).

No phenotypic differences were observed between SNF cases belonging to “pure”, “partial” or “multiple Phenotype” families as well between SNF patients included in 19 families and the others SNF patients.

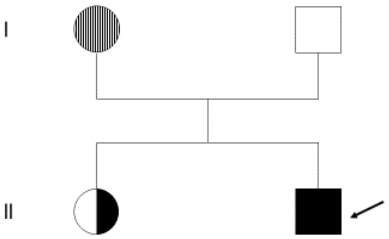
Family 1



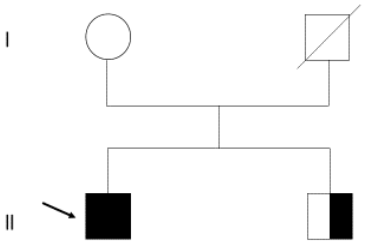
Family 2



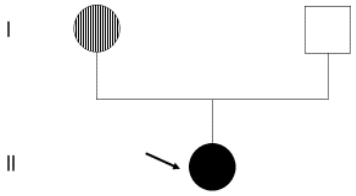
Family 3



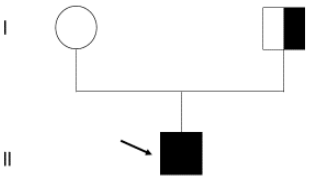
Family 4



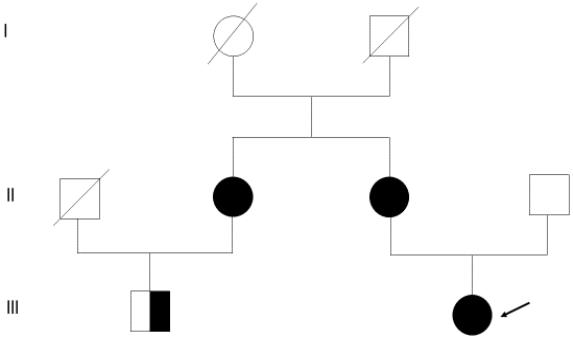
Family 5



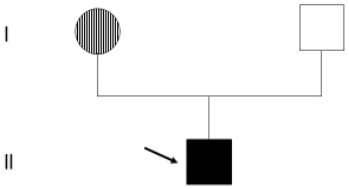
Family 6



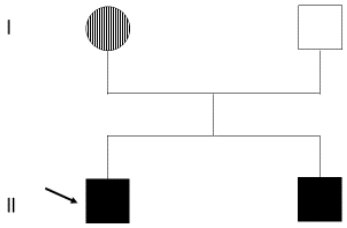
Family 7



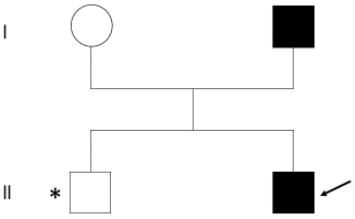
Family 8



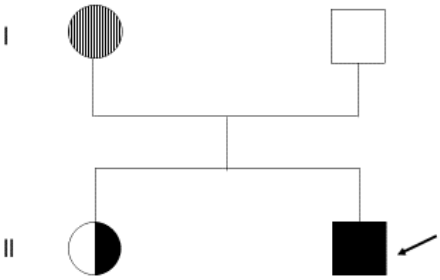
Family 9



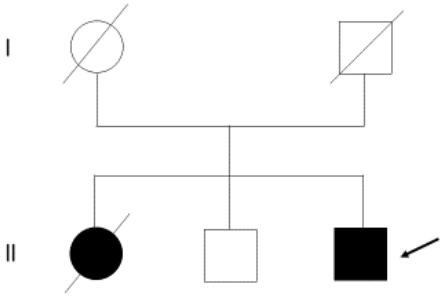
Family 10



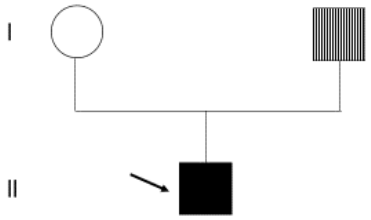
Family 11



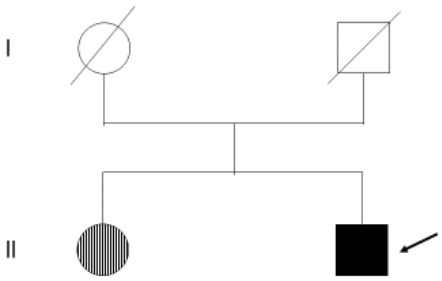
Family 12



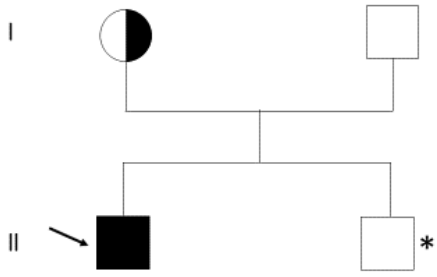
Family 13



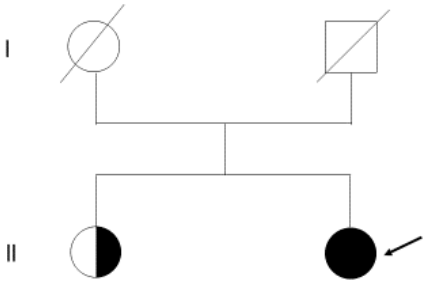
Family 14



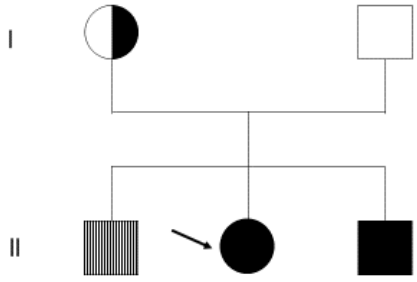
Family 15



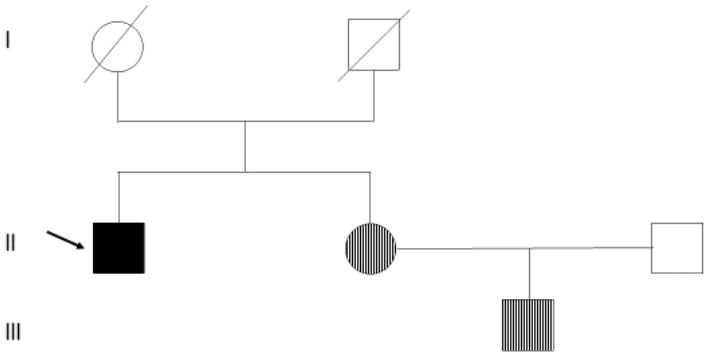
Family 16



Family 17



Family 18



Family 19

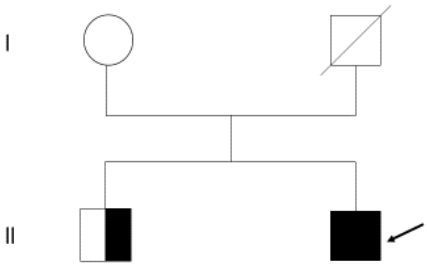
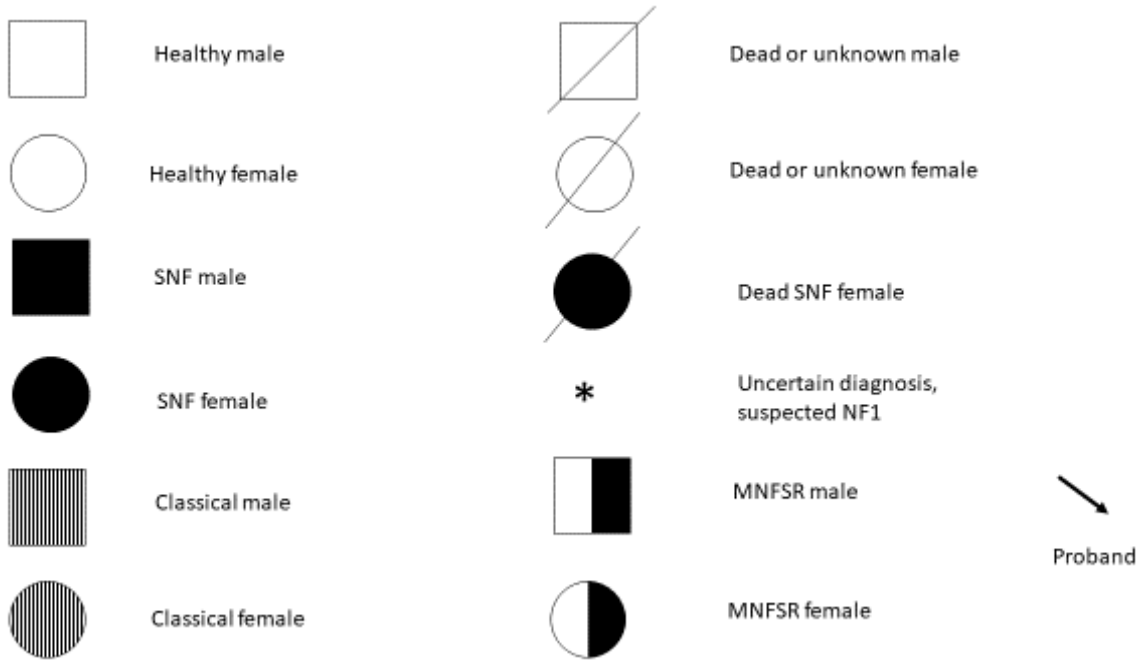


Figure 10. Genetic pedigrees of the 19 SNF families.

Legend:



Neurofibromatosis type 1 is a pathology that is expressed in different forms. Classical NF1 and spinal appear to be determined by mutations in the NF1 gene. To date, the genetic causes responsible for the classical or spinal form are not known, even if the two forms are clearly distinguishable from a clinical point of view. Having a large range of cases related to the two forms, we looked for mutations in genes related to the RAS pathway, neurofibromin interactors and genes of the 17q11.2 region by Targeted Resequencing. Genes from the 17q11.2 region were also included in the gene panel, since patients with microdeletion syndrome, whose results are

not the subject of this thesis, were also analyzed. To check for the presence of mutational profiles specific to the two forms, gDNA were sequenced from patients with classical NF1 and from patients with the rarest spinal form (SNF) by NGS targeted resequencing (panel NGStr3). Additional data from 24 SNF patients, previously analyzed with an NGS targeted resequencing panel (NGStr2) in 2017, were added to the statistical analyses.

4.2 Indicators of the quality of sequences obtained by NGStr3 and NGStr2 sequencing

The sequencing obtained with the NGStr3 panel produced an average number of reads of 3689056,479 and of these more than 99% were mapped correctly. After eliminating duplicates due to PCR steps, the percentage of mapped reads, dropped to 98% on average, achieving an average depth of sequencing (*mean depth*), or the coverage of each individual base, of 54.867 (MIN: 1.385; MAX: 105.354). The target was covered with an average of 97364 (*mean coverage*).

The sequencing obtained with the NGStr2 panel produced an average number of reads of 345522.0 and of these more than 99% were mapped correctly. After eliminating duplicates due to PCR steps, the percentage of mapped reads, dropped to 98.5% on average, achieving an average depth of sequencing (*mean depth*), or the coverage of each individual base, of 54.8 (MIN: 45.176; MAX: 141.371). The target was covered with an average of 98452 (*mean coverage*).

The raw reads of panel NGStr3 data are available in NCBI Short-read Archive (SRA, https://www.ncbi.nlm.nih.gov/sra) under the accession number PRJNA8509016 and the raw reads of panel NGStr2 under the accession number PRJNA688415.

4.3 Variants' annotation by Annovar

Variants of all NGStr3 and NGStr2 panel genes of 100 classical NF1 patients (6 patients out of 106 had large *NF1* mutations, detected by MPLA) and 63 spinal (7 patients out of 74 had large *NF1* mutations, detected by MPLA, and 4 patients were negative for *NF1* mutations) patients characterized by a suitable qualitative process (pass), and which are annotated in the Annovar software, were selected. The total number of variants in the classic NF1 patients is 11347, while in the spinal it is 8653. The type of variants has been defined according to the classification present in the Reference Sequence Database: splicing, exonic, intronic, intergenic, ncRNA, upstream, downstream, UTR3 and UTR5 variants. The exonic variants can be divided into

synonyms, not synonymous, start/stop loss, stop gain, insertions and frameshift and non-frameshift deletions.

4.4 Comparative analysis of *NF1* variants in classical and spinal patients

To increase the number of patients with spinal form, in the study of variants in the *NF1* gene, an additional 24 SNF patients, previously sequenced with the NGStr2 panel in 2017, were added to the 50 SNF patients analyzed with the NGStr3 gene panel. Of these 74 patients, 55 are sporadic SNF, while 19 are probands belonging to families with pure SNF or mixed families in which there are other subjects suffering from MNFSR or classical form of NF1 (figure 10). Family members were excluded from the comparison between the population of classical and spinal patients.

First, an analytical study was carried out to understand the pathogenic significance of the SNPs and indel variants present in the *NF1* gene as the causative gene of neurofibromatosis.

For the analysis of the *NF1*gene, the filters represented in figure 11 were applied on 100 classical and 63 spinal patients.

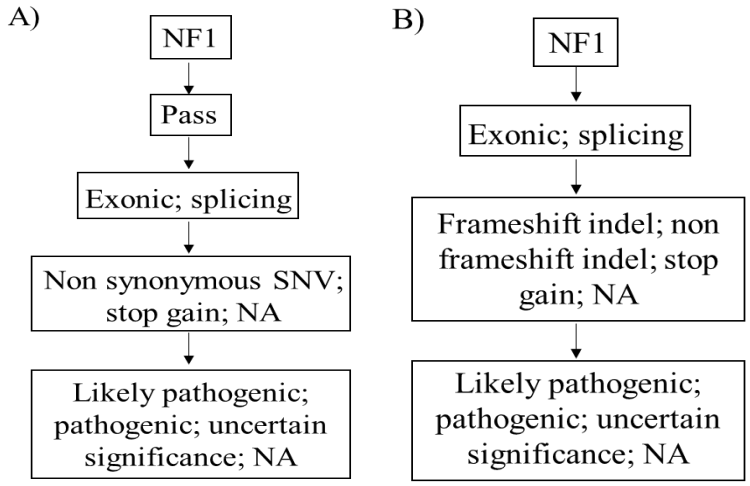


Figure 11. Pipeline for the selection of SNPs (A) and indel (B) variants of the NF1 gene

4.4.1 Mutational analysis

NF1 gene mutations were detected in 186 patients and in 2 relatives by NGS and in 18 cases by MLPA method. The MLPA/NGS approach was validated by Sanger sequencing on DNA or RNA of both positive and negative cases at Besta Institute. We identified 160 different NF1 gene mutations. *NF1* mutations observed are reported (Table 16a, 16b, 16c) with molecular details (DNA, RNA, protein change) and the classification of the variants by type, tertile [60] and domains [61]. In Table 16a are described nineteen families with at least one affected by SNF, in Table 16b 55 SNF sporadic cases, and in Table 16c, 106 classical NF cases.

17 mutations, detected in classical patients and 28 in SNF (13 belonging to the SNF families), were never reported as new mutations, the others were already described (Table 16).

Five mutations were present in both unrelated SNF and classical patients c.288+1delG; c.1318C>T; c.2033_2034dupC; c.5546G>A; c.6789_6792delTTAC.

Four SNF patients were negative for *NF1* mutations (also by MPLA approach).

Interestingly, five SNF patients show more than one *NF1* variant. For three familial cases belonging to three families, it was possible to infer whether one or both NF1 alleles were affected. Family 1, family 17 and family 18 (Figure 10) were informative to answer the above question.

Precisely, in family 1, the SNF patient 1136, who carries the *NF1*:c.62T>A *(p.Leu21His)* missense variant inherited from his MNFSR father (1139) and shared by his SNF brother (1140), shows a second *NF1:*c.528T>A *(p.Asp176Glu)* missense mutation with uncertain clinical significance, inherited from his mother, indicating that the two missense variants are *in trans*. Despite his mother apparently not affected, the *p.Asp176Glu* substitution was predicted to be damaging by 9/20 predictors (Annovar) and may have a subclinical significance. Accordingly, patient 1136 shows a more severe phenotype than his affected father and brother.

In family 17, the SNF proband patient N4, who presents the *NF1:*c.3314+2T>C splicing variant inherited from her MNFSR affected mother (N5) shows a second *NF1:*c.7595C>T (*p.Ala2532Val – rs148154172*) missense variant, with uncertain clinical significance, predicted to be damaging by 10/20 predictors (Annovar), inherited from her father (never clinically evaluated) and shared by her brother displaying a cutaneous NF1 form. Also in this case the second mutation could have a subclinical effect that could worsen the clinical phenotype in the proband carrying mutations on both *NF1* alleles.

In family 18, the SNF proband N8 shows the pathogenic c.1595T>G *(p.Leu532Arg)* and the c.3242C>G *(p.Ala1081Gly) NF1* missense variant, with uncertain clinical significance, shared by his sister patient N9 and inherited by his nephew patient N11, indicating that the two variants are

in cis. The evidence that both his sister and her child are affected by classical NF1 suggests that this double mutated allele is not specifically associated with a specific NF form.

Two sporadic SNF patients are carriers of two variants in the *NF1* gene, but we were not able to define their inheritance was *in cis* or *in trans*, because their parents were not available.

Precisely, the patient 2207 presents a pathogenic stop gain mutation *NF1*: c.1246 C>T *(p.Arg4016*)* and the missense variant and c.403C >T *(p.Arg135Trp)*. Even if the second variant was reported as "uncertain" in Clinvar, this mutation has been classified as potentially damaging by 19 predictors out of the 20 interrogated by Annovar and is absent from controls in the GnomAD and in the 1000 genomes (1000g2015aug_eur) databases. Moreover, the mutation replaces the conserved basic amino acid arginine at residue 135 to polar-neutral tryptophan.

The patient 891 has a pathogenic frameshift *NF1* mutation c.6346_6347insA *(p.Ser2116Tyr*6)* and a second missense variant *NF1* c. 5221 G >A *(p.Val1741Ile)*, classified as potentially damaging by 8 out of 20 predictors questioned (Annovar) and not reported either in the GnomAD or in the 1000 genomes databases.

