



Turun yliopisto
University of Turku

NEUROFIBROMATOSIS TYPE 1: FROM GENE TO POPULATION AND CANCER

Elina Uusitalo



Turun yliopisto
University of Turku

NEUROFIBROMATOSIS TYPE 1: FROM GENE TO POPULATION AND CANCER

Elina Uusitalo

University of Turku

Faculty of Medicine

Institute of Biomedicine

Cell Biology and Anatomy and

Dermatology and Venereal Diseases

Turku Doctoral Programme of Molecular Medicine (TuDMM)

Supervised by

Professor Juha Peltonen, MD, PhD
Institute of Biomedicine
University of Turku, Turku, Finland

Docent Sirku Peltonen, MD, PhD
Department of Dermatology
University of Turku, Turku, Finland

Reviewed by

Professor Jaakko Kaprio, MD, PhD
University of Helsinki and Institute for
Molecular Medicine Finland (FIMM),
Helsinki, Finland

Docent Sirpa Kivirikko, MD, PhD
Department of Clinical Genetics, Helsinki
University Hospital, Helsinki, Finland

Opponent

Professor Eric Legius, MD, PhD
Department of Human Genetics,
KU Leuven,
Leuven, Belgium

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

ISBN 978-951-29-7155-8 (PRINT)

ISBN 978-951-29-7156-5 (PDF)

ISSN 0355-9483 (Print)

ISSN 2343-3213 (Online)

Painosalama Oy - Turku, Finland 2018

To my family

ABSTRACT

Elina Uusitalo: NEUROFIBROMATOSIS TYPE 1: FROM GENE TO POPULATION AND CANCER University of Turku, Faculty of Medicine, Institute of Biomedicine, Cell Biology and Anatomy & Dermatology and Venereal Diseases, Turku Doctoral Programme of Molecular Medicine (TuDMM). *Annales Universitatis Turkuensis, Medica-Odontologica*, 2018.

Neurofibromatosis type 1 (NF1) is an autosomal dominant syndrome caused by mutations in the large and complex *NF1* tumor suppressor gene on chromosome 17. NF1 is the most frequent hereditary tumor predisposition syndrome. The diagnosis of NF1 is usually based on clinical findings, such as tumors of the peripheral nervous system called neurofibromas and hyperpigmentary abnormalities such as café-au-lait pigment spots. This thesis investigated the molecular diagnostics, epidemiology, and cancer biology associated with the NF1 cancer syndrome.

In this thesis, a new method for *NF1* molecular diagnostics exploiting next-generation sequencing was developed. This is important, because mutation analysis is currently not available for every patient, and in the majority of the patients, the diagnosis is still merely based on clinical manifestations.

A national NF1 cohort of 1,404 patients was used in this thesis, and a retrospective register-based total population study was carried out to evaluate the epidemiology and cancer incidence of NF1 in Finland. The results revealed that NF1 incidence is higher than previously accepted. The results with a birth incidence of 1/2,000 challenge the generally accepted NF1 incidence of ~1/3,000. The NF1 cancer incidence was studied with data from the Finnish Cancer Registry. A five-fold increase in cancer incidence was observed, which is the highest cancer incidence reported so far. In addition, cancers in the NF1 patients have a worse prognosis than the corresponding cancers in the general population. In NF1 patients the risk for breast cancer is also elevated, particularly under the age of 40. NF1-related breast cancer has poor prognosis, which is not solely explained by occurrence at young age or by histopathological type. Our results suggest that *NF1* mutations are an independent factor contributing to low survival of patients with breast cancer. Active surveillance of NF1 patients and awareness of the NF1-related cancer risk are needed for early detection of the tumors and improved prognosis.

Keywords: Neurofibromatosis type 1, next-generation sequencing, incidence, cancer, breast cancer, survival

TIIVISTELMÄ

Elina Uusitalo: NEUROFIBROMATOOSI TYYPPI 1: GEENISTÄ YLEISYYTEEN JA SYÖPÄÄN. Turun yliopisto, Lääketieteellinen tiedekunta, Biolääketieteen laitos, Solubiologia ja anatomia & Iho- ja sukupuolitautioppi, Turun molekyyli­lääketieteen tohtoriohjelma (TuDMM). Annales Universitatis Turkuensis, Medica-Odontologica, 2018.