Table 16a *.NF1* mutations in 55 SNF sporadic patients

N°	ID Code	Age	Sex	DNA change	RNA change	Protein change	Type	Truncating	Exon/ intron	Tertile
1	368	30	M	c.31C>T	*r.(?)*	*p.(Gln11*)*	NS	yes	1	1
2	367	28	M	c.58C>T	*r.[58c>u, 57_60del4]*	*p.[Gln20Glufs*16, Gln20*]*	NS/SS	yes	1	1
3	1185/35	33	M	c.288+1137 C>T	*r.288_289ins288+ 1019_288+1136ins118*	*p.Gly96_Glu97ins39+fs*10*	SS	yes	IVS 3	1
4	1547	45	F	c.288+1del G	*r.288delg*	*p.Gln97Asnfs*6*	SS	yes	IVS 3	1
5	1069	42	M	c.586+2T>G	*r.480_586del107*	*p.Leu161Asnfs*4*	SS	yes	5	1
6	1304	45	M	c.61-?_586+?del	*r.(?)*	*p.(Leu21Lysfs*9)*	LD	yes	2-3-4-5	1
7	1638	35	M	c.730+4A>G	*r.655_730del76*	*p.Ala219Asnfs*37*	SS	yes	IVS 7	1
8	51 B	25	M	c.801delG	*r.(?)*	*p.(Trp267Cysfs*14)*	FS	yes	8	1
9	741	38	M	c.945_946delGCinsAA	*r.889_1062del174*	*p.Lys297_Lys354del*	DEL-INS	no, in-frame	9	1
10	2207	23	F	c.1246C>T c.403C>T	*r.1246c>u r.(?)*	*p.Arg416* p.(Arg135Trp)*	NS MS	yes no	11 4	1 1
11	425	57	M	c.1318C>T	*r.1318c>u*	*p.Arg440**	NS	yes	12	1
12	2171/2213	35	F	c.1711T>A	*r.(?)*	*p.(Trp571Arg)*	MS	no	15	1
13	692	31	M	c.1885G>A	*r.1846_1886del41*	*p.Gln616fs*4*	SS	yes	17	1
14	2146	48	M	c.2033_2034dupC	*r.2033dupc*	*p.Ile679Aspfs*21*	FS	yes	18	1
15	1708	16	M	c.2252G>T	*r.(?)*	*p.(Gly751Val)*	MS	no	19	1
16	1386	25	M	c.2326-3T>G	*r.2326_2409del84*	*p.Ala776_803Glndel*	SS	no, in-frame	IVS 19	1
17	1493	37	M	c.2446C>T	*r.2446c>u*	*p.Arg816**	NS	yes	21	1
18	1498/61	29	F	c.2509T>C	*r.2509u>c*	*p.Trp837Arg*	MS	no	21	1
19	1367	52	M	c.2810T>A	*r.2810u>a*	*p.Leu937**	NS	yes	21	1
20	509	54	M	c.3737_3740delTGTT	*r.(?)*	*p.(Phe1247fs*18)*	FS	yes	28	2
21	834	37	M	c.3827G>A	*r.3827g>a*	*p.Arg1276Gln*	MS	no	28	2

N°	ID Code	Age	Sex	DNA change	RNA change	Protein change	Type	Truncating	Exon/ intron	Tertile
22	584	57	M	c.3827G>C	*r.(?)*	*p.(Arg1276Pro)*	MS	no	28	2
23	1145	42	M	c.3888T>G	*r.3888u>g*	*p.Tyr1296**	NS	yes	29	2
24	1099	45	M	c.4267A>G	*r.4267a>g*	*p.Lys1423Glu*	MS	no	31	2
25	268	46	M	c.4480C>T	*r.(?)*	*p.(Gln1494*)*	NS	yes	33	2
26	1382	28	F	c.4719_4720dup AC	*r.4719_4720dupac*	*p.Gln1574Thrfs*30*	FS	yes	35	2
27	1430	35	M	c.4773-2A>C	*r.4773_5065del293*	*p.Phe1592Leufs*7*	SS	yes	IVS 35	2
28	918	28	F	c.4973_4978delTCTATA	*r.4973_4978delucuaua*	*p.Ile1658_Tyr1659del*	DEL	no, in-frame	36	2
29	1521/39	21	M	c.5199delT	*r.5199delu*	*p.Ile1734Leufs*10*	FS	yes	36	2
30	1263	37	F	c.5615dupT	*r.5615dupu*	*p.Glu1873Argfs*19*	FS	yes	38	2
31	1803	24	M	c.5630delT	*r.(?)*	*p.(Leu1877Tyrfs*27)*	FS	yes	38	2
32	1450	20	M	c.5704 A>C	*r.5704a>c*	*p.Thr1902Pro*	MS	no	38	2
33	1242	27	M	c.5923delA	*r.5923dela*	*p.Ile1975Tyrfs*16*	FS	yes	39	3
34	319	25	M	c.5943G>T	*r.(?)*	*p.(Gln1981His)*	MS	yes	39	3
35	1478	41	M	c.5943+1G>A	*r.5901_5943del43*	*p.Met1967Ilefs*9*	SS	yes	IVS 39	3
36	197	53	F	c.6084G>C	*r.(?)*	*p.(Lys2028Asn)*	MS	no	40	3
37	334	30	M	c.6085-2A>G	*r.6085_6364del280*	*p.Val2029Lysfs*7*	SS	yes	IVS 40	3
38	1573	69	M	c.6085G>T	*r.6085_6364del280*	*p.Val2029Lysfs*7*	SS	yes	41	3
39	2281	42	M	c.6088_6090delAATinsCTTTACA	*r.6088_6090delauuinscuuuaca*	*p.Ile2030Leufs*10*	FS	yes	41	3
40	571	36	F	c.6311T>C	*r.6311u>c*	*p.Leu2104Pro*	MS	no	41	3
41	891	23	M	c.6346_6347insA c.5221G>A	*r.(?)* *r.(?)*	*p.(Ser2116Tyrfs*6)* *p.(Val1741Ile)*	FS MS	yes no	41 38	3 2
42	531	43	M	c.6364+2T>A	*r.spl*	*p.(?)*	SS	yes	IVS 41	3
43	7	55	F	c.6688delG	*r.(?)*	*p.(Val2230Serfs*14)*	FS	yes	44	3

N°	ID Code	Age	Sex	DNA change	RNA change	Protein change	Type	Truncating	Exon/ intron	Tertile
44	829	40	M	c.6791dupA	*r.6791dupa*	*p.Tyr2264**	FS	yes	45	3
45	1877	15	F	c.6789_6792delTTAC	*r.6789_6792deluuac*	*p.Tyr2264Thrfs*5*	FS	yes	45	3
46	65	31	F	c.7846C>T	*r.7846c>u*	*p.Arg2616**	NS	yes	54	3
47	NF 220	22	F	c.8051-1G>C	*r.(?)*	*p.(?)*	SS	?	IVS 55	3
48	981	49	M	c.7127-?_8314+?del	*r.(?)*	*p.(?)*	LD	no, in-frame	49-57	3
49	NF 291	23	M	c.-718-?_8375+?del	*?*	*?*	LD	?	/	/
50	607	25	M	/	/	/	LD	/	/	1-2-3
51	M.E.	34	F	/	/	/	LD	/	/	1-2-3
52	1773	55	F	NEGATIVE	/	/	/	/	/	/
53	1357	26	M	NEGATIVE	/	/	/	/	/	/
54	1390	58	F	NEGATIVE	/	/	/	/	/	/
55	1085	23	M	NEGATIVE	/	/	/	/	/	/

NS= non-stop mutation; SS= splicing mutation; LD= large deletion; MS= missense mutation; FS= frameshift mutation,; DEL= deletion; INS= insertion

				Table 16b. *NF1* mutations in 106 Classical sporadic patients						
N°	**ID Code**	**Age**	**Sex**	**DNA change**	**RNA change**	**Protein change**	**Type**	**Truncating**	**Exon/ intron**	**Tertile**
1	623	39	F	c.200dupA	*r.200dupa*	*p.Asn67Lysfs*10*	FS	yes	2	1
2	1318	52	F	c.204+2T>G	*r.100_204del105*	*p.Val34_Met68*	SS	no inframe	IVS 2	1
3	136	71	F	c.288+1delG	*r.288_288delg*	*p.Gln97Asnfs*6*	SS	yes	IVS 3	1
4	767	38	M	c.493delA	*r.493dela*	*p.Thr165Leufs*13*	FS	yes	5	1
5	323	46	F	c.499_502delTGTT	*r.499_502deluguu*	*p.Cys167Glnfs*10*	FS	yes	5	1
6	412	32	F	c.499_502delTGTT	*r.499_502deluguu*	*p.Cys167Glnfs*10*	FS	yes	5	1
7	1455	34	F	c.499_502delTGTT	*r.499_502deluguu*	*p.Cys167Glnfs*10*	FS	yes	5	1
8	738	57	F	c.574C>T	*r.574c>u*	*p.Arg192**	NS	yes	5	1
9	1490	45	M	c.574C>T	*r.574c>u*	*p.Arg192**	NS	yes	5	1
10	384	39	F	c.652_653delAAinsG	*r.(?)*	*p.(Lys218Glyfs*7)*	FS	yes	6	1
11	858	62	F	c.653delA	*r.653delA*	*p.Lys218Argfs*7*	FS	yes	6	1
12	1967	57	M	c.725delT	*r.725delu*	*p.Met242Argfs*39*	FS	yes	7	1
13	1781/1888	50	M	c.587-?_888+?dup	*r.(?)*	*p.(?)*	DUP/ FS	?	6-7-8	1
13	1435	40	M	c.908T>C	*r.908u>c*	*p.Leu303Pro*	MS	no	9	1
14	765	60	F	c.910C>T	*r.910c>u*	*p.Arg304**	NS	yes	9	1
15	329/1020	53	F	c.932_933delG	*r.932_933delg*	*p.Gly311Glufs*6*	FS	yes	9	1
16	1504	56	F	c.943C>T	*r.943c>u*	*p.Glu315**	NS	yes	9	1
17	501	66	M	c.1019_1020delCT	*r.1019_1020delcu*	*p.Ser340Cysfs*12*	FS	yes	9	1
18	752	43	F	c.1019_1020delCT	*r.1019_1020delcu*	*p.Ser340Cysfs*12*	FS	yes	9	1
19	1542	37	F	c.1019_1020delCT	*r.1019_1020delcu*	*p.Ser340Cysfs*12*	FS	yes	9	1
20	1488	52	F	c.1185+2delT	*r.1063_1185del123*	*p.Asn355_Lys395del*	SS	no inframe	IVS 10	1
21	356	62	M	c.1318C>T	*r.1318c>u*	*p.Arg440**	NS	yes	12	1
22	199	49	F	c.1466A>G	*r.1466_1527del62*	*p.Tyr489**	SS	yes	13	1

N°	ID Code	Age	Sex	DNA change	RNA change	Protein change	Type	Truncating	Exon/ intron	Tertile
23	1548	43	F	c.1466A>G	*r.1466_1527del62*	*p.Tyr489**	SS	yes	13	1
24	1669	47	F	c.1466A>G	*r.1466_1527del62*	*p.Tyr489**	SS	yes	13	1
25	1328	41	F	c.1541_1542delAG	*r.1541_1542delag*	*p.Gln514Argfs*43*	FS	yes	14	1
26	915	73	M	c.1658A>G	*r.1658A>G*	*p.His553Arg*	MS	no	15	1
27	2019	54	M	c.1907_1908delCT	*r.1907_1908delcu*	*p.Ser636**	FS	yes	17	1
28	1577	40	F	c.1925_1931delAAATGTC	*r.1925_1931delaaauguc*	*p.Gln642Profs*44*	FS	yes	17	1
29	663	35	M	c.2033dupC	*r.2033dup*	*p.Ile679Aspfs*21*	FS	yes	18	1
30	1238	66	F	c.2033dupC	*r.2033dup*	*p.Ile679Aspfs*21*	FS	yes	18	1
31	1601	46	F	c.2041C>T	*r.2041c>u*	*p.Arg681**	NS	yes	18	1
32	860	63	M	c.2041C>T	*r.2041c>u*	*p.Arg681**	NS	yes	18	1
33	1754	37	M	c.2076C>A	*r.(?)*	*p.(Tyr692*)*	NS	yes	18	1
34	524	54	M	c.2106delT	*r.2106delu*	*p.Val703Phefs*45*	FS	yes	18	1
35	489	44	M	c.2205T>G	*r.(?)*	*p.(Tyr735*)*	NS	yes	18	1
36	171	51	M	c.2326-1G>C	*r.2252_2325del74*	*p.Arg752Leufs*17*	SS	yes	IVS 19	1
37	558	30	F	c.2356delC	*r.2356delc*	*p.Gln786Lysfs*5*	FS	yes	20	1
38	507	45	F	c.2492_2493dupCA	*r.2492_2493dup*	*p.Asp832Glnfs*10*	FS	yes	21	1
39	290	73	M	c.2540T>C	*r.2540u>c*	*p.Leu847Pro*	MS	no	21	1
40	1590	37	F	c.2546_2546delG	*r.2546_2546delg*	*p.Gly849Glufs*29*	FS	yes	21	1
41	53	42	F	c.2850+1G>T	*r.2618_2850del*	*p.Lys874Phefs*4*	SS	yes	IVS 21	1
42	764	39	F	c.2851-2AT	*r.2851_2990del140*	*p.Leu952Cysfs*22*	SS	yes	IVS 21	1
43	1377	38	M	c.2953C>T	*r.2952_2990del39*	*p.Gly984_Arg997del*	SS	no inframe	22	2
44	1165	47	F	c.2991-2A>G	*r.2991_3113del123*	*p.Tyr998_Arg1038del*	SS	no inframe	IVS 22	2
45	2022	46	M	c.2991-2A>T	*r.2991_3113del123*	*p.Tyr998_Arg1038del*	SS	no inframe	IVS 22	2
46	459	45	F	c.3384_3390delTGGCAGG	*r.(?)*	*p.(Gly1129Asnfs*11)*	FS	yes	26	2
47	2111	60	F	c.3485delT	*r.(?)*	*p.(Met1162Serfs*4)*	FS	yes	26	2

N°	ID Code	Age	Sex	DNA change	RNA change	Protein change	Type	Truncating	Exon/ intron	Tertile
48	1566	44	M	c.3586C>T	*r.3586c>u*	*p.Leu1196Phe*	MS	no	27	2
49	1491	36	M	c.3644T>G	*r.3644u>g*	*p.Met1215Arg*	MS	no	27	2
50	73	46	F	c.3708+1G>C	*r.3497_3708del212*	*p.Leu1167**	SS	yes	IVS 27	2
51	876	75	F	c.3785delC	*r.3785delc*	*p.Ser1262Leufs*4*	FS	yes	28	2
52	1594	50	M	c.3826C>T	*r.3826c>u*	*p.Arg1276**	NS	yes	28	2
53	1420	46	M	c.3870+1G>C	*r.3845_3870del26*	*p.Lys1283fs*22*	SS	yes	IVS 28	2
54	744	40	M	c.3888T>G	*r.(?)*	*p.(Tyr1296*)*	NS	yes	29	2
55	1452	39	F	c.3892C>T	*r.3892c>u*	*p.Gln1298**	NS	yes	29	2
56	700	51	F	c.3916C>T	*r.3916c>u*	*p.Arg1306**	NS	yes	29	2
57	809	32	F	c.3916C>T	*r.3916c>u*	*p.Arg1306**	NS	yes	29	2
58	705	58	F	c.3941G>A	*r.3941g>a*	*p.Trp1314**	NS	yes	29	2
59	1320	39	F	c.3975-1G>A	*r.3975_3959delguuag*	*p.Arg1325Asnfs*16*	SS	yes	IVS 29	2
60	1194	58	M	c.4077delT	*r.4077delu*	*p.Gln1360Asnfs*25*	FS	yes	30	2
61	134	45	M	c.4084C>T	*r.4084c>u*	*p.Arg1362**	NS	yes	30	2
62	2018	54	M	c.4269+1G>C	*r.4111_4269del159*	*p.Val1371_Lys1423del*	SS	no inframe	IVS 31	2
63	1984	51	F	c.4368-1G>T	*r.4368_4384del17*	*p.Arg1456Serfs*3*	SS	yes	IVS 32	2
64	733	35	F	c.4402_4406delAGTGA	*r.4402_4406delaguga*	*p.Ser1468Cysfs*5*	FS	yes	33	2
65	965	47	F	c.4435A>G	*r.4368_4435del68*	*p.Phe1457**	SS	yes	33	2
66	936	61	M	c.4537C>T	*r.4537c>u*	*p.Arg1513**	NS	yes	34	2
67	1978	61	M	c.4537C>T	*r.4537c>u*	*p.Arg1513**	NS	yes	34	2
68	170	47	F	c.4538C>T	*r.(?)*	*p.(Arg1513*)*	NS	yes	34	2
69	273	43	F	c.4630delA	*r.4630dela*	*p.Thr1544Profs*9*	FS	yes	34	2
70	1353	54	F	c.4637C>G	*r.4637c>g*	*p.Ser1546**	NS	yes	34	2
71	1428	48	F	c.4854T>A	*r.4854u>a*	*p.Tyr1618**	NS	yes	36	2
72	1500	61	M	c.4917dupT	*r.4917dupu*	*p.Lys1640**	FS	yes	36	2
73	919	58	F	c.4973_4978delTCTATA	*r.4973_4978delucuaua*	*p.Ile1658Tyr1659del*	SS	no inframe	36	2

N°	ID Code	Age	Sex	DNA change	RNA change	Protein change	Type	Truncating	Exon/ intron	Tertile
74	1358	58	F	c.4981T>C	*r.4981u>c*	*p.Cys1661Arg*	MS	no	36	2
75	32	51	F	c.5154_5157(dupATCC)	*r.(?)*	*p.(His1720Ilefs*17)*	FS	yes	36	2
76	822	35	M	c.5242C>T	*r.5242c>u*	*p.Arg1748**	NS	yes	37	2
77	1749	43	F	c.5470A>T	*r.5470a>u*	*p.Ile1824Phe*	MS	no	37	2
78	1213	52	F	c.5495C>G	*r.5495c>g*	*p.Thr1832Arg*	MS	no	37	2
79	966	50	M	c.5513_5514delTA	*r.5513_5514del*	*p.Leu1838Serfs*2*	FS	yes	37	2
80	213	32	F	c.5546G>A	*r.5206_5546del341*	*p.Gly1737Serfs*4*	SS	yes	37	2
81	1214	43	M	c.5546G>A	*r.5206_5546del341*	*p.Gly1737Serfs*4*	SS	yes	37	2
82	502	40	F	c.5546+5G>C	*r.[5206_5546del341, 5206_5749del544]*	*p.[Gly1737Serfs*4, Gly1737Leufs*3]*	SS	yes	IVS 37	2
83	474	52	F	c.5750-177A>C	*r.5749_5750ins5750-174_5750-108*	*p.Ser1917Argfs*25*	SS	yes	IVS 38	2
84	620	47	F	c.5839C>T	*r.5839c>u*	*p.Arg1947**	NS	yes	39	3
85	946	56	F	c.5839C>T	*r.5839c>u*	*p.Arg1947**	NS	yes	39	3
86	49	38	F	c.5890G>T	*r.(?)*	*p.(Glu1964*)*	NS	yes	39	3
87	1169	34	F	c.6084+1G>A	*r.5944_6084del141*	*p.Ile1982_Lys2028del*	SS	no inframe	IVS 40	3
88	541	60	F	c.6641+1G>A	*r.6580_6641del62*	*p.Ala2194fs*	SS	yes	IVS 43	3
89	133	60	F	c.6709C>T	*r.6709c>u*	*p.Arg2237**	NS	yes	44	3
90	1777	48	M	c.6709C>T	*r.6709c>u*	*p.Arg2237**	NS	yes	44	3
91	603	39	F	c.6760delC	*r.6760delc*	*p.Glu2255Argfs*4*	FS	yes	45	3
92	1327	61	M	c.6789_6792delTTAC	*r.(?)*	*p.(Tyr2264Glnfs*5)*	FS	yes	45	3
93	183	50	F	c.6789_6792delTTAC	*r.6789_6792deluuac*	*p.Thr2264fs*	FS	yes	45	3
94	29	47	F	c.6999+1G>C	*r.spl*	*p.(?)*	SS	yes	IVS 46	3
95	1467	50	F	c.7151_7161delTTGTTGCAAGA	*r.7151_7161deluuguugcaaga*	*p.Ile2384Asnfs*13*	FS	yes	48	3

N°	ID Code	Age	Sex	DNA change	RNA change	Protein change	Type	Truncating	Exon/ intron	Tertile
96	702	39	F	c.7422dupC	*r.7422dupc*	*p.Ser2475Leufs*6*	FS	yes	50	3
97	1533	43	M	c.7486C>T	*r.7486c>u*	*p.Arg2496**	NS	yes	50	3
98	1608	42	M	c.7500delC	*r.7500delc*	*p.Met2501**	FS	yes	50	3
99	846	51	F	c.7926_7929delTAAG	*r.7926_7929deluaag*	*p.Lys2643Serfs*14*	FS	yes	54	3
100	870	31	M	c.1007G>A	*r.1007g>a*	*p.Trp336**	NS	yes	9	1
101	727	65	F	/	/	/	LD	/	/	1-2-3
102	621	32	F	/	/	/	LD	/	/	1-2-3
103	/	42	M	/	/	/	LD	/	/	1-2-3
104	/	44	F	/	/	/	LD	/	/	1-2-3
105	/	46	M	/	/	/	LD	/	/	1-2-3
106	/	48	F	/	/	/	LD	/	/	1-2-3

NS= non-stop mutation; SS= splicing mutation; LD= large deletion; MS= missense mutation; FS= frameshift mutation,; DEL= deletion; INS= insertion

Table 16c. *NF1* mutations in SNF probands and relatives

Family	ID Code	Age	Sex	Subject	Phenotype	DNA change	RNA change	Protein change	Type	Truncating	Exon/ intron	Tertile
1	1136	38	M	Proband*§	SNF	c.62T>A* c.528T>A§	*r.62u>a r.528u>a*	*p.Leu21His p.Asp176Glu*	MS MS	no no	2 5	1 1
	1140	37	M	Brother*	SNF							
	1139	63	M	Father*	MNFSR							
2	451	9	M	Proband	SNF	c.1393-?_2325+?del	*r.(?)*	*p.(Ser465_Glu775del)*	LD	no, in-frame	13-19	1
	494	54	F	Mother	Classical							
3	1153/ 176	22	M	Proband	SNF	c.6364+1G>A	*r.6085_6364del280*	*p.Val2029Lysfs*7*	SS	yes	IVS 41	3
	1202	38	F	Sister	MNFSR							
	1228	57	F	Mother	Classical							
4	926	38	M	Brother	MNFSR	c.2329T>A	*r.2329u>a*	*p.Trp777Arg*	MS	no	20	1
	258	42	M	Proband	SNF							
5	46 B	19	F	Proband	SNF	c.5543T>A	*r.(?)*	*p.(Leu1748*)/ p.(Leu1848*)*	NS	yes	38	3
	47 B	55	F	Mother	Classical							
6	/	36	M	Proband	SNF	c.2297T>G	*r.(?)*	*p.(Ile766Ser)*	MS	no	19	1
	/	71	M	Father	MNFSR							
7	/	45	F	Proband	SNF	c.7126+3A>T	*r.(?)*	*p.(Gly2334fs*14)*	SS	yes	IVS 47	3
	/	70	F	Mother	SNF							
	/	74	F	Aunt	SNF							
	/	46	M	Cousin	MNFSR							
8	1434	23	M	Proband	SNF	c.7079dupA	*r.7079dupa*	*p.Asp2360Lysfs*5*	FS	yes	47	3
	1436	57	F	Mother	Classical							
9	1931	28	M	Proband	SNF	c.7395-?_7552+?del	*r.7395_7552del158*	*p.Thr2466Asnfs*6*	LD	yes	51	3
	1813/ 1912	23	M	Brother	SNF							
	1814	55	F	Mother	Classical							
10	1271	28	M	Proband	SNF	*c.3827G>A*	*r.3827g>a*	*p.Arg1276Gln*	MS	no	22	2
	1276	71	M	Father	MNFSR							

Family	ID Code	Age	Sex	Subject	Phenotype	DNA change	RNA change	Proteinchange	Type	Truncating	Exon/ intron	Tertile
11	1957	39	M	Proband	SNF	c.5546G>A	*r.5206_5546del341*	*p.Gly1737Serfs*4*	SS	yes	37	2
	1065	44	F	Sister	MNFSR							
	1660	73	F	Mother	Classical							
12	550	31	M	Proband	SNF	c.6791dupA	*r.6791dupa*	*p.Tyr2264**	FS	yes	45	3
	277	37	F	Sister	SNF							
13	1649	43	M	Proband	SNF	c.2523_2524insT	*r.2523_2524insu*	*p.Gly842Trpfs*23*	FS	yes	21	1
	1650	69	M	Father	Classical							
14	1086/ 2277	45	M	Proband	SNF	c.6085-2A>C	*r.6085_6364del280*	*p.Val2029Lysfs*7*	SS	yes	IVS 40	3
	2198	49	F	Sister	Classical							
15	392	20	M	Proband	SNF	c.1381C>T	*r.1381c>u*	*p.Arg461**	NS	yes	12	1
	2191	48	F	Mother	MNFSR							
16	/	49	F	Proband	SNF	c.1527+5G>T	*r.(?)*	*p.(?)*	SS	?	IVS 13	1
	/	52	F	Sister	MNFSR							
17	N03	27	M	Brother*	SNF	c.3314+2T>C* c.7532C>T§	*r.spl r.(?)*	*p.(?)p.(Ala2511Val)*	SS MS	yes no	IVS 25 51	2 3
	N04	34	F	Proband*§	SNF							
	N05	57	F	Mother*	MNFSR							
	N06	40	M	Brother§	Classical							
18	N8	52	M	Proband*§	SNF	c.1595T>G * c.3242C>G§	*r.(?) r.(?)*	*p.(Leu532Arg) p.(Ala1081Gly)*	MS MS	no no	14 25	1 2
	N9	44	F	Sister*§	Classical							
	N10	14	M	Nephew*§	Classical							
19	/	38	M	Brother	MNFSR	c.7881_7882del	*r.(?)*	*p.(Val2627fs*)*	FS	yes	57	3
	/	33	M	Proband	SNF							

NS= non-stop mutation; SS= splicing mutation; LD= large deletion; MS= missense mutation; FS= frameshift mutation,; DEL= deletion; INS= insertion

4.4.2 Higher prevalence of missense NF1 mutations in SNF

We performed a comparative analysis of *NF1* mutations in classical and spinal patients by considering in the SNF cohort 55 SNF unrelated sporadic patients and 19 SNF probands (for a total of 74 SNF patients), belonging to 19 unrelated SNF families, and compared this cohort with the classical cohort composed by 106 unrelated NF1 patients.

We studied the different *NF1* mutations (microdeletion, frameshift, missense, nonsense, splicing, small deletion /insertion) classes in the SNF and classical cohorts. We included in the analysis only the causative *NF1* mutations (Annovar annotation) and excluded the *NF1* variants with uncertain, benign and likely benign clinical significance (Annovar annotation).

The proportion of missense mutations was higher in the SNF cohort than in the classical NF1 group (p=0,001), while the proportion of nonsense was lower (p=0.03). After applying Benjamini-Hochberg (B-H) correction for multiple testing with a false discovery rate at 0.025 and 0.01, the first differences remained statistically significant, while the second only with a false discovery rate of 0.025 (Table 17).

Table 17. Distribution of the *NF1* mutation classes between SNF and classical groups

Mutation type	SNF n (%) (n= 70)	Classical NF n (%) (n=106)	P-value	OR (95% CI)	Total numbers
Large deletions	7 (10)	6 (5.7)	0.28	1.85 (0.59– 5.77)	13
Frameshift	17 (24.3)	38 (35.8)	0.10	0.57 (0.29 -1.13)	55
Missense	15 (21.4)	8 (7.5)	**0.007***	3.34 (1.33 -8.38)	28
Nonsense	10 (14.2)	28 (26.4)	0.055	0.46 (0.20-1.03)	38
Splicing	19 (27)	26 (24.5)	0.70	1.15 (0.58 -2.28)	45
Deletion/insertion	2 (2.9)	0	0.16	7.7 (0.97 - 16.44)	2

n = number of *NF1* mutations; 4 SNF patients were negative for *NF1* mutations

*Statistical significant P values, with false discovery rate of 0.05 after correction for multiple testing using Benjamin Hochberg procedure.

Because the SNF is a rare form of NF1 and relatively few SNF patients have been described, we carried out a combined analysis aimed at verifying the occurrence of specific classes of *NF1* mutations in SNF-described patients joining the data obtained in our SNF cohort. Taking in account that Ruggieri [24], by the application of rigorous criteria, diagnosed SNF to 49 patients out of the 98 described overall in the literature, we considered in this casuistry, only the mutations of the unrelated patients, reducing the described cohort from 49 to 25. The combined analysis with our data (Table 18) showed a statistically significant increase of missense mutations (25.3% vs. 7.5%, p value =0.001 (OR 4.14; CI= 1.76 -9.75) in the SNF cohort compared to our classical patient cohort with the p value remaining statistically significant also after correcting with Benjamini-Hochberg method for multiple tests, with a false discovery rate of 0.05, 0.025 and 0.01.

Table 18. Distribution of the *NF1* mutation types between SNF and classical groups including cases reported by Ruggieri et al.

Mutation type	SNF n (%) (n= 70)	SNF n (%) Ruggieri et al. (n=25)	total SNF n (%) (n=95)	Classical NF n (%) n=106	P-value	OR (95% CI)	Total numbers
Large deletions	7 (10)	1 (4)	8 (8.4)	6 (5.7)	0.44	1.53 (0.51- 4.59)	14
Frameshift	17(24.3)	3 (12)	20 (21)	38 (35.8)	**0.02**	0.47 (0.25- 0.89)	58
Missense	15(21.4)	9 (36)	24 (25.3)	8 (7.5)	**0.001** §	4.14 (1.76 -9.75)	32
Nonsense	10(14.2)	3 (12)	13 (13.7)	28 (26.4)	**0.025**	0.44 (0.21 - 0.91)	41
Splicing	19 (27)	6 (24)	25 (26.3)	26 (24.5)	0.77	1.09 (0.58 -2.07)	51
Deletion/insertion	2 (2.9)	3 (12)	5 (5.3)	0	0.023		5

n= number of *NF1* mutations

§Statistical significant P value, with false discovery rate of 0.05, 0.025 and 0.01 after correction for multiple testing using Benjamin Hochberg.

4.4.3 Higher prevalence of 3' NF1 tertile mutations in SNF as compared to classical cohort

As previously described, there is a higher prevalence of missense *NF1* mutations in SNF as compared to the classical NF1 cohort. The distribution of mutations seems to be not random. If we divide the *NF1* gene into tertiles, as described by Sharif and colleagues [61], the 5'tertile, corresponds to the exons 1-21, the middle tertile to the exons 22-38, and the 3' tertile to the exons 39-57 and we can determine the corresponding distribution of mutations in classical and spinal patients. The prevalence of mutations occurring in the 3' tertile of the NF1 gene is significantly higher in patients with SNF (34.3%) than that observed in the classical ones (16%) (p=0.006 (OR 2.277; CI = 1.31–5.7), while the prevalence of those harboring in the middle tertile was lower (p=0.038 (OR 0.49; CI = 0.25–0.96) (Table 19).

Table 19. Comparison of *NF1* mutations based on their position along *NF1* gene

Tertile	SNF n (%) n= 67	Classical NF n (%) n=100	P-value	OR (95% CI)	Total numbers
5′ tertile(1-21)	27 (40.3)	43 (43)	0.73	0.89 (0.48 - 1.68)	70
Middle tertile (22-38)	17 (25.4)	41(41)	**0.038**	0.49 (0.25 – 0.96)	58
3′ tertile (39–57)	23 (34.3)	16 (16)	**0.006#**	2.77 (1.31 - 5.72)	39

#Statistical significant P values with false discovery rate of 0.05 and 0.025 after correction for multiple testing using Benjamin Hochberg procedure. n= number of *NF1* mutations. The deletions spanning more than one tertile were excluded from the analyses.

The combined analysis of *NF1* variants distribution in tertiles, by adding the data from Ruggieri et al. [24], as previously described, confirmed the higher prevalence of *NF1* mutations in the 3' tertile in SNF as compared to classical NF1, patients with an increasing of statistical significance (P=0.0016, with false discovery rate of 0.05, 0.025 and 0.01 after correction for multiple testing using Benjamin Hochberg procedure).

4.4.4 Distribution of different variant classes within the NF1 tertiles

The prevalence in 5' tertile of frameshift and missense mutations differs significantly, also after B-H correction (p=0.00619 and p= 0.0045 respectively), between classical NF1 and SNF patients (Figure 12a). The distribution in the middle tertile showed no statistical differences between the two groups of patients (Figure 12b). The prevalence of 3' tertile *NF1* stop gain mutations was lower in SNF (Figure 12c) as compared to NF1 classical patients (p= 0.0127) and remain significant after B-H correction. In the 3' tertile we found no missense mutations in the classical patients (0%) as compared to SNF (5%). For this reason, we could not apply the Chi square or Fisher exact test.

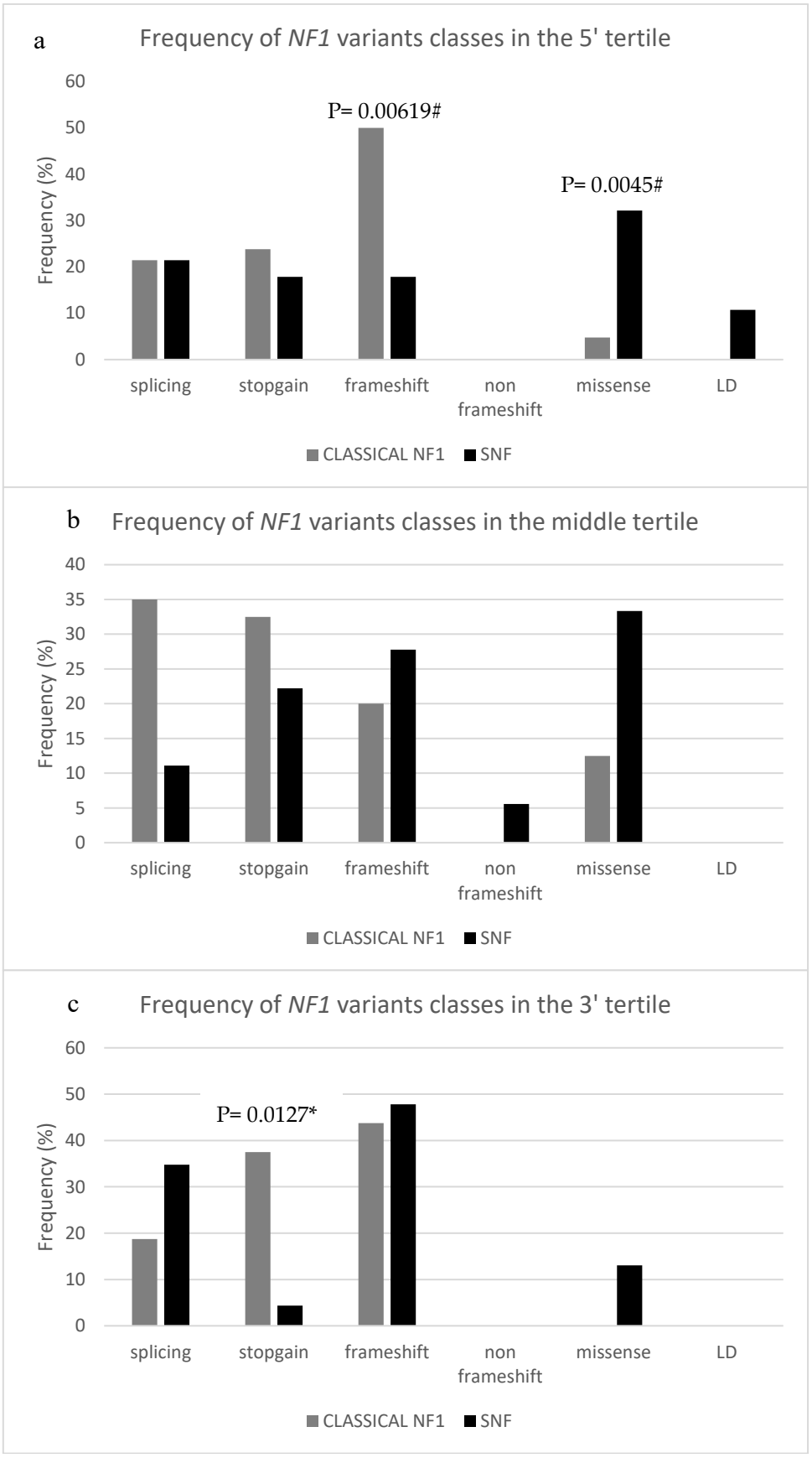

a
Frequency of NF1 variants classes in the 5' tertile
Frequency (%)
60
50
40
30
20
10
0
P= 0.00619#
P= 0.0045#
splicing
stopgain
frameshift
non frameshift
missense
LD
CLASSICAL NF1
SNF
b
Frequency of NF1 variants classes in the middle tertile
Frequency (%)
40
35
30
25
20
15
10
5
0
splicing
stopgain
frameshift
non frameshift
missense
LD
CLASSICAL NF1
SNF
c
Frequency of NF1 variants classes in the 3' tertile
Frequency (%)
60
50
40
30
20
10
0
P= 0.0127*
splicing
stopgain
frameshift
non frameshift
missense
LD
CLASSICAL NF1
SNF

Figure 12. Distribution of splicing, stopgain, frameshift insertion-deletion, non-frameshift insertion-deletion, missense, large deletion (LD) and of *NF1* gene mutations in the 5' (a), middle (b) and 3' (c) tertile of the *NF1* gene. Statistically significant P values obtained by Fisher exact test or Chi square and after correction for multiple testing using Benjamin Hochberg procedure are showed above the bars. #Significant with a FDR of 0,05 and 0,025 after B-H correction for multiple tests; ***** significant with a FDR of 0,05 after B-H correction for multiple tests; **§** significant with a FDR of 0,05 and 0,025 and 0.01 after B-H correction for multiple tests

4.4.5 Variants of neurofibromin interactors in SNF and classical NF1 patients

We hypothesized a functional significance of the prevalence of 3' tertile *NF1* mutations in SNF. Accordingly, we verified the presence of variants in syndecans, for which the neurofibromin binding domain, SBR, is known, in SNF in comparison to NF1 patients. These interactors are the genes belonging to the syndecan family: *SDC1, SDC2, SDC3,* and *SDC4*. We searched for rare variants with MAF <0.01 by means of ANNOVAR annotation. Six variants in the four genes encoding syndecans were identified in 5 SNF and 1 classical NF1 patient (Table 20). We assessed the clinical significance according to the ACMG/AMP criteria, which led to their classification into two groups (Table S1, Supplementary materials): 1 "Uncertain" (4/6, 67%), when evidence was not sufficient to draw definitive conclusions on pathogenicity, including c.215C>T *(p.Thr72Met)* in *SDC1, c.*923C>T *(p.Pro308Leu)* and c. 721A>G *(p.Thr241Ala)* in *SDC3*, c. A92G *(p.Asp31Gly)* in *SDC4*. 4. "Likely pathogenic" (2/6, 33%), when evidence supporting pathogenicity were concordant among several different *in silico* predictors although at least one major pathogenicity criterium, such as either detection in other patients with similar phenotypes or variant functional validation, was still missing. This group includes the variants c.830G>A *(p.Arg277His)* and c.449 T>C *(p.Ile150Thr)* affecting the *SDC1* and *SDC2* genes, respectively. The 830G>A *(p.Arg277His)* in *SDC1* has never been reported in gnomAD v.3.1.1 and 1000 Genomes database and both the variants are in the cytoplasmic syndecan domain and predicted as damaging by most predictors (18 out of 20).