Neurofibromatoosi tyyppi 1 (NF1) on vallitsevasti periytyvä syövälle altistava oireyhtymä, joka aiheutuu mutaatiosta kromosomissa 17 sijaitsevassa *NF1*-geenissä. *NF1*-geeni on hyvin suurikokoinen ja sen koko alueella havaitaan haitallisia mutaatioita. NF1 esiintyy noin 1/3,000 syntyvästä lapsesta ja diagnoosi perustuu yleensä kliinisiin kriteereihin. Tärkeimpiä kliinisiä kriteereitä ovat ihon vaaleanruskeat maitokahviläiskät sekä neurofibroomat eli hermon hyvänlaatuiset sidekudoskasvaimet. NF1 oireyhtymälle tyypillisiä syöpätyyppejä ovat aivojen ja ääreishermoston kasvaimet. Tämä väitöskirja keskittyy erityisesti NF1:n diagnostiikkaan, epidemiologiaan sekä syöpäbiologiaan.

Väitöskirjatutkimuksessa kehitettiin uusi menetelmä NF1:n molekyyli­diagnoosiin. Menetelmässä hyödynnettiin kohdennetusti uuden sukupolven sekvensointimenetelmiä. Uusi menetelmä on tärkeä, sillä tällä hetkellä mutaatio­analyysiä ei ole tarjolla kaikille potilaille ja diagnoosi perustuu yleensä kliinisiin oireisiin, joiden kehittyminen voi viedä vuosia.

Väitöskirjassa koottiin kansallinen 1,404 NF1 potilaan kohortti. Väitöskirjatyön tulokset osoittavat, että NF1 on selvästi aiemmin luultua yleisempi sairaus, sillä sen ilmaantuvuudeksi saatiin tutkimuksessa 1/2,000. Potilaiden syöpäilmaantuvuutta sekä -kuolleisuutta analysoitiin Suomen Syöpärekisterin tietojen avulla. Tutkimuksen tulokset osoittivat että NF1 potilailla syövän ilmaantuvuus on viisinkertainen muuhun väestöön verrattuna. Tämän lisäksi havaittiin, että NF1 syöpäpotilailla on muuta väestöä huonompi ennuste. NF1:lle tyypillisten kasvainten lisäksi havaittiin kohonnut rintasyöpäriski jo nuorilla, sekä huonompi viiden vuoden elossaolo-osuus kuin verrokkipotilailla. Paras keino vähentää NF1 potilaiden syöpä­kuolleisuutta on lisätä tietoisuutta syöpäriskistä sekä aktiivinen seuranta syöpien löytämiseksi varhaisessa vaiheessa. *NF1*-mutaatiot ovat yleisiä myös vertailu­väestön syövissä. Siksi NF1 potilaiden syöpien tutkiminen voi auttaa ymmärtämään *NF1*-geenin osuutta syövän synnyssä laajemmin.

Avainsanat: Neurofibromatoosi tyyppi 1, uuden sukupolven sekvensointimenetelmät, ilmaantuvuus, syöpä, rintasyöpä, eloonjäämisanalyysi

TABLE OF CONTENTS

ABSTRACT	4
TIIVISTELMÄ.....	5
ABBREVIATIONS.....	8
LIST OF ORIGINAL PUBLICATIONS	10
1 INTRODUCTION	11
2 REVIEW OF LITERATURE.....	12
2.1 Neurofibromatosis type 1 syndrome	12
2.1.1 NF1 symptoms	12
2.1.2 NF1 diagnosis.....	14
2.1.3 Other NF1 subtypes	15
2.2 The <i>NF1</i> gene	16
2.2.1 Neurofibromin	18
2.3 <i>NF1</i> mutations	20
2.4 <i>NF1</i> mutation analysis	23
2.4.1 Next-generation sequencing.....	24
2.5 Epidemiology of NF1	26
2.5.1 NF1 prevalence and incidence	26
2.5.2 Mortality in NF1.....	27
2.5.3 Syndromes with phenotypic overlap with NF1.....	29
2.6 Cancer	31
2.6.1 Cancer types	31
2.6.2 Biology of cancer.....	31
2.6.3 NF1 gene alterations in cancer	32
2.7 Cancer in NF1	33
2.7.1 NF1 cancer risk.....	33
2.7.2 MPNST	34
2.7.3 Brain tumors	35
2.7.4 Other cancers with possible risk.....	36
2.8 Breast cancer and NF1	37
2.8.1 Breast cancer	37
2.8.2 Breast cancer risk factors	37
2.8.3 Molecular profiling of breast cancer	38
2.8.4 Breast cancer incidence in NF1.....	39
3 AIMS OF THE STUDY.....	41
4 MATERIALS AND METHODS.....	42
4.1 DNA sampling (I).....	42