All the syndecan variants but one, were co-present in patients carrying mutations of the 5 'or middle *NF1*tertile. Two out of the six variants were predicted as damaging from most predictors interrogated by Annovar (Dampred= 18,2) (Table S1, section "6. Supplementary").

Table 20. Syndecans variants identified in *NF1* mutated patients

Patient phenotype (patient ID)	**SyndecanVariant**	***NF1* mutation tertile**	**DamPred**	**Clinical significance**
SNF (NF220)	SDC1(NM_002997.5):c.830G>A *(p.Arg277His)*	3’	18.2	Likely pathogenic
Classical NF1 (1358)	SDC1(NM_002997.5):c.215C>T *(p.Thr72Met)*	middle	7.2	uncertain
SNF (1271)	SDC2(NM_002998.4:c.449T>C *(p.Ile150Thr)*	5’	18.2	Likely pathogenic
SNF (1803)	SDC3(NM_014654.4)*c.*923C>T *(p.Pro308Leu)*	middle	11.2	uncertain
SNF (425)	SDC3(NM_014654.4):c.721A>G *(p.Thr241Ala)*	5’	4.2	uncertain
SNF (258)	SDC4(NM_002999.4):c.A92G *(p.Asp31Gly)*	5’	2.2	uncertain

Dampred= Damage prediction score calculated by Annovar

We searched also for variants in *NF1* 5’ and middle tertile interactors, applying the above-described pipeline and found 4 rare variants. By applying the above-described criteria, we classified the 4 variants into two groups (Table S2, section "6. Supplementary"): 1 “Uncertain”, including one variant in *SPRED1*, a GRD domain interactor, found in one classical patient.

2. “Likely pathogenic” including three variants in the *APP* gene, a GRD domain interactor, in two classical (1214 and 1165) and in one SNF patient (1085). Interestingly, The SNF patient does not carry *NF1* mutations.

4.5 Syndecan transcripts expression in SNF and in classical NF1

To investigate the expression of syndecans genes in SNF and NF1 patients, we performed qPCR on RNA peripheral blood samples from 16 SNF patients, 16 patients with classical NF1, and 16 healthy controls. For qPCR assays we selected the genes of *SDC2, SDC3* and *SDC4*, with an expression level in peripheral blood greater than 0.5 TPM (transcripts per million). The SDC1 gene didn't result expressed in SNF and NF1 patients as expected, according to the GTex reported data for peripheral blood, being <0.5 TPM. The average value of the quantitative expression levels ($2^{-\Delta Ct}$) of SDC2 and SDC3 were significantly higher in SNF and NF1 patients compared to controls, after application of the Student's t-test and the B-H correction for multiple tests (Table S3, section "6. Supplementary", and Figure 13 a and b). The average value of the quantitative expression levels ($2^{-\Delta Ct}$) of *SDC4* was found to be significantly higher in classical NF1 patients compared to controls (Table S3 section "6. Supplementary", and Figure 13 c).

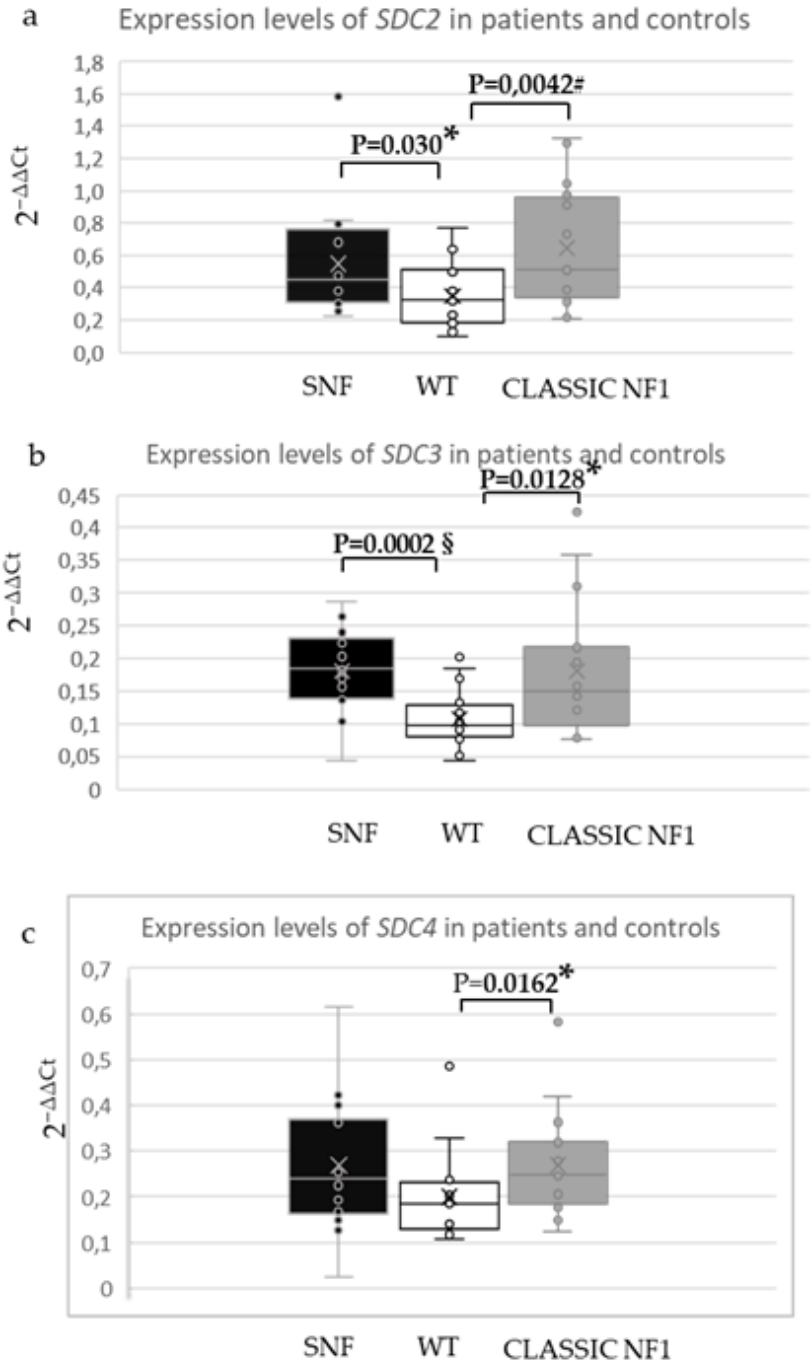


Figure 13. The box plots show the dispersion and the quantitative expression levels of the gene expression values ($2^{-\Delta Ct}$) analyzed by qPCR of the syndecans genes *SDC2, SDC3* and *SDC4* in peripheral blood from 16 patients with SNF (shown in black), 16 controls (WT, shown in white) and 16 patients with classical NF1 (shown in grey). *SDC2, SDC3* and *SDC4* were statistically significantly hyper-expressed, even after B-H correction for multiple tests, in SNF and classical NF1 patients as compared to controls, Student's t-test. The boxes represent the 25th and 75th percentiles. The whiskers show the minimum and maximum value of the distribution, excluding the outliers. The big horizontal lines represent the median value. The outliers are represented as spots outside of the boxes and excluded from the Student's t-test analysis. Statistically significant P values obtained by Student's t-test are showed above the bars.

significant with a FDR of 0.05 and 0.025 after B-H correction for multiple tests

* significant with a FDR of 0.05 after B-H correction for multiple tests

§ significant with a FDR of 0.05, 0.025 and 0.01 after B-H correction for multiple tests

As both SNF and classical NF1 patients showed increased levels of syndecans compared to controls, we verify if this hyperexpression was possibly associated with the presence of *NF1* mutations in a specific *NF1*tertile, without distinguishing between the specific NF1 form. We compared the average value of the quantitative expression levels ($2^{-\Delta Ct}$) of the *SDC2, SDC3* and *SDC4* in 39 NF1 patients (to the 32 previously analyzed patients we added 7 NF1 patients to enlarge the casuistry) subdivided in the three subgroups according to the specific *NF1* tertile.
SDC2, *SDC3* and *SDC4* were significantly hyper-expressed in patients with *NF1* mutations in the 3' tertile as compared to controls (Figure 14). Moreover, the *SDC2* and *SDC4* genes were significantly hyper-expressed in patients with *NF1* mutations in the 3' tertile as compared to patients with *NF1* mutations in the middle tertile. (Table S4, section "6. Supplementary" and Figure 14 a and c). These data assess that the presence of *NF1* mutations in the 3' tertile, including the SBR domain, is associated with the increasing of the *SDC2* and *SDC4* expression.

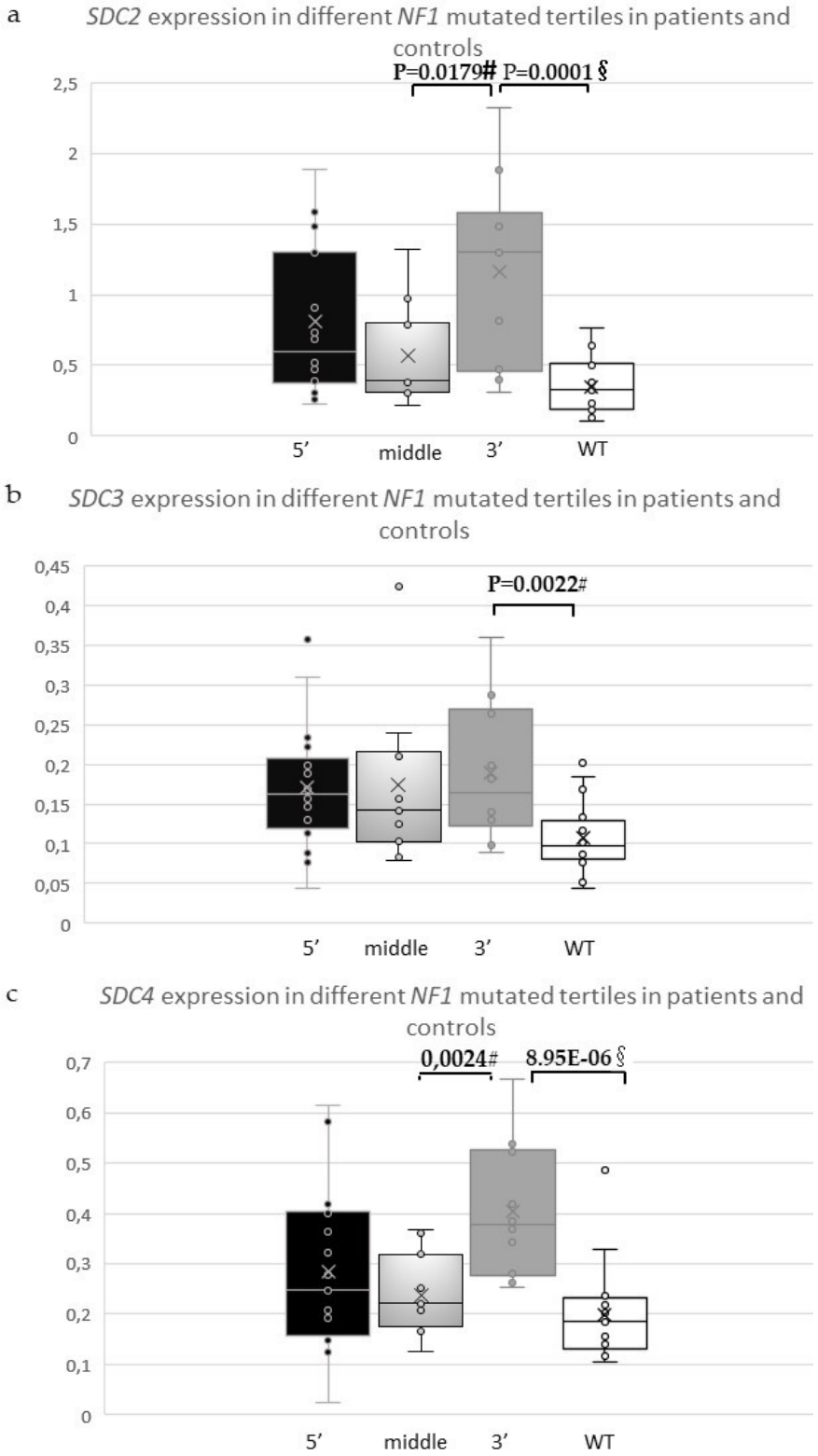


Figure 14. Syndecan expression in different *NF1* mutated tertiles in patients and controls The box plots show the dispersion and the quantitative expression levels of the gene expression values ($2^{-\Delta Ct}$) analyzed by qPCR of the syndecans genes *SDC2* (a), *SDC3* (b)and *SDC4* (c) in peripheral blood from 18 patients with *NF1* mutations in the 5' tertile (5', shown in black), 11 patients with *NF1* mutations in the middle tertile (middle, shown in light grey), 10 patients with *NF1* mutations in the 3' tertile (3', shown in

grey) and 16 healthy controls (WT, shown in white). *SDC2, SDC3* and *SDC4* were statistically significantly hyper-expressed, even after B-H correction for multiple tests, in patients with *NF1* mutations of the 3' tertile as compared to controls. *SDC2* and *SDC4* were statistically significantly hyper-expressed in patients with *NF1* mutations of the 3' tertile as compared with patients carrying *NF1* mutations in the middle tertile, Student's t-test. The boxes represent the 25th and 75th percentiles. The whiskers show the minimum and maximum value of the distribution, excluding the outliers. The big horizontal lines represent the median value. The outliers are represented as spots outside of the boxes and excluded from the Student's t-test analysis. Statistically significant P values obtained by Student's t-test are showed above the bars.

significant with a FDR of 0.05 and 0.025 after B-H correction for multiple tests

§ significant with a FDR of 0.05 , 0.025 and 0.01 after B-H correction for multiple tests

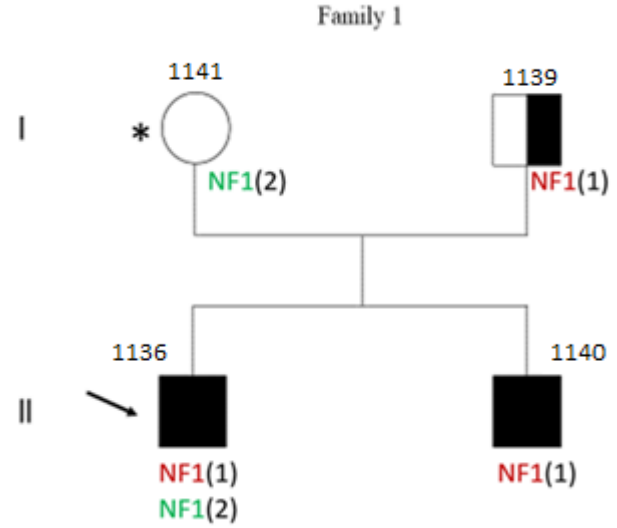
4.6 Identification of familial and sporadic compound heterozygotes for the NF1 gene

Two sporadic SNF patients and three probands from SNF families carry two variants in the NF1 gene. By contrast, double mutations were not found in classical patients. To verify whether the two variants were on the same allele (*incis*) or on two different alleles (*in trans*), an analysis of the segregation of the variants from parents to children in families 1, 17 and 18 was carried out (figure 15).

The proband of family 1 (1136) carries a *NF1*:c.62T>A *(p.Leu21His)* missense mutation inherited from his father MNFSR (1139) and shared with his brother SNF (1140). Furthermore, the patient 1136 has a second missense variant*NF1*:c.528T>A *(p.Asp176Glu)* inherited from the mother with clinical suspicion of NF1 (3 CALMs). The two variants are inherited *in trans* by the proband.

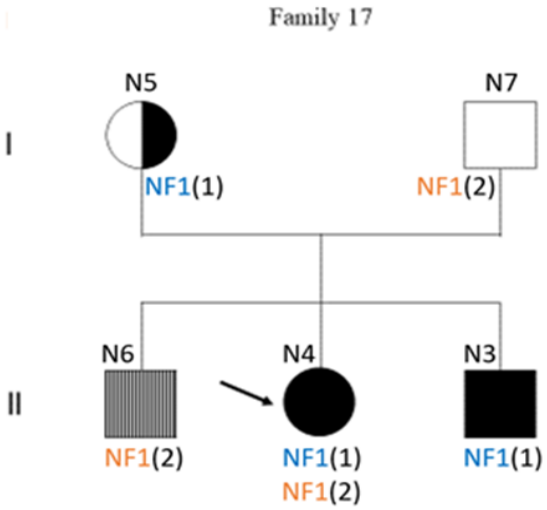
In family 17, the SNF proband (N4), who has a splice mutation c.3314 + 2T> C inherited from her mother (N5) with MNFSR, shows a second missense variant *NF1*: c.7532C>T *(p.Ala2511Val)* inherited from the presumably healthy father (he did not reach clinical observation) and shared with his brother, who is affected by classical NF1. The two variants are inherited in *trans* by the proband.