4.2	Sequence capture and sequencing (I)	42
4.3	Data collection and patient ascertainment (II-IV)	45
4.4	Statistical analyses (II-IV)	46
4.5	Breast cancer samples (IV)	47
4.6	Ethics (I-IV)	49
5	RESULTS	50
5.1	Mutation analysis of the <i>NF1</i> gene (I).....	50
5.1.1	Sequencing	50
5.1.2	Identification of mutations	50
5.2	Formation of the population-based NF cohort and NF1 incidence in Finland (II-IV).....	53
5.2.1	Search for NF1 patients (II)	53
5.2.2	Higher NF1 incidence than previously reported (II).....	56
5.3	Mortality (II-IV)	56
5.4	Cancer outcomes (III)	57
5.5	Breast cancer in NF1 (IV).....	58
6	DISCUSSION.....	61
6.1	New <i>NF1</i> mutation analysis method exploiting NGS (I).....	61
6.2	Total population-based NF1 cohort (II-IV)	62
6.3	NF1 birth incidence in Finland (II)	63
6.4	Interpretation of the NF1 cancer and mortality study results (II-III)..	64
6.5	Breast cancer (III-IV)	65
6.6	Strengths and limitations (I-IV)	66
6.7	Suggestions for future research.....	68
7	CONCLUSIONS	70
	ACKNOWLEDGEMENTS.....	71
	REFERENCES.....	73
	ORIGINAL PUBLICATIONS	95

ABBREVIATIONS

AKT	Protein kinase B	GRB2	xxx Growth factor receptor-bound protein 2
ATM	Ataxia-telangiectasia mutated	GRD	GAP-related domain
BMI	Body mass index	GTP	Guanosine triphosphate
bp	Base pair	H&E	Hematoxylin and eosin
BRCA1	Breast cancer 1	HER2	human epidermal growth factor receptor 2
BRCA2	Breast cancer 2	HILMO	Hospital Discharge Register
CDKN2A	Cyclin-dependent kinase inhibitor 2A	ICD	International Classification of Diseases
cDNA	Complementary DNA	ICGC	International Cancer Genome Consortium
CHEK2	Checkpoint kinase 2	JMML	Juvenile myelomonocytic leukemia
chr17	Chromosome 17	kb	Kilo base
CI	Confidence interval	kDa	KiloDalton
CK	Cytokeratin	LZTR1	Leucine Zipper Like Transcription Regulator 1
CMMRD	Constitutional mismatch repair deficiency	Mb	Mega base
DAB	3,3'- diaminobenzidine	MEK	Mitogen-activated protein kinase kinase
dbSNP	The Single Nucleotide Polymorphism database	MLPA	Multiplex ligation dependent probe amplification
DNA	Deoxyribonucleic acid	MMR	Mismatch repair
EGFR	Epidermal growth factor receptor	MPNST	Malignant peripheral nerve sheath tumor
emPCR	Emulsion PCR	MRI	Magnetic resonance imaging
ER	Estrogen receptor	mRNA	Messenger RNA
ERK	Extracellular signal regulated kinase	mTOR	Mechanistic target of rapamycin
<i>EVI2A/B</i>	Ecotropic viral integration site	NF1	Neurofibromatosis type 1
FISH	Fluorescent <i>in situ</i> hybridization	<i>NF1</i>	human <i>NF1</i> gene
GABA	gamma- Aminobutyric acid	<i>Nf1</i>	mouse <i>Nf1</i> gene
GAP	GTPase activating protein	NF2	Neurofibromatosis type 2
gDNA	genomic DNA	NGS	Next-generation sequencing
GDP	Guanosine diphosphate		
GIST	Gastrointestinal stromal tumor		

Abbreviations

NIH	National Institutes of Health	SIR	Standardized incidence ratio
OMG	Oligodendrocyte myelin glycoprotein	SMARCB1	SWI/SNF Related, Matrix Associated, Actin Dependent Regulator Of Chromatin, Subfamily B, Member 1
OPG	Optic pathway glioma		
PALB2	partner and localizer of BRCA2		
PCR	Polymerase chain reaction	SMR	Standardized mortality ratio
PDGFRA	Platelet-derived growth factor receptor alpha	SNP	Single nucleotide polymorphism
PGM	Personal Genome Machine	SOS1	SOS Ras/Rac Guanine Nucleotide Exchange Factor 1
PH domain	Pleckstrin homology		
PI3K	Phosphatidylinositol-4,5-bisphosphate 3-kinase	SPRED1	Sprouty-related, EVH1 domain-containing protein 1
PMR	Proportionate mortality ratio	SUZ12	SUZ12 Polycomb Repressive Complex 2 Subunit
PR	Progesterone receptor	TCGA	The Cancer Genome Atlas
RAF	Rapidly Accelerated Fibrosarcoma	TP53	Tumor protein p53
RR	Rate ratio		

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following original publications, which are referred to in the text by the Roman numerals I-IV:

- I Elina Uusitalo, Anna Hammamais, Elina Palonen, Ville-Veikko Mäkelä, Roope Kallionpää, Eeva-Mari Jouhilahti, Minna Pöyhönen, Annika Brandt, Juhani Soini, Juha Peltonen, Sirkku Peltonen. Neurofibromatosis type 1 gene mutation analysis using sequence capture and high-throughput sequencing. *Acta Derm Venereol* 2014; 94(6): 663–6.
- II Elina Uusitalo*, Jussi Leppävirta*, Anna Koffert, Sakari Suominen, Jussi Vahtera, Tero Vahlberg, Minna Pöyhönen, Juha Peltonen, Sirkku Peltonen. Incidence and mortality of neurofibromatosis. A total population study in Finland. *J Invest Dermatol* 2015; 135(3), 904–6.
- III Elina Uusitalo, Matti Rantanen, Roope A. Kallionpää, Minna Pöyhönen, Jussi Leppävirta, Heli Ylä-Outinen, Vincent M. Riccardi, Eero Pukkala, Janne Pitkäniemi, Sirkku Peltonen, Juha Peltonen. Distinctive cancer associations in patients with neurofibromatosis type 1. *J Clin Oncol* 2016; 34(17), 1978–86.
- IV Elina Uusitalo, Roope A. Kallionpää, Samu Kurki, Matti Rantanen, Janne Pitkäniemi, Pauliina Kronqvist, Pirkko Härkönen, Riikka Huovinen, Olli Carpen, Minna Pöyhönen, Sirkku Peltonen, Juha Peltonen. Breast cancer in neurofibromatosis type 1: overrepresentation of unfavourable prognostic factors. *BR J Cancer* 2017; 116(2), 211-7.

* equal contribution

The original publications are reproduced with the permission of the copyright owners.

1 INTRODUCTION

In the 1860s, an Austrian monk named Gregor Mendel introduced his famous theory of inheritance based on his experimental work in garden. His experiments with pea plants showed that the inheritance of certain traits follows particular patterns. Earlier, most people believed that inheritance was due to a diluted blending of parental ‘substances,’ just like mixing blue and yellow paint will produce a green color. To date, we know that cells store their hereditary information in the form of double-stranded molecules of DNA. In eukaryotic cells, the DNA is packaged into a set of chromosomes, and chromosomes carry genes, which are the functional units of heredity.

Monogenic disorders, also often referred to as Mendelian, are caused by a single defective gene. Dominant diseases are monogenic disorders that involve damage to only one gene copy, and thus an abnormal gene from one parent is sufficient to cause the disease. This happens even when the matching gene from the other parent is normal. However, also healthy parents may have an affected child if a new mutation occurs spontaneously in the germ line cells of one or the other parent.

Neurofibromatosis type 1 (NF1) is one of the most common autosomal dominant syndromes but yet a rare disease. It is caused by mutations in the large *NF1* gene in chromosome 17 (Wallace et al., 1990, Xu et al., 1990). The diagnosis of NF1 is usually based on clinical findings, but genetic testing is needed to confirm the diagnosis in several situations. Rare diseases, in general, are poorly understood; and there are many questions unanswered.