In family 18, the SNF proband (N8) shows the missense mutation*NF1*:c.1595T>G *(p.Leu532Arg)* and the missense variant *NF1*:c.3242C>G *(p.Ala1081Gly)* in the NF1 gene, shared with the sister and subsequently inherited by the nephew, indicating that the two variants are inherited in *cis.*



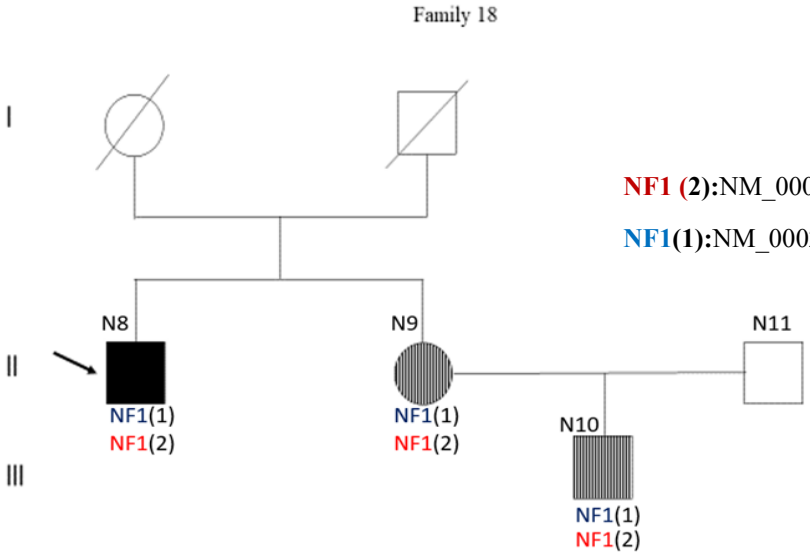
NF1(1): NM_000267:exon2:c.62T>A *(p.Leu21His)*

NF1(2) : NM_000267:exon5:c.528T>A *(p.Glu176Asp)*



NF1(1) :NM_000267:exon25:c.3314+2T>C

NF1 (2) :NM_000267:exon50:c.7532C>T*:(p.Ala2511Val)*

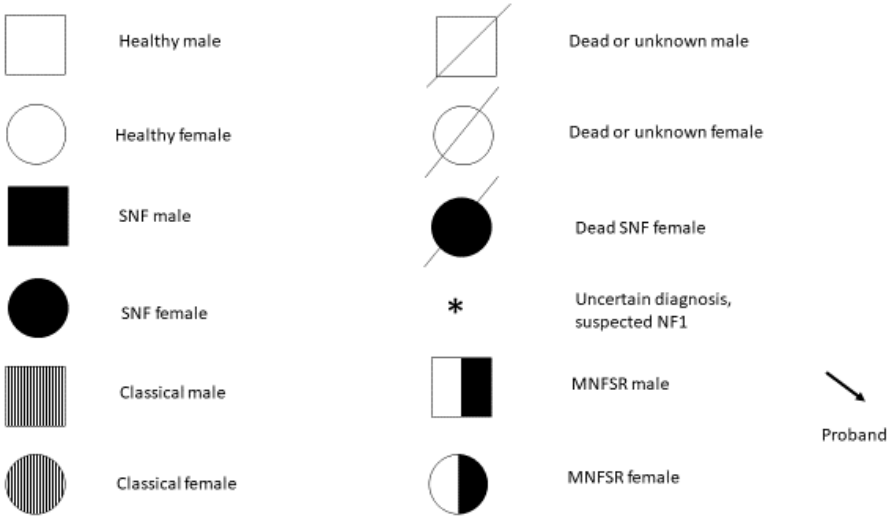


NF1 (2):NM_000267:exon14:c.1595T>G*:(p.Leu532Arg)*

NF1(1):NM_000267:exon25:c.3242C>G*:(p.Ala1081Gly)*

Figure 15. Pedigrees of heterozygous families for double *NF1* mutations

Legend:



Moreover, two sporadic SNF patients carry two variants of the *NF1* gene: a pathogenetic mutation and a missense variant classified as "uncertain" in the Intervar database. Since it is not possible to recover the gDNA of the parents, we could not establish whether the variants on the NF1 gene of these two patients are in *cis* or *trans*. To notice, no one of the classical NF1 patients showed double *NF1* mutations in our cohort of 106 patients.

4.7 Expression analysis with digital PCR of the double *NF1* mutations in family 1

To establish whether the mutated alleles were both expressed in the analysed members of family 1 (the biological samples of the family 17 were not available), we performed digital PCR (dPCR) with specific labelled probes and primers selected to detect both the isoform I and II of NF1, on all the family members, confirming that the mutations present in the gDNA of the three patients and of the healthy mother are expressed in the mRNA (Fig 16a and b).

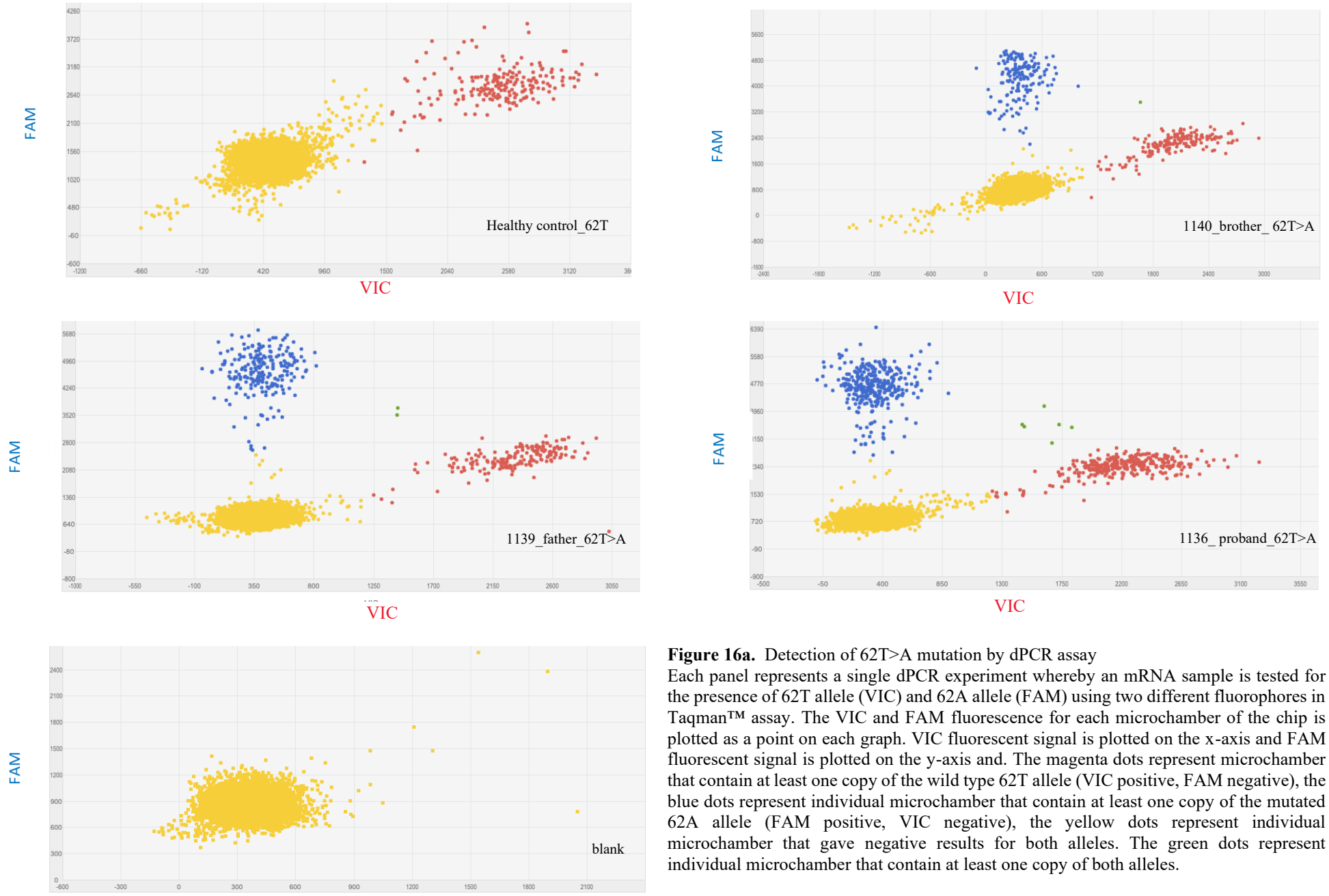


Figure 16a. Detection of 62T>A mutation by dPCR assay
Each panel represents a single dPCR experiment whereby an mRNA sample is tested for the presence of 62T allele (VIC) and 62A allele (FAM) using two different fluorophores in Taqman™ assay. The VIC and FAM fluorescence for each microchamber of the chip is plotted as a point on each graph. VIC fluorescent signal is plotted on the x-axis and FAM fluorescent signal is plotted on the y-axis and. The magenta dots represent microchamber that contain at least one copy of the wild type 62T allele (VIC positive, FAM negative), the blue dots represent individual microchamber that contain at least one copy of the mutated 62A allele (FAM positive, VIC negative), the yellow dots represent individual microchamber that gave negative results for both alleles. The green dots represent individual microchamber that contain at least one copy of both alleles.

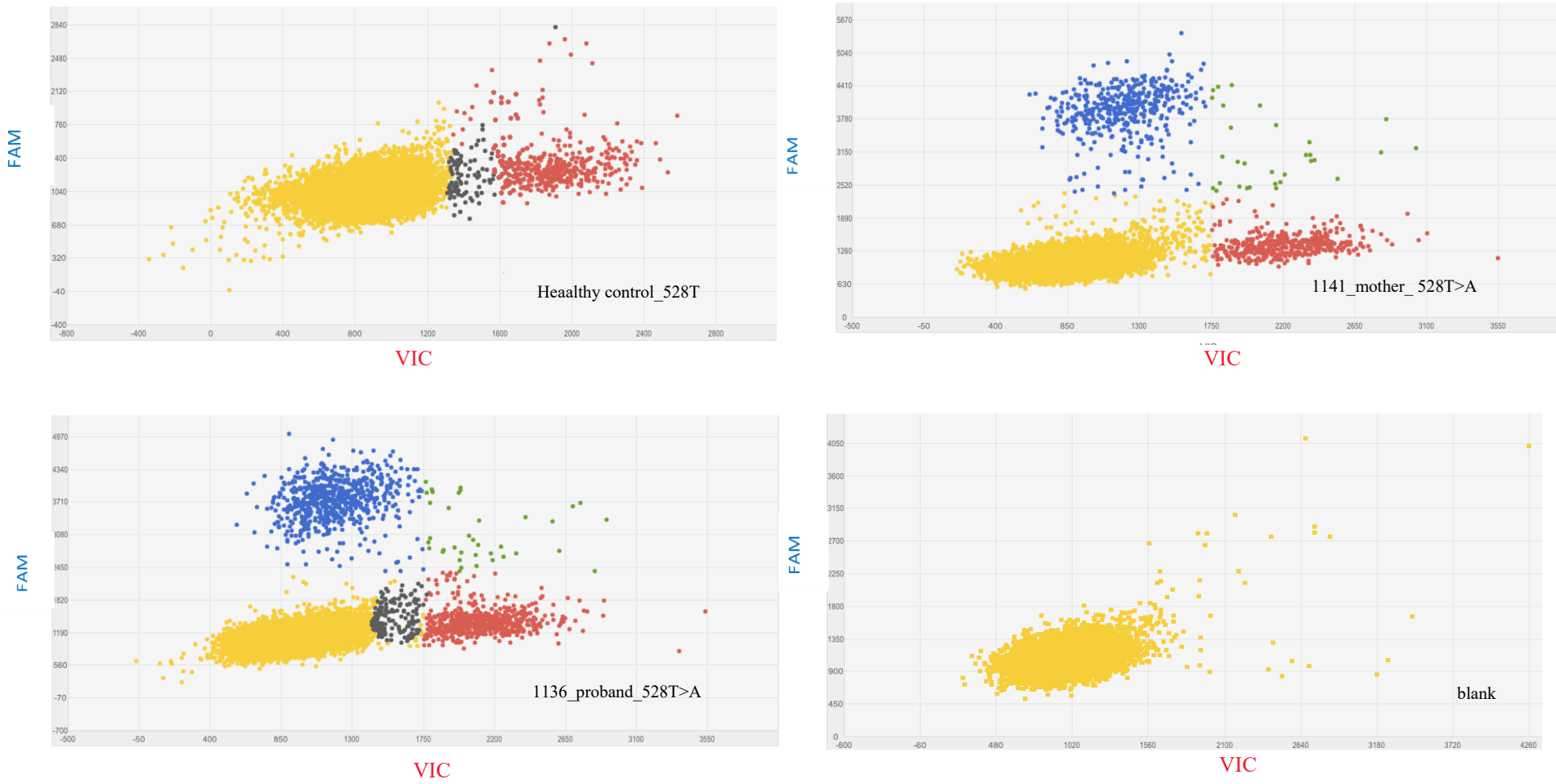
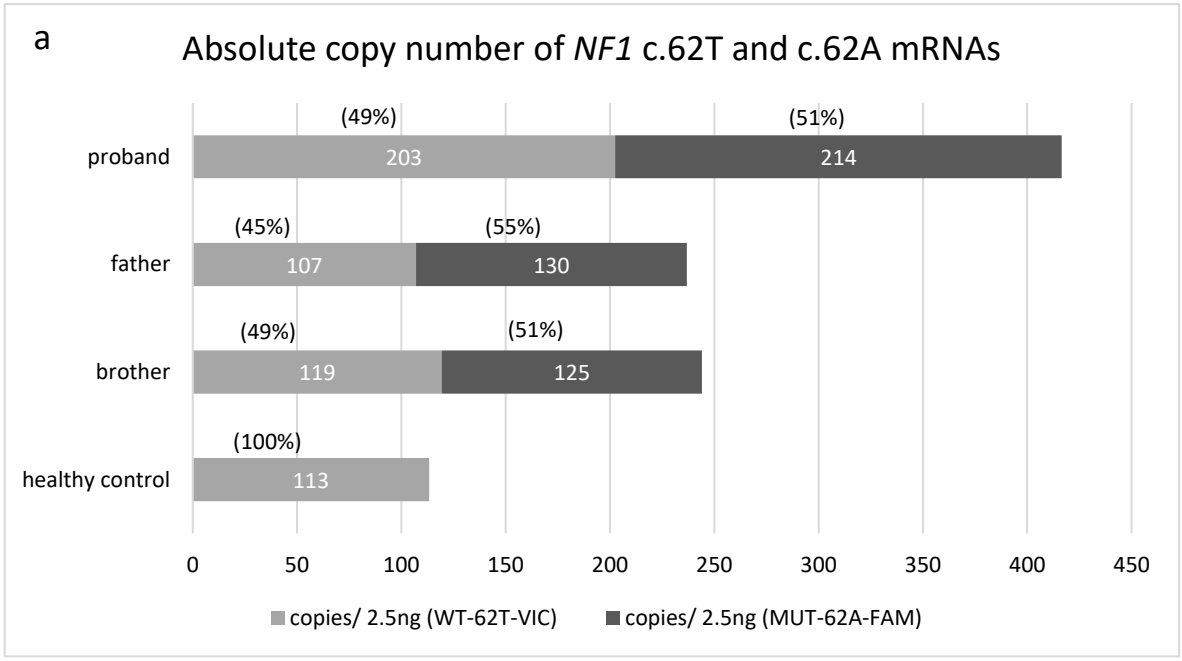


Figure 16b. Detection of 528T>A mutation by dPCR assay
Each panel is a single dPCR experiment whereby an mRNA sample is tested for the presence of 528T allele (VIC) and 528A allele (FAM) using two different fluorophores in Taqman™ assays. The VIC and FAM fluorescence for each well of the chip is plotted as a data point on each graph. VIC fluorescent signal is plotted on the x-axis and FAM fluorescent signal is plotted on the y-axis. The magenta dots are wells that contain at least one copy of the wild type 528T allele (VIC positive, FAM negative), the blue dots represent individual wells that contain at least one copy of the mutated 528A allele (FAM positive, VIC negative), the yellow dots represent individual wells that gave negative results for both alleles. the green dots are individual wells that contain at least one copy of both alleles. The grey dots represent microchambers excluded from the analysis.

The dPCR analyses have been carried out by using specific probes for 62T and 62A alleles of the *NF1*: c. 62T>A *(p.Leu21His)* mutation and for the 528T and 528A alleles of the *NF1*: c.528T>A *(p.Asp176Glu)* mutation, to establish the amount of each transcript in the proband, in the relatives carrying one of the two mutated *NF1* copies, and in a male healthy control. The overall amount of the mutated transcripts is about 4 times higher in the proband compared to the wild type transcripts of the control, 1.7 times higher compared to his father and brother (Fig.17a), and 1.4 times compared to his mother (Fig 17b). The overall amount of wild-type and mutated transcripts present in the relatives is more than twice expressed compared to the *NF1* expression in the healthy control (Figure 17a and 17b). Moreover, the mutated alleles are expressed at comparable level in the compound heterozygous, as well as the mutated and wild-type alleles in the relatives. Interestingly, not only the mutated transcripts, but also the wild-type transcripts are hyperexpressed in the relatives compared to the control.



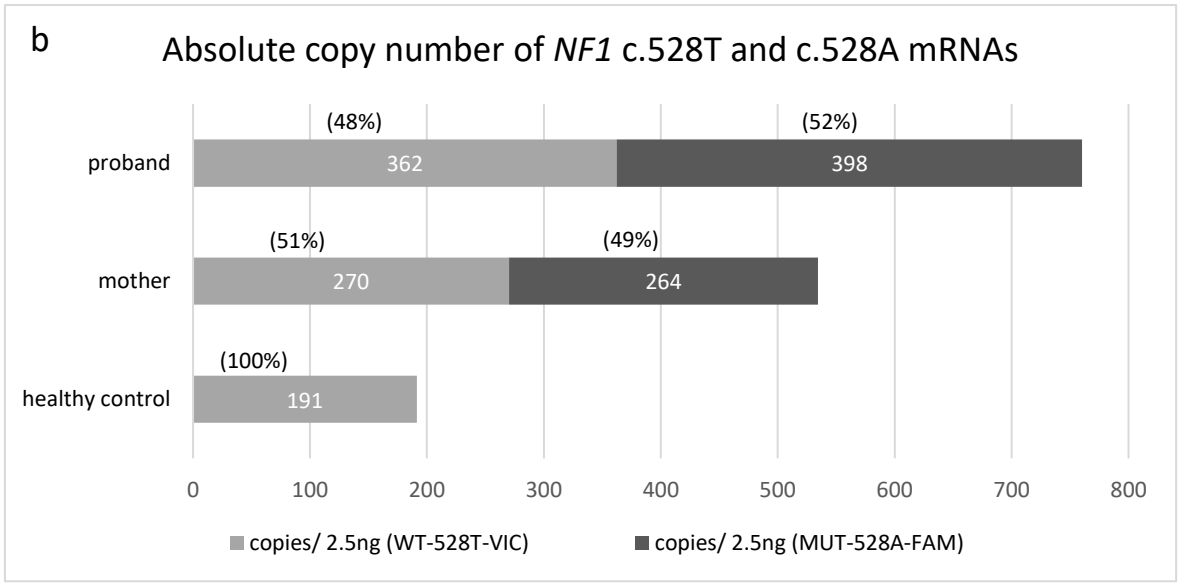


Figure 17. Absolute quantification and percentage of the 62 TA (a) and 528 TA (b) alleles in 2.5 ng of cDNA template in family 1 and in a healthy control

All the data concerning the expression analysis of alleles 62T, 62A, 528T and 528A of *NF1* in family 1 are reported in Table S5 (section "6. Supplementary").

To evaluate possible different effects on the *NF1* expression, associated to missense or to stop gain mutations, we carried out a dPCR assay (Figure 18) in the patient 1493, carrier of the *NF1* stop gain mutation c.2446C>T *(p.Arg816*)* in heterozygous condition. We used as healthy control the same sample analyzed for the dPCR study carried out in family 1 and an RNA pool from five male healthy controls (healthy controls pool), to estimate the expression variability of *NF1* gene.

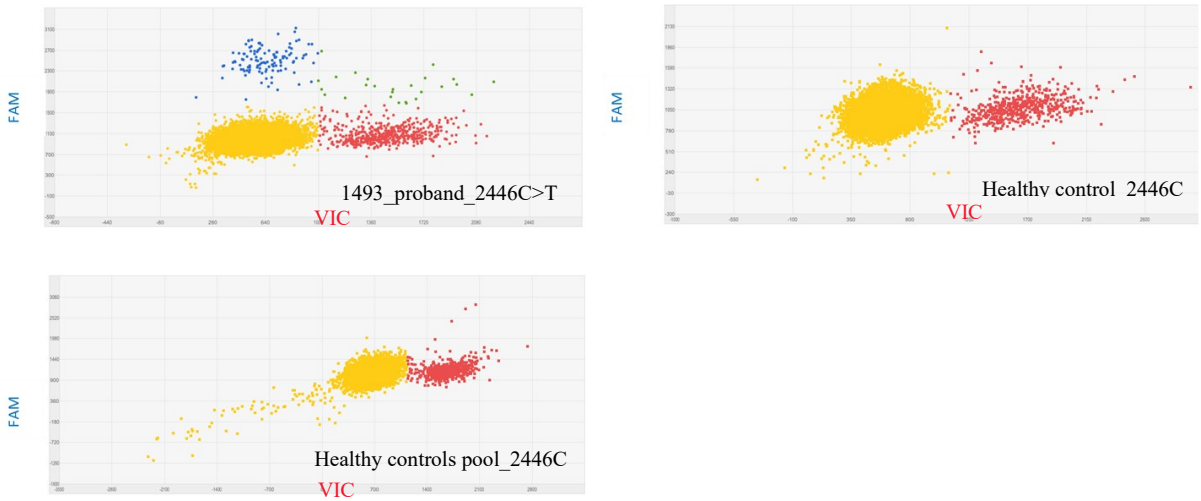


Figure 18. Detection of *NF1*: c.2446C>T *(p.Arg816*)* mutation by dPCR assay. Each panel represents a single dPCR experiment whereby an RNA sample is tested for the presence of 2446C allele (VIC) and 2446T allele (FAM) using two different fluorophores in Taqman™ assays. The VIC and FAM fluorescence for each well of the chip is plotted as a data point on each graph. VIC fluorescent signal is plotted on the x-axis and FAM fluorescent signal is plotted on the y-axis. The magenta dots represent wells that contain at least one copy of the wild type 2446C allele (VIC positive, FAM negative), the blue dots represent individual wells that contain at least one copy of the mutated 2446T allele (FAM positive, VIC negative), the yellow dots represent individual wells that gave negative results for both alleles. the green dots represent individual wells that contain at least one copy of both alleles.

The overall amount of wild type and mutated transcript in the patient 1493 is 1.3 times higher compared to the *NF1* expression in the healthy control and 1.5 in healthy controls' RNAs. The dPCR showed that the wild type allele 2446T is expressed at 78% and the mutated allele 2446C at 22% , (Figure 19).

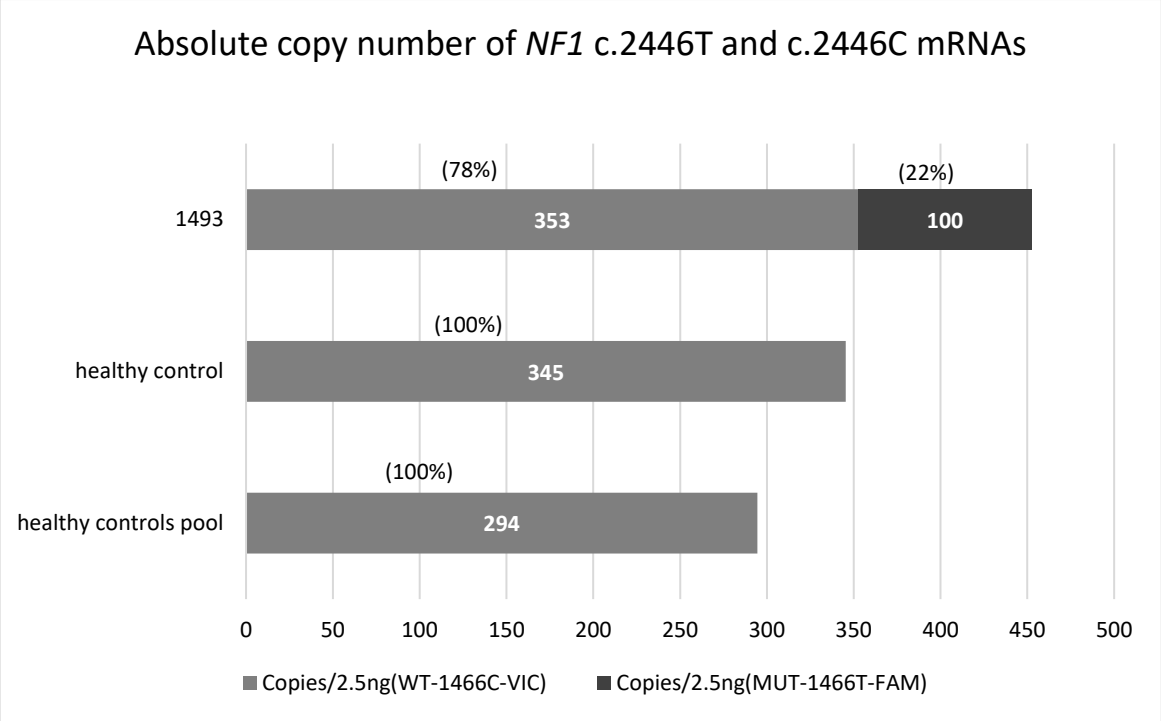


Figure 19. Absolute quantification and percentage of the 2446C and 2446T alleles in 2.5 ng of cDNA template in the SNF patient (1493), in a healthy control and a in pool of five healthy controls

All the data concerning the expression analysis of the wild type allele 2446C and of the mutated allele 2446T of *NF1* in patient 1493 are reported in Table S6 (Supplementary materials).

4.8 Loss Of Heterozygosity (LOH) study in patients' tumoral DNA:

In family 1 we also carried out a LOH study on the DNA of dermal neurofibromas resected in the patient (1136) and its brother (1140). No LOH of one of the *NF1* mutated alleles was observed in the tumoral tissues of the composite heterozygous proband (1136) and in his SNF brother (1140, carrying the germinal 62TA mutation). The DNA sequence of the proband (1136) showed the two mutations (Figure 20 a and b) while the brother carries the 62 T>A mutation (Figure 20c).

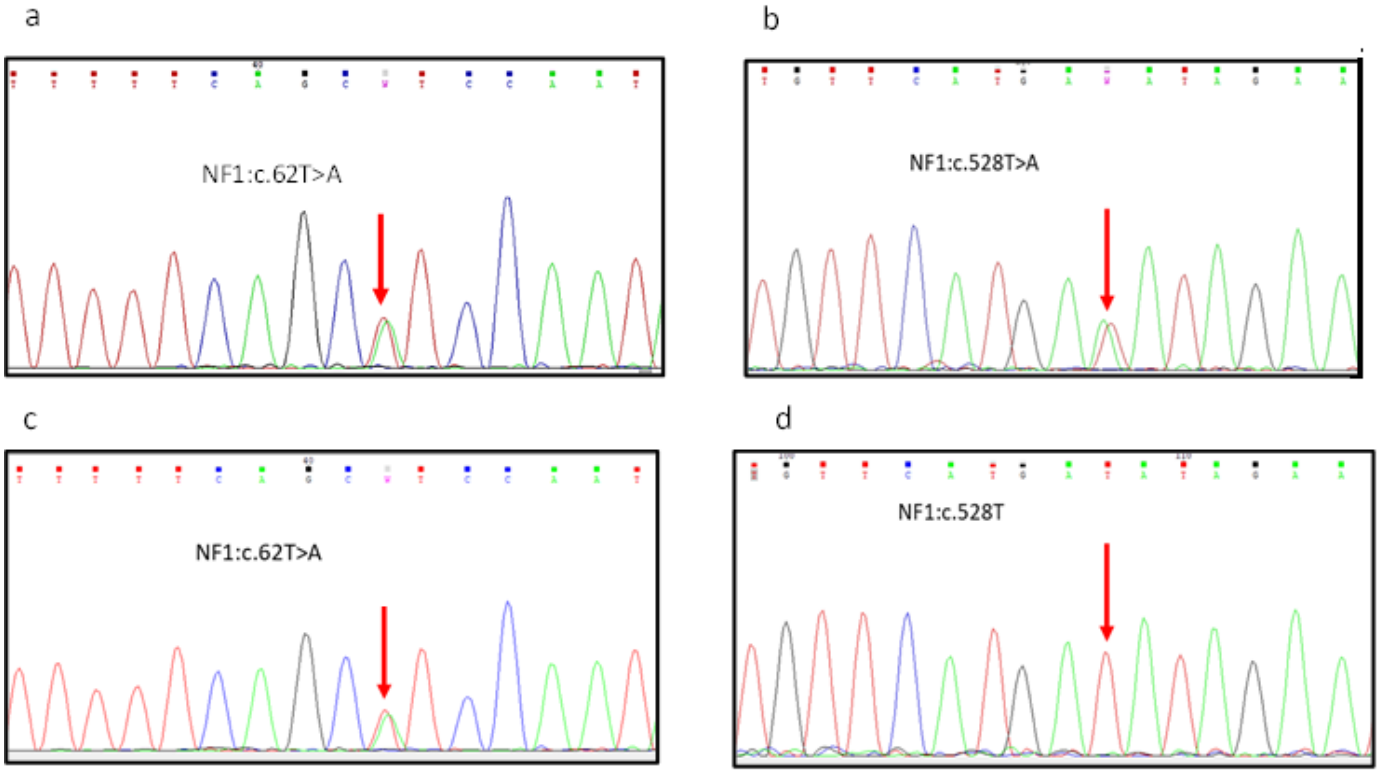


Figure 20. Partial chromatograms of *NF1* exon 2 (panels a and c) and exon 5 (panels b and d) of sequences in patients of family 1 neurofibromas derived DNA

(a). The DNA from subject 1136 (proband) is heterozygous for the *NF1*: c.62T>A transition in exon 2. (b).The DNA from subject 1136 (proband) is heterozygous for *NF1*: c.528T>A transition in exon 5. (c). The DNA from subject 1140 (brother) is heterozygous for the *NF1*: c.62T>A transition in exon 2. (d). The DNA from subject 1140 (brother) is homozygous for the wild type allele *NF1*: c.62T.

5. DISCUSSION

SNF is a distinct clinical entity of NF1. It can be distinguished from NF1 by means of spinal MRI, nevertheless the molecular bases of the two forms are still unknown [3]. Our study seems to indicate that two differential *NF1* mutational spectra emerge in SNF and classical NF1. A deregulation of syndecans' expression is also observed, suggesting their involvement in NF1 pathogenesis.

Here we confirmed, in a large SNF cohort, the previous observations reported in the literature [24], indicating the prevalence of *NF1* missense mutations in SNF. With the present work we increased the power and significance of the statistical analysis by joining our SNF cohort and the SNF patients clinically and genetically described in literature, by a combined analysis. Our data confirms that SNF and classical NF1 are characterized by two different mutational spectra, specifically enriched in missense mutations in SNF patients. The presence of *NF1* missense mutations could indicate that their functional significance could lead to a gain-of-function of mutant nerofibromin, impairing other additional pathways besides RAS signal activation, typically involving loss-of-function *NF1* mutations underlying classical NF1. Truncating and frameshift mutations, proportionally more frequent in classical patients, result in a loss of protein function, while missense mutations, observed more frequently in patients with spinal form, could lead to a gain in neurofibromin functions. A recent study showed that [63] full-length neurofibromin dimerizes with high affinity both in vitro and in human cells. To identify the domains of the protein involved in dimerization, the authors carried out co-immunoprecipitation experiments by dividing it into six domains (A, B, C, D, E, F). The results obtained suggest that the E domain, containing the HLR domain and consisting of HEAT-like repeats, commonly involved in protein/protein interactions, and the C domain, are involved in the formation of the dimer. Since it has been shown in several studies that in heterozygous cells for truncated *NF1* mutations or frameshifts there is an amount of neurofibromin less than 50%, the same authors have hypothesized that truncating mutations or frameshifts are involved in the degradation of the wild-type protein, which forms a dimer with the mutated protein. This mechanism may be not applied to missense mutations found in spinal patients. Consistently, in 2019 Frayling et al hypotized[64]

that *NF1* missense mutation could affect the function of neurofibromin acting within the cell as a dimer, perhaps by acting dominant-negatively.

Moreover, our results indicate the involvement of a specific neurofibromin domain in SNF. The prevalence of missense *NF1* mutations and the prevalent localization of *NF1* mutations in the 3' tertile strongly support the hypothesis of gain of function missense mutations targeting the C-terminal neurofibromin domains in the pathogenesis of SNF.

Interestingly, the prevalence of mutations in the 3' tertile including much of the E domain, could lead to an increased or decreased efficiency in the formation of the dimer, with a possible gain or loss of protein functions. The coding regions of the 3' tertile, where the mutations of spinal cases fall more frequently than the classic ones, corresponds to the C-terminal of neurofibromin. In this protein portion HLR, NLS and SBR domains are present. The first two domains are necessary for the nuclear localization of neurofibromin, while SBR is involved in translocation of neurofibromin along the membrane, by binding syndecans. The interaction between neurofibromin and syndecans is important for cell differentiation and proliferation and for synaptic plasticity[65]. We speculate that mutations in the 3' tertile or in its interactors could have different role in development of SNF rather that classical NF1.

The colocalization of neurofibromin near the pre- or post-synaptic membrane could promote the GTPase activity of RAS or contribute to NF1 specific phenotypes, currently unknown. Data from single cell RNAseq show that, the levels of distribution in the different cell types of *NF1* and the syndecan transcripts are comparable (Gtex portal). Moreover, the four syndecans interact in turn with CASK, a membrane-associated guanylate cyclase expressed in the brain in the embryonic and postnatal phase. The following binding with neurofibromin causes the formation of the neurofibromin-syndecan-CASK, a ternary protein complex, whose role has not yet been identified[65]. Given this evidence in the literature, we investigated the presence of genetic variants in *NF1* 3' tertile interactors and found the co-presence of syndecans' genetic variants with an uncertain and likely pathogenic clinical significance only in SNF patients carrying mutations in the other two *NF1* tertiles. No classical patient presents this type of variants. We at now do not know if the presence of missense variants in the 3' tertile of the NF1 gene and the co-presence of variants in the syndecans with *NF1* mutations of the 5' and middle tertile could contribute to determining the spinal phenotype and to elucidate the role of syndecans in NF1, but the possible role of this NF1 region can be a challenge for studies aimed at identifying not only new

diagnostic, but also pharmacological targets in NF1 disease. Finally, we found that the syndecans are hyper-expressed in SNF and classical NF1, suggesting their possible role in NF1 pathogenesis. The data presented here led us to formulate two hypotheses: the impairment of NF1 3' tertile may i) be less associated with classic NF1 or ii) have a pathogenic significance in SNF. However, the present study on syndecans expression is performed on pheripheral blood mRNA and we don't know if the sydecans' hyperexpression is also present in the peripheral nervous system or in the skin, the two tissues mainly involved in NF1. Further expression studies on other tissues and also functional studies are necessary to confirm these preliminary data. Another aspect unravels from the present work, that seem to differentiate SNF from the classical NF1 form, is the presence of a "second" *NF1* subclinical variant in SNF. At our knowledge, besides this study, only one study describes a *NF1* compound heterozygous patient[48]. Two studies reported on two double mutated NF1 patients, where the two mutations are present *in cis*, and the other *NF1* copies are WT[49,50]. Interestingly, the double *NF1* mutated patients of family 1 and 17 display a more severe phenotype than relatives carrying one *NF1* mutation, showing proband 1 and proband 17.

In family 1 the c.62T>A (*pLeu21His)* missense mutation was shared by the father (1139), the proband (1136) and his brother (1140) seems to be associated to a SNF condition. In fact, the pathogenicity predictive results (ANNOVAR) show 17/20 deleterious predictors. Furthermore, the variant was never described in ClinVar and it is not present in the populations of the 1000 genomes and Exac databases. In the same family, the mother (1141), shares with the proband the missense variant c.528T> A *(p.Asp176Glu)*. In fact, the mother at 55 years aged had no tumor in her medical history, at physical examination only two CAL spots, one on left arm and one on the chest, were observed, no Lisch nodules were detected at eye examination. MRI with gadolinium showed several small (diameter less than 1 cm) nodular enhancing lesions in the laterocervical soft tissues, suggestive for neurofibromas, but no further sign of NF1. Thus, she did not meet the diagnostic criteria for NF1. The association of this variant with a subclinical phenotype is consistent with the pathogenicity predictive results showing 10/20 deleterious predictors. Furthermore, according to the InterVar analysis, the variant is classified as benign, while ClinVar predict the variant with conflicting interpretation of pathogenicity. Accordingly, patient the proband of family 1 shows a more severe phenotype than his affected father and brother.

According to Mauda-Avakuk [66], in patients with NF1 with spinal involvement two features correlate with clinical presentation and outcome: tumor burden and a particular NF subtype, in

particular the presence of kissing neurofibromas at cervical level is a risk factor for a greater morbidity. Furthermore, an association between clinical outcome and the presence of neurofibromas at cervical region and intradural involvement have been reported. [67,68]
In family 1 both brothers demonstrated a severe burden of spinal disease presenting with multilevel disease. However, the composite heterozygous proband, had symmetrical bilateral kissing neurofibromas and showed an early and quick tumor growth according to REINS criteria [69] and a greater morbidity, due to cervical myelopathy.

Similarly, in the family 17, the proband shows a more severe phenotype in respect to the relatives carrying one or the two *NF1* mutations. The splicing mutation c.3314 + 2T> C is present in the SNF or MNFSR patients (N3, N4 and N5), indicating an association with the SNF form, while the second missense variant c.7532C> T *(p.Ala2511Val)* is present in the NF1 classical subject (N6), and in the father (referred healthy, but never clinically evaluated). Interestingly, the proband (N4) of family 17, carrying both *NF1* mutations, presents, besides the SNF form, *cafè-au-lait* spots (n= 6-10), axillary freckles, UBOs, Lisch nodules and cutaneous (n= 1-10) neurofibromas, clinical signs more frequent in the classical form of NF, as observed in our casuistry (Table 15) and in the literature [70, 71, 72].

The co-presence of the SNF-associated allele with a second *NF1* mutated allele in both probands, correlated to a more severe phenotype, indicates that the neurofibromin encoded by one of the two alleles maintain a partial function. Consistently, the loss-of-function *NF1* mutations were never detected in homozygous status, having been demonstrated lethal in the null $nf1^{-/-}$ mouse [73,74]. Previosly reported *NF1* missense mutations showed hyperexpression in an experimental assay in which the ability to produce mature neurofibromin in transformed $NF1^{-/-}$cells by mutant mouse cDNAs carrying variants found in NF1-affected individuals, were established[75] .

Interestingly, the compound heterozygotes here described and the only other reported in literature [48], are affected by SNF. One explanation could be that they carry a missense mutation, recently associated to SNF form and probably causing a partial neurofibromin function, being all the described cases compound heterozygous for NF1 locus. Nevertheless, additional cases should be studied to verify a specific association between the heterozygous compound genotype and SNF.

The condition of *NF1* compound heterozygotes is not usually verified in NF1 patients because of the complex setting up of analytical methods of *NF1* mutation detection, before NGS diagnostic

application. The *NF1* variants with uncertain pathogenic significance are probably underestimated, as well as the *NF1* compound heterozygotes that could account for the variable expressivity of the disease even in familial cases[76].

Our study provided data about the expression of mutated and WT *NF1* alleles applying dPCR in RNA samples from patients of family 1. Little is known about the differential allelic expression of the mutated and WT *NF1* alleles in NF1 patients[77]. We observed the hyperexpression of both mutated alleles in the proband and of mutated and WT alleles in the relatives. The hyperexpression of WT allele in patients carrying one *NF1* mutation could be caused by an epigenetic mechanism. As the analyzed *NF1* mutated transcripts carry missense mutations, we performed a differential allelic expression study in a patient carrying an *NF1* stop-gain mutation. Our findings indicate that the expression of stop-gain mutated allele is strongly reduced compared to the WT allele, as expected for non-sense-mediate-decay mechanisms, commonly active when stop-gain mutations occur. Moreover, we included a WT control derived by a pull of five healthy individuals allowing us to verify the stability of the WT expression level.

We also checked in the family 1, the *NF1* LOH in tumour samples, on the DNA derived from neurofibromas of the cervical roots, surgically removed in the proband (1136), and on the DNA derived from neurofibromas of the peripheral nerves, removed in its brother (1140). Most of the studies addressing the LOH of *NF1* in the literature are performed on cutaneous neurofibromas, reporting a percentage of LOH in the tumoral tissues ranging from 2,26% to the 32% [78,79,80]. The only study performed on spinal neurofibromas in SNF patients, by Upadhyaya et al.[81], identified *NF1* LOH in 8/22 of the studied spinal tumour tissues. The absence of LOH in our patients is not in contrast with the previous reported studies and could be consistent with a gain of function significance of both variants, thus contributing to the tumoral phenotype of the patients.

6. CONCLUSIONS:

- The SNF is characterized by a higher prevalence of *NF1* missense mutations, confirming data already obtained in previous studies carried out in small casuistries.
- SNF is also characterized by a higher prevalence of *NF1* mutations in the 3' tertile of the gene. These results indicate that the two forms of neurofibromatosis could depend on altered levels of neurofibromin in the classical or on the presence of a mutated neurofibromin with new functions in the spinal form. The significative presence of mutations in the 3' tertile of *NF1* in SNF patients suggests that the functional domains SBR, NLS and HLR, whose activity is unknown, may play a role in the development of the spinal form of NF1.
- In our large SNF casuistry, we detected a significative occurrence of mutations in both copies of *NF1* gene: beside the causative *NF1* mutation, we found a second *NF1* variant with uncertain clinical significance.
- The expression analysis of both *NF1* mutated alleles in one SNF family by dPCR assay showed that both the variants are expressed at comparable levels in the composite heterozygous proband and theìr carrier relatives. We also showed that the overall amount of the the mutated transcripts is higher in the composite heterozygous proband compared to the wild type transcripts of the control.

These results need to be confirmed in studies with larger case series, aimed at evaluating the differential allelic expression of *NF1*, to understand the contribution of each variant in determining the spinal or classic phenotype. The occurrence of compound heterozygotes NF1 patients, that were previously diagnosed with other methods, can be also reevaluated by the application of massive sequencing techniques, to assess the frequency of compound heterozygotes. In fact, the occurrence of a second *NF1* variant, partially affecting the neurofibromin function, could be related to variable expressivity of the phenotype even in intra-family cases.

The above results suggest the presence of two mutational *NF1* spectra characterizing the classical and spinal form of NF1. Spinal neurofibromas, observed in SNF, can remain asymptomatic for years, making it difficult to make an early diagnosis. Moreover, once severe neurological deficits have developed, the likely success of any surgical treatment is greatly reduced. Our study, carried out in a large series, strongly indicates that *NF1* missense mutations could be a prognostic marker

for the spinal form of the disease, useful to an early differential diagnosis, as well asthe presence of mutations in the 3' tertile of *NF1* or of a second subclinical variant of *NF1*, once our results are confirmed in larger series and by functional studies.

Further results obtained in this study, are indicative of a role of the interactors of *NF1* gene 3' tertile in the pathogenesis of the two forms of neurofibromatosis type I:

- We found rare variants in the syndecans genes, coding for interactors of the sydecan binding domain of neurofibromin, in SNF patients.
- The expression analysis of the sydecans genes showed a hyper-expression of the syndecans in NF1 patients presenting mutations in the *NF1* 3' tertile, as compared to controls. There are no data in the literature that confirm or help us to understand the syndecan deregulation observed in NF1 patients. Moreover, functional studies are necessary to confirm the biological significance of our findings. To understand the possible pathogenic role of syndecans in NF1, iPSCs from PBMC of NF1 patients with mutations in 3' *NF1* tertile, could be generated, differentiated in specific cellular lineage involved the NF1 (keratinocytes, cortical neurons) and used in functional studies in which the syndecans levels could be modulated with standard techniques (for example siRNA), also addressing the identification of potential pharmacological targets for NF1 treatments. Moreover, nf1a -/-; nf1b-/- zebrafish commercial lines could be studied before and after syndecans' dysregulation (sdc2 and sdc3 has a homology of 50% about with SDC2 and SDC3) to verify alterations in the animal model phenotypes[82]. Knowing that syndecans are adhesion molecules, their functions could be inhibited by specific pharmacological treatments. Interestingly, monoclonal antibodies and enzymatic inhibitors of the syndecans are largely studied in various types of cancer with inhibiting effects on cell growth and migration [83]. These molecules could also be used in Induced pluripotent stem cells (iPSCs)[84] derived from NF1 patients to downregulate the high level of syndecans observed in patients, to assess the role of the syndecans in the pathology and to identify potential pharmacological treatments for NF1.

Currently, the only promising pharmacological treatment in NF1 seems to be the MEK inhibitors [84, 85] used in phase II clinical trialsand in mouse animal models [86] studying the reduction or growth inhibition of plexiform neurofibromas; however, there are no effective pharmacological treatments for the cutaneous, subcutaneous, or spinal neurofibromas. If conclusive, the results obtained in functional studies on syndecans, could address future

studies aimed at the identification of new pharmacological targets in the context of the neurofibromatosis type I.

7. SUPPLEMENTARY

Table S1. Evaluation of the clinical significance of the *syndecans* variants

Patient	Phenotype	Variant	OMIM ID	MIM ID	SNV localization (GRCh37/hg19)	Ref	Alt	Missense variant	dbSNP Build 154	gnomAD v.3.1.2 (NFE)	1000 genomes euro 2015
190NF1sp	SNF	*SDC1*:NM_002997.5:exon5:c.G830A	186355	no	chr2:20402630-20402630	C	T	p.R277H	rs745498874	not reported	not reported
34NF1c	CLASSICAL NF1	*SDC1*: NM_002997.5: exon3:c.C215T	186355	no	chr2: 20403986-20403986	G	A	p. T72M	rs766762575	0.00003030	not reported
N46	SNF	*SDC2*: NM_002998.4: exon5: c. T449C	142460	no	chr8:97621619-97621619	T	C	p. I150T	rs143693366	0.0001176	not reported
157NF1sp	SNF	*SDC3*: NM_014654.4: exon4:c.C923T	186357	#601665: {Obesity, association with} AD, AR, Mu	chr1: 31347383-31347383	C	T	p. P308L	rs374473992	0.00007350	not reported
N33	SNF	*SDC3*: NM_014654.4: exon3: c. A721G	186357	#601665: {Obesity, association with} AD, AR, Mu	chr1:30876701-30876701	A	G	p. T241A	rs369847197	0.00005882	0,001
N1	SNF	*SDC4*: NM_002999.4: exon2: c. A92G	600017	no	chr20: 45335889-45335889	A	G	p. D31G	rs200999058	0.0001029	not reported

Patient	Phenotype	Variant	DECIPHER v.11.13	COSMIC v.92	ClinVar	Novel	Conserved aminoacid	functional domain	DamagePred Count by ANNOVAR	Clinical significance
190NF1sp	SNF	*SDC1*:NM_002997.5:exon5:c.G830A	not reported	COSM6773650, reported in 2 samples of adenocarcinoma from large intestine as pathogenic	not reported	no	yes	yes	18,2	**likely pathogenic**
34NF1c	CLASSICAL NF1	*SDC1*: NM_002997.5: exon3:c.C215T	not reported	COSM3380002, reported in a sample of adenocarcinoma from large intestine as neutral	not reported	no	yes	no	7,2	**uncertain**
N46	SNF	*SDC2*: NM_002998.4: exon5: c. T449C	not reported	COSM8912385, reported in a sample of endometrial carcinoma as pathogenic	not reported	no	yes	yes	18,2	**likely pathogenic**
157NF1sp	SNF	*SDC3*: NM_014654.4: exon4:c.C923T	not reported	COSM4878319, reported in 2 samples of prostatic adenocarcinoma as pathogenic	not reported	no	yes	no	11,2	**uncertain**
N33	SNF	*SDC3*: NM_014654.4: exon3: c. A721G	not reported	not reported	not reported	no	no	no	4,2	**uncertain**
N1	SNF	*SDC4*: NM_002999.4: exon2: c. A92G	not reported	not reported	not reported	no	no	no	2,2	**uncertain**

Table S2. Evaluation of the clinical significance of variants in NF1 interactors of middle and 5' tertiles

Patient	Phenotype	Variant	OMIM ID	MIM ID	SNV localization (GRCh37/hg19)	Ref	Alt	Missense variant	dbSNP Build 154	gnomAD v.3.1.2 (NFE)	1000 genomes_eur 2015
1214	CLASSICAL NF1	APP (NM_001136131.2):c.1445C>T	104760	104300 (Alzheimer disease); 605714 (Cerebral amyloid angiopathy)	chr 21: 27284187-27284187	G	A	p.A482V	rs754790092	0.00001758	notreported
1165	CLASSICAL NF1	APP (NM_201413.3):c.298C>T	104760	104300 (Alzheimer disease); 605714 (Cerebral amyloid angiopathy)	chr 21: 27462316-27462316	G	A	p.R100W	rs200347552	0.00002939	notreported
1085	SNF	APP (NM_000484.4): c.1754G>A	104760	104300 (Alzheimer disease); 605714 (Cerebral amyloid angiopathy)	chr 21: 27284208-27284208	C	T	p.R585K	notreported	notreported	notreported
876	CLASSICAL NF1	SPRED1(NM_152594.3):c.1313G>T	609291	611431 (Legius syndrome)	chr 15:8643843-8643843	G	T	p.G438V	rs573603750	0.0000792	notreported

Patient	Phenotype	Variant	DECIPHER v.11.13	COSMIC v.92	ClinVar	Novel	functional domain	DamagePredCount by ANNOVAR	Clinical significance
1214	CLASSICAL NF1	APP (NM_001136131.2):c.1445C>T	notreported	Not reported	uncertain	no	yes	16,2	Likely pathogenic
1165	CLASSICAL NF1	APP (NM_201413.3):c.298C>T	notreported	COSM5057538:reported in a sample of intestinal adenocarcinoma as pathogenic	RCV000813633:reported in a patient with Alzheimer disease with uncertain significance	no	yes	17,2	Likely pathogenic
1085	SNF	APP (NM_000484.4): c.1754G>A	notreported	Not reported	RCV002020226: reported in a patient with Alzheimer disease with uncertain significance	no	yes	17,2	Likely pathogenic
876	CLASSICAL NF1	SPRED1(NM_152594.3):c.1313G>T	notreported	Not reported	RCV001044046:reported in a patient with reported in a patient with Legiu syndrome with uncertain significance	no	no	19,2	Likelypathogenic

NFE= Not Finnish European

Table S3. Statistical analysis of real-time PCR data

Gene	Patients	N	2^(?ΔCt) mean	SD	SEM	95% CI		Student's t-test P value		
						Lower	Upper	SNF vs CLASSICAL	SNF vs WT	CLASSICAL vs WT
SDC2	SNF	16	0,4773	0,2034	0,0507	0,3688	0,5856	0,1304	**0,030 ***	**0,0042#**
	CLASSICAL	16	0,6454	0,3689	0,0893	0,4489	0,842			
	WT	16	0,3501	0,1963	0,0475	0,2455	0,4547			
SDC3	SNF	16	0,1811	0,0638	0,0155	0,1471	0,2151	0,5340	**0,0002§**	**0,0128***
	CLASSICAL	16	0,1644	0,0835	0,0208	0,1199	0,2089			
	WT	16	0,1014	0,0379	0,0095	0,6343	-0,4315			
SDC4	SNF	16	0,2689	0,1456	0,0352	0,1913	0,3464	0,6164	0,0367	**0,0162***
	CLASSICAL	16	0,2472	0,0818	0,0204	0,2036	0,2908			
	WT	16	0,1802	0,0599	0,0149	0,1483	0,2121			

In bold are shown the comparisons remaining statistically significant after B-H correction method for multiple tests

 # significant with a FDR of 0,05 and 0,025 after B-H correction for multiple tests

*** significant with a FDR of 0,05 after B-H correction for multiple tests**

§ significant **with a FDR of 0,05 and 0,025 and 0.01 after B-H correction for multiple tests**

SNF, patients with SNF; classical, patients with the classical form of NF1; WT, healthy controls,

N, number of patients; CI, confidence interval ; p, p-value; SD, standard deviation

Table S4. Statistical analysis of real-time PCR data by *NF1* tertiles

Gene	Patients	N	2^(?ΔCt) mean	SD	SEM	95% CI		Student's t-test P value			
						Lower	Upper	3'vs WT	3' vs middle	3' vs 5'	5' vs middle
SDC2	5'	18	0,8112	0,5287	0,1211	0,4000	0,7900	**0,0001§**	**0,0179#**	0,1376	0,1898
	middle	11	0,5686	0,3527	0,1014	0,2500	0,6200				
	3'	10	1,1608	0,3527	0,1988	0,4600	1,2100				
	WT	16	0,3501	0,1963	0,0475	0,1500	0,3000				
SDC3	5'	18	0,1594	0,0648	0,01524	0,048	0,0990	**0,0022#**	0,2491	0,3199	0,7029
	middle	11	0,1499	0,0563	0,0169	0,0390	0,1000				
	3'	10	0,1894	0,0886	0,0292	0,0710	0,1800				
	WT	16	0,1014	0,0379	0,0095	0,0280	0,0600				
SDC4	5'	18	0,2838	0,1583	0,0363	0,1200	0,2400	**§ 8,95E-06**	**0,0024#**	0,0537	0,3654
	middle	11	0,2362	0,0805	0,0232	0,0560	0,1400				
	3'	10	0,4039	0,1352	0,0406	0,0930	0,2500				
	WT	16	0,1802	0,0599	0,0149	0,0440	0,0940				

In bold are shown the comparisons remaining statistically significant after B-H correction method for multiple tests

significant with a FDR of 0,05 and 0,025 after B-H correction for multiple tests

§ significant with a FDR of 0,05 and 0,025 and 0.01 after B-H correction for multiple tests

SNF, patients with SNF; classical, patients with the classical form of NF1; WT, healthy controls, N, number of patients;
CI, confidence interval ; p, p-value; SD, standard deviation

Table S5. Expression analysis of 62TA and 528TA *NF1*mutations in family 1 and a healthy control by dPCR data

Sample	MUT-62A-FAM/Total	CI MUT-62A-FAM/Total	Copies MUT-62A-FAM/microliter	CI Copies/microliter
WT	0%	NA	0	NA
brother	51,09%	42.747% -- 61.008%	16,623	14.409 -- 19.178
father	54,80%	46.445% -- 64.593%	17,301	15.177 -- 19.722
proband	51,39%	44.911% -- 58.739%	28,542	25.651 -- 31.758

Sample	Copies (WT-62T-VIC) /microliter	CI Copies/microliter
WT	15,098	13.139 -- 17.349
brother	15,912	13.749 -- 18.415
father	14,272	12.357 -- 16.485
proband	27,002	24.195 -- 30.134
Sample	**Copies (WT-62T-VIC)/2.5ng**	**Copies (MUT-62A-FAM) /2.5ng**
WT	113,235	0
brother	119,34	124,6725
father	107,04	129,7575
proband	202,515	214,065

Sample	MUT-528A-FAM/Total	CI (MUT-528A-FAM)/Total	Copies (MUT-528A-FAM)/microliter	CI Copies/microliter (FAM)
WT	0%	0% -- 0%	0	0 -- 0
H2O	0%	0% -- 0%	0	0 -- 0
proband	52,35%	47.519% -- 57.613%	53,066	49.228 -- 57.204
mother	49,44%	43.993% -- 55.514%	35,235	32.119 -- 38.655

Sample	Copies (WT-528A-VIC)/microliter	CI Copies/microliter (VIC)
WT	25,509	22.892 -- 28.425
H2O	0	0 -- 0
proband	48,311	44.659 -- 52.262
mother	36,033	32.879 -- 39.489

Sample	Copies (WT-528A-VIC)/ 2.5ng	copies (MUT-528A-FAM)/ 2.5ng (FAM)
WT	191,3175	0
proband	362,3325	397,995
mother	270,2475	264,2625

WT= healthy control, CI = Confidence Interval, MUT-62T-FAM= mutated allele 62T FAM dye labeled, WT-62A-VIC= wild type allele 528A VIC dye labeled, MUT-528T-FAM= mutated allele 528T FAM dye labeled, WT-528A-VIC= wild type allele 528A VIC dye labeled

Table S6. Expression analysis of wild type 2446C and mutated 2446T alleles of *NF1* in patient 1493 and in healthy controlsby dPCR data

Sample	MUT-2446T-FAM/Total	CI MUT-2446T-FAM/Total	Copies MUT-2446T-FAM/microliter	CI Copies/microliter (FAM)
patient 1493	22,07%	18.408% -- 26.424%	13,315	11.334 -- 15.643
WT	0%	0% -- 0%	0	0 -- 0
WT pool	0%	0% -- 0%	0	0 -- 0

Sample	Copies (WT-2446C-VIC) /microliter	CI Copies/microliter (VIC)
patient 1493	47,021	43.133 -- 51.259
WT	46,056	42.353 -- 50.083
WT pool	39,262	35.95 -- 42.878

Sample	Copies(WT-2446C-VIC)/2.5ng	Copies (MUT-2446T-FAM)/2.5ng
patient 1493	352,6575	99,8625
WT	345,42	0
WT pool	294,465	0

WT= healthy control; WT pool = pool of five healthy controls; MUT-2446T-FAM= mutated alllele 2446T FAM dye labelled; WT2446C-VIC= wild type allele 2446T VIC dye labelled; CI = Confidence Interval

8. BIBLIOGRAPHY AND SITEOGRAPHY

1. National Institute of Neurological Disorders and Stroke. www.ninds.nih.gov.
2. Boyd, K. P., Korf, B. R. & Theos, A. Neurofibromatosis type 1. *Journal of the American Academy of Dermatology***61**, 1–14 (2009).
3. Ruggieri, M., Praticò, A. D., Caltabiano, R. & Polizzi, A. Early history of the different forms of neurofibromatosis from ancient Egypt to the British Empire and beyond: First descriptions, medical curiosities, misconceptions, landmarks, and the persons behind the syndromes. *American Journal of Medical Genetics, Part A***176**, 515–550 (2018).
4. Smith, R. W. A treatise on the pathology, diagnosis and treatment of neuroma. 1849. *Clinical Orthopaedics and Related Research* 3–9 (1989)
5. R, V. Uber die Reform der pathologischen und therapeutischen Anschauungen durch die mikroskopischen Untersuchungen. in 207–255 (1847).
6. Virchow R. Die krankhaften Geschwulste. in *Die krankhaften Geschwulste* 325–327 (1863).
7. v. R. F. Uber die Multiplen Fibrome der Haut und ihre Beziehung zu Multiplen Neuromen. *Berlin: August Hirschwald***44**, 1–138 (1882).
8. Statement, C. Neurofibromatosis. (2015).
9. Dulai, S. *et al.* Decreased bone mineral density in neurofibromatosis type 1: Results from a pediatric cohort. *Journal of Pediatric Orthopaedics***27**, 472–475 (2007).
10. Oderich, G. S. *et al.* Vascular abnormalities in patients with neurofibromatosis syndrome type I: Clinical spectrum, management, and results. *Journal of Vascular Surgery***46**, 475–485 (2007).
11. Madanikia, S. A., Bergner, A., Ye, X. & Blakeley, J. O. N. Increased risk of breast cancer in women with NF1. *American Journal of Medical Genetics, Part A***158 A**, 3056–3060 (2012).
12. Zinnamosca, L. *et al.* Neurofibromatosis type 1 (NF1) and pheochromocytoma: Prevalence, clinical and cardiovascular aspects. *Archives of Dermatological Research***303**, 317–325 (2011).
13. Ohtsuki, Y. *et al.* Duodena I Carcino id (Soma tosta t inoma) Corn bined with Vo n Reck li ng ha usen ' s Disease A Case Report and Review of the Literature. 141–146 (1989).
14. Seminog, O. O. & Goldacre, M. J. Risk of benign tumours of nervous system, and of malignant neoplasms, in people with neurofibromatosis: Population-based record-linkage study. *British Journal of Cancer***108**, 193–198 (2013).
15. Coffin, C. M., Davis, J. L. & Borinstein, S. C. Syndrome-associated soft tissue tumours. *Histopathology***64**, 68–87 (2014).
16. Takazawa, Y., Sakurai, S., Sakuma, Y. & Ikeda, T. Gastrointestinal Stromal Tumors of Neurofibromatosis. **29**, 755–763 (2005).
17. Listernick, R., Ferner, R. E., Liu, G. T. & Gutmann, D. H. Optic pathway gliomas in neurofibromatosis-1: Controversies and recommendations. *Annals of Neurology***61**, 189–198 (2007).
18. Ostendorf, A. P., Gutmann, D. H. & Weisenberg, J. L. Z. Epilepsy in individuals with neurofibromatosis type 1. *Epilepsia***54**, 1810–1814 (2013).
19. Cimino, P. J. & Gutmann, D. H. Neurofibromatosis type 1. *Handbook of clinical neurology***148**, 799–811 (2018).

20. Evans, D. G. R. *et al.* Malignant peripheral nerve sheath tumours in neurofibromatosis 1. *Journal of medical genetics***39**, 311–314 (2002).

21. Leeds, N. E. & Jacobson, H. G. Spinal neurofibromatosis. *AJR. American journal of roentgenology***126**, 617–623 (1976).

22. Pulst, S.-M., Riccardi, V. M., Fain, P. & Korenberg, J. R. Familial spinal neurofibromatosis. *Neurology***41**, 1923 LP – 1923 (1991).

23. Poyhonen, M., Leisti, E. L., Kytölä, S. & Leisti, J. Hereditary spinal neurofibromatosis: a rare form of NF1? *Journal of medical genetics***34**, 184–187 (1997).

24. Ruggieri, M. *et al.* The natural history of spinal neurofibromatosis: A critical review of clinical and genetic features. *Clinical Genetics***87**, 401–410 (2015).

25. Rauen, K. A. The RASopathies. *Annual Review of Genomics and Human Genetics***14**, 355–369 (2013).

26. Tajan, M., Paccoud, R., Branka, S., Edouard, T. & Yart, A. The RASopathy family: Consequences of germline activation of the RAS/MAPK pathway. *Endocrine Reviews***39**, 676–700 (2018).

27. Maik-Rachline, G., Hacohen-Lev-Ran, A. & Seger, R. Nuclear erk: Mechanism of translocation, substrates, and role in cancer. *International Journal of Molecular Sciences***20**, 1–18 (2019).

28. Bergoug, M. *et al.* Neurofibromin Structure, Functions and Regulation. *Cells***9**, (2020).

29. Marchuk, D. A. *et al.* cDNA cloning of the type 1 neurofibromatosis gene: Complete sequence of the NF1 gene product. *Genomics***11**, 931–940 (1991).

30. Daston, M. M. *et al.* The protein product of the neurofibromatosis type 1 gene is expressed at highest abundance in neurons, Schwann cells, and oligodendrocytes. *Neuron***8**, 415–428 (1992).

31. Luo, G., Kim, J. & Song, K. The C-terminal domains of human neurofibromin and its budding yeast homologs Ira1 and Ira2 regulate the metaphase to anaphase transition. *Cell Cycle***13**, 2780–2789 (2014).

32. Koliou, X., Fedonidis, C., Kalpachidou, T. & Mangoura, D. Nuclear import mechanism of neurofibromin for localization on the spindle and function in chromosome congression. *Journal of Neurochemistry***136**, 78–91 (2016).

33. Melloni, G. *et al.* Risk of optic pathway glioma in neurofibromatosis type 1: No evidence of genotype–phenotype correlations in a large independent cohort. *Cancers***11**, (2019).

34. Anastasaki, C., Woo, A. S., Messiaen, L. M. & Gutmann, D. H. Elucidating the impact of neurofibromatosis-1 germline mutations on neurofibromin function and dopamine-based learning. *Human Molecular Genetics***24**, 3518–3528 (2015).

35. Hakimi, M. A., Speicher, D. W. & Shiekhattar, R. The motor protein kinesin-1 links neurofibromin and merlin in a common cellular pathway of neurofibromatosis. *Journal of Biological Chemistry***277**, 36909–36912 (2002).

36. de Schepper, S. *et al.* Neurofibromatosis type 1 protein and amyloid precursor protein interact in normal human melanocytes and colocalize with melanosomes. *Journal of Investigative Dermatology***126**, 653–659 (2006).

37. Couchman, John R., Syndecans: proteoglycan regulators of cell-surface microdomains? *Nature Reviews. Molecular Cell Biology***4** (12); 926-37 (2003). 38.

38. Palaiologou M, Delladetsima I and Tiniakos D. CD138 (syndecan-1) expression in health and disease. *Histol Histopathol***29** (2): 177,89 (2014).

39. Hsueh Y-P, Roberts A.M, Volta M., Sheng M. and Roberts R.G. Bipartite Interaction between Neurofibromatosis Type 1 Protein (Neurofibromin) and Syndecan Trnsmembrane Hepparan Sulfate Proteoglycans. *Journal of Neuroscience* **21** (11): 3764-70 (2001).

40. Yu, H. *et al.* Expression of NF1 pseudogenes. *Human Mutation***26**, 487–488 (2005).

41. LOVD. http://www.lovd.nl/3.0/home.43.

42. Pasmant, E. *et al.* NF1 microdeletions in neurofibromatosis type 1: From genotype to phenotype. *Human Mutation***31**, (2010).

43. Upadhyaya, M. *et al.* 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. *American Journal of Human Genetics***80**, 140–151 (2007).

44. Pinna, V. *et al.* P.Arg1809Cys substitution in neurofibromin is associated with a distinctive NF1 phenotype without neurofibromas. *European Journal of Human Genetics***23**, 1068–1071 (2015).

45. Trevisson, E. *et al.* The Arg1038Gly missense variant in the NF1 gene causes a mild phenotype without neurofibromas. *Molecular Genetics and Genomic Medicine***7**, 1–6 (2019).

46. Koczkowska, M. *et al.* Genotype-Phenotype Correlation in NF1: Evidence for a More Severe Phenotype Associated with Missense Mutations Affecting NF1 Codons 844–848. *American Journal of Human Genetics***102**, 69–87 (2018).

47. Koczkowska, M. *et al.* Clinical spectrum of individuals with pathogenic NF1 missense variants affecting p.Met1149, p.Arg1276, and p.Lys1423: genotype–phenotype study in neurofibromatosis type 1. *Human Mutation***41**, 299–315 (2020).

48. C. Fauth, H. Kehrer-Sawatzki, A. Zatkova, S. Machherndl-Spandl, L. Messiaen, G. Amann, J.A. Hainfellner, K. Wimmer, Two sporadic spinal neurofibromatosis patients with malignant peripheral nerve sheath tumour, European Journal of Medical Genetics,Volume **52**, Issue 6,2009, Pages 409-414.

49. E Hernández-Imaz 1, B Campos, F J Rodríguez-Álvarez, O Abad, G Melean, J Gardenyes, Y Martín, C Hernández-Chico 50. Characterization of NF1 allele containing two nonsense mutations in exon 37 that segregates with neurofibromatosis type 1. Clin Genet 2013 May;**83**(5):462-6.

50. Terzi YK, Sirin B, Hosgor G, Serdaroglu E, Anlar B, Aysun S, Ayter S. Two pathogenic NF1 gene mutations identified in DNA from a child with mild phenotype. Childs Nerv Syst. 2012 Jun;**28**(6):943-6

51. Ferrari L, Mangano E, Bonati MT, Monterosso I, Capitanio D, Chiappori F, Brambilla I, Gelfi C, Battaglia C, Bordoni R, Riva P. Digenic inheritance of subclinical variants in Noonan Syndrome patients: an alternative pathogenic model? Eur J Hum Genet. 2020 Oct;**28(**10):1432-1445

52. Babraham Bioinformatics FastQC. http://www.bioinformatics.babraham.ac.uk/ projects/fastqc.

53. Burrows-Wheeler Aligner. http://bio-bwa.sourceforge.net/index.shtml.

54. Samtools. http://samtools.sourceforge.net/.

55. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010 Sep;**20**(9):1297-303

56. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research***38**, 1–7 (2010).

57. Li, Q. & Wang, K. InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. *American Journal of Human Genetics***100**, 267–280 (2017). https://wintervar.wglab.org/.

58. IntAct: https://www.ebi.ac.uk/intact

59. Ratner N, Miller SJ. A RASopathy gene commonly mutated in cancer: the neurofibromatosis type 1 tumour suppressor. Nat Rev Cancer. 2015 May;**15**(5):290-301 54.

60. McGaughran JM, Harris DI, Donnai D, Teare D, MacLeod R, Westerbeek R, Kingston H, Super M, Harris R, Evans DG. A clinical study of type 1 neurofibromatosis in northwest England. J Med Genet. 1999 Mar;**36**(3):197-203

61. Sharif S, Upadhyaya M, Ferner R, Majounie E, Shenton A, Baser M, Thakker N, Evans DG. A molecular analysis of individuals with neurofibromatosis type 1 (NF1) and optic pathway gliomas (OPGs), and an assessment of genotype-phenotype correlations. J Med Genet. 2011 Apr;**48**(4):256-60

62. Bonneau F, Lenherr ED, Pena V, Hart DJ, Scheffzek K. Solubility survey of fragments of the neurofibromatosis type 1 protein neurofibromin. Protein Expr Purif. 2009 May;**65**(1):30-7

63. Sherekar, M.; Han, S.W.; Ghirlando, R.; Messing, S.; Drew, M.; Rabara, D.; Waybright, T.; Juneja, P.; O'Neill, H.; Stanley, C.B.; et al. Biochemical and Structural Analyses Reveal That the Tumor Suppressor Neurofibromin (NF1) Forms a High-Affinity Dimer. Journal of Biological Chemistry 2020, **295**, 1105–1119

64. Frayling, ian M.; Mautner, V.-F.; van Minkelen, rick; Kallionpaa, roope; aktaş, S.; Baralle, D.; Ben-Shachar, S.; callaway, alison; cox, H.; eccles, D.M.; et al. Breast Cancer Risk in Neurofibromatosis Type 1 Is a Function of the Type of NF1 Gene Mutation: A New Genotype-Phenotype Correlation. *J Med Genet* 2019, **56**, 209–219,

65. Hu, H.-T.; Shih, P.-Y.; Shih, Y.-T.; Hsueh, Y.-P. The Involvement of Neuron-Specific Factors in Dendritic Spinogenesis: Molecular Regulation and Association with Neurological Disorders. **2016**, 2016.

66. Mauda-Havakuk, M.; Shofty, B.; Ben-Shachar, S.; Ben-Sira, L.; Constantini, S.; Bokstein,F. Spinal and Paraspinal Plexiform Neurofibromas in Patients with Neurofibromatosis Type 1: A Novel Scoring System for Radiological-Clinical Correlation. *Am. J. Neuroradiol*. **2017**, *38*, 1869–1875.

67. Taleb FS, Guha A, Arnold PM, et al. Surgical management of cervical spine manifestations of neurofibromatosis type 1: long-term clinical and radiological follow-up in 22 cases. J Neurosurg Spine 2011; 14:356–66

68. Patronas NJ, Courcoutsakis N, Bromley CM, et al. Intramedullary and spinal canal tumors in patients with neurofibromatosis 2: MR imaging findings and correlation with genotype. Radiology 2001; 218:434–42

69. Scott R. Plotkin, Jaishri O. Blakeley, Eva Dombi, Michael J. Fisher, C. Oliver Hanemann, Karin S. Walsh, Pamela L. Wolters, Brigitte C. Widemann. Achieving consensus for clinical trials. The REiNS International Collaboration. Neurology Nov 2013, 81 (21 supplement 1) S1-S5

70. Huson, S.M.; Compston, D.A.; Clark, P.; Harper, P.S. A genetic study of von Recklinghausen neurofibromatosis in south east Wales. I. Prevalence, fitness, mutation rate, and effect of parental transmission on severity. J. Med. Genet. 1989, **26**, 704–711.

71. Friedman, J.M.; Birch, P.H. Type 1 neurofibromatosis: a descriptive analysis of the disorder in 1,728 patients. Am. J. Med. Genet. 1997, **70**, 138–143.

72. Cnossen, M.H.; de Goede-Bolder, A.; van den Broek, K.M.; Waasdorp, C.M.; Oranje, A.P.; Stroink, H.; Simonsz, H.J.; van den Ouweland, A.M.; Halley, D.J.; Niermeijer, M.F. A prospective 10 year follow up study of patients with neurofibromatosis type 1. Arch. Dis. Child. 1998, **78**, 408–412, doi:10.1136/adc.78.5.408.

73. Vogel KS, Klesse LJ, Velasco-Miguel S, Meyers K, Rushing EJ, Parada LF. Mouse tumor model for neurofibromatosis type 1. Science. 1999 Dec 10;**286**(5447):2176-9

74. Costa RM, Yang T, Huynh DP, Pulst SM, Viskochil DH, Silva AJ, Brannan CI. Learning deficits, but normal development and tumor predisposition, in mice lacking exon 23a of Nf1. Nat Genet. 2001 Apr;**27**(4):399-405

75. Long A, Liu H, Liu J, Daniel M, Bedwell DM, Korf B, Kesterson RA, Wallis D. Analysis of patient-specific NF1 variants leads to functional insights for Ras signaling that can impact personalized medicine. Hum Mutat. 2022 Jan;**43**(1):30-41

76. Sabbagh A, Pasmant E, Imbard A, Luscan A, Soares M, Blanché H, Laurendeau I, Ferkal S, Vidaud M, Pinson S, Bellanné-Chantelot C, Vidaud D, Parfait B, Wolkenstein P. NF1 molecular characterization and neurofibromatosis type I genotype-phenotype correlation: the French experience. Hum Mutat. 2013 Nov;**34**(11):1510-8

77. Anastasaki C, Woo AS, Messiaen LM, Gutmann DH. Elucidating the impact of neurofibromatosis-1 germline mutations on neurofibromin function and dopamine-based learning. Hum Mol Genet. 2015 Jun 15;**24**(12):3518-28

78. Colman SD, Williams CA, Wallace MR (1995). Benign neurofibromas in type I neurofibromatosis (NF1) show somatic deletion of the NF1 gene. Nat Genet **11**:90–92.

79. Serra E, Puig S, Otero D, et al. (1997). Confirmation of the double-hit model for the NF1 gene in benign neurofibromas. Am J Hum Genet **6**1:512–519.CrossRefPubMedWeb of ScienceGoogle Scholar

80. Daschner K, Assum G, Eisenbath I, et al. (1997). Clonal origin of tumour cells in a plexiform neurofibroma with LOH in intron 38 and in dermal neurofibromas without LOH of the NF1 gene. Biochem Biophys Res Commun **234**:346–350.

81. Upadhyaya M, Spurlock G, Kluwe L, Chuzhanova N, Bennett E, Thomas N, Guha A, Mautner V. The spectrum of somatic and germline NF1 mutations in NF1 patients with spinal neurofibromas. Neurogenetics. 2009 Jul;**10**(3):251-63

82. Shin J, Padmanabhan A, de Groh ED, Lee JS, Haidar S, Dahlberg S, Guo F, He S, Wolman MA, Granato M, Lawson ND, Wolfe SA, Kim SH, Solnica-Krezel L, Kanki JP, Ligon KL, Epstein JA, Look AT. Zebrafish neurofibromatosis type 1 genes have redundant functions in tumorigenesis and embryonic development. Dis Model Mech. 2012 Nov;**5**(6):881-94.

83. Malavaki CJ, Roussidis AE, Gialeli C, Kletsas D, Tsegenidis T, Theocharis AD, Tzanakakis GN, Karamanos NK. Imatinib as a key inhibitor of the platelet-derived growth factor receptor mediated expression of cell surface heparan sulfate proteoglycans and functional properties of breast cancer cells. FEBS J. 2013 May; **280** (10):2477-89.

84. Dombi E., Baldwin A., Marcus L.J. Activity of selumetinib in neurofibromatosis type 1-related plexiform neurofibromas. N Engl J Med. 2016; 375: 2550-2560

85. Weiss B.D., Wolters P.L., Plotkin S.R.NF106: A neurofibromatosis clinical trials consortium phase II trial of the MEK inhibitor mirdametinib (PD-0325901) in adolescents and adults with NF1-related plexiform neurofibromas. *J Clin Oncol.* 2021; **39**: 797-806

86. Mo J, Anastasaki C, Chen Z, Shipman T, Papke J, Yin K, Gutmann DH, Le LQ. Humanized neurofibroma model from induced pluripotent stem cells delineates tumor pathogenesis and developmental origins. J Clin Invest. 2021 Jan 4;131(1)

9. RINGRAZIAMENTI

Ringrazio infinitamente la mia tutor e correlatrice Prof.ssa Paola Riva, per avermi seguito e supportato nella realizzazione di questo progetto. Le sono profondamente riconoscente per questa esperienza preziosa.

Un sincero grazie al Prof. Eugenio Monti, Direttore del Corso di Dottorato, e alla Prof.ssa Giuseppina De Petro, precedente Direttrice del Corso, per avermi dato la possibilità di svolgere la mia attività di Dottorato, e il Prof. Monti in particolar modo per la sua disponibilità e tempestività ad ogni mia richiesta.

Un particolare ringraziamento va alle colleghe del Laboratorio dell'Università di Milano Marinella, Emanuela, Viviana, Elisa, Luigia e Francesca, per il loro prezioso sostegno e aiuto.

Ringrazio il gruppo di lavoro della Dr.ssa Marica Eoli dell'Istituto Besta di Milano e quello della Dr.ssa Federica Natacci del Policlinico Mangiagalli di Milano per la loro preziosa collaborazione, i pazienti dell'Istituto Besta e del Policlinico e le loro famiglie per la loro disponibilità a partecipare allo studio.

Un Ringraziamento ad Eleonora, Roberta, Tania, Francesca M. e Francesca C. e Cristina del CNR per il loro supporto alle varie fasi di elaborazione del mio progetto di tesi.

Ringrazio il Dirigente scolastico Andrea Carrara per avermi dato questa importante oppurtunità di crescita professionale.

Alle amiche di sempre, e tra queste in particolare ad Anna, grazie per il vostro sostegno e per i momenti felici che abbiamo condiviso.

Alla mia famiglia, ed in particolare ad Adrian, Francesco e Riccardo, senza di voi tutto questo non sarebbe stato possibile. Grazie per avermi sostenuta nei momenti più difficili.