In Finland, the use of comprehensive national health care registries provides a unique opportunity to study diseases with low incidence. Record-keeping in general has a long tradition in Finland, for example, the first nation-wide, computerized disease register, the Finnish Cancer Registry was founded in 1952.

In the present study, a new method for NF1 molecular diagnostics was developed. In addition, the epidemiology and cancer incidence of NF1 in Finland was evaluated with a national NF1 cohort and retrospective register-based total population study. The intention of this study was also to advance the knowledge about breast cancer related to NF1.

2 REVIEW OF LITERATURE

2.1 Neurofibromatosis type 1 syndrome

Neurofibromatosis type 1 (NF1) is an autosomal dominant tumor predisposition syndrome. NF1, previously known as von Recklinghausen's disease after Friedrich Daniel Von Recklinghausen, is caused by mutations in the *NF1* tumor suppressor gene on chromosome 17 (Wallace et al., 1990, Xu et al., 1990). NF1 is one of the most common monogenic syndromes. The autosomal dominant inheritance pattern of NF1 has been confirmed for many years, meaning that NF1 patients have a 50/50 chance of an affected child. Besides that, the mutation rate of the *NF1* gene is one of the highest observed resulting in about half of the NF1 cases caused by a *de novo* mutation (Huson et al., 1989).

2.1.1 *NF1* symptoms

NF1 is a multisystem disease causing a varying combination of symptoms. The symptoms can impact essentially every organ system. However, as a neurocutaneous syndrome, its hallmark features involve particularly the skin, the central nervous system, and the peripheral nervous system. The major defining features of NF1 are hyperpigmentary abnormalities shown in Figure 1 such as skinfold freckling, café-au-lait macules, and iris Lisch nodules (Huson, 2008). A significant feature of this syndrome is the formation of multiple tumors of the peripheral nervous system called neurofibromas (Figure 1B) (Jouhilahti et al., 2011). These features are present in the majority of the NF1 patients (Peltonen and Pöyhönen, 2012).



Figure 1 The hallmark symptoms of NF1. A) Skinfold freckles; B) Café-au-lait macules (asterisk) and neurofibromas (arrowhead); and C) Lisch nodules (iris hamartomas, arrowhead). Photo courtesy of: A) Sirkku Peltonen, B) Eeva-Mari Jouhilahti, and C) Vesa Aaltonen.

NF1 is a syndrome first manifesting itself during childhood. The first signs of the syndrome are often café-au-lait macules of the skin, which are present in 95% of the patients by the age of 1 year (DeBella et al., 2000) and in 99% of the patients by adulthood (Huson, 2008, McGaughran et al., 1999). Café-au-lait macules are hyperpigmented areas of the skin, which often become larger and more numerous during childhood but may fade during adulthood (Shah, 2010). Solitary café-au-lait macules are also common in the general population. A study in the United States found café-au-lait macules in 2% of all newborns (Kanada et al. 2012). Consistent with the current knowledge, they were most common in African-American children.

Skinfold freckling occurs in non-sun-exposed skin and usually appears in NF1 patients later than the café-au-lait macules (Evans et al., 2017). Also the appearance of iris Lisch nodules, which are benign hamartomas, typically occurs early in childhood. Over 90% of adult NF1 patients have Lisch nodules (Evans et al., 2017, Peltonen and Pöyhönen, 2012).

Other symptoms related to the syndrome include cognitive impairment, skeletal abnormalities, and different malignancies (Gutmann et al., 2017, Peltonen and Pöyhönen, 2012). Many NF1 children have cognitive deficits, learning disabilities, and behavioral problems (Lehtonen et al., 2013). Cognitive deficits in NF1 children are an essential burden and weak performance in verbal and nonverbal tasks and impaired global language belong to the characteristic features (Krab et al., 2008, Mautner and Boltshauser, 2008). Moreover, NF1 children frequently have a general delay in development and problems with attention and impulse control (Mautner and Boltshauser, 2008). A high prevalence of social behavioral problems and autism spectrum disorders has also been reported (Garg et al., 2013, van der Vaart et al., 2016).

NF1 children can also develop bone manifestations such as pseudarthrosis and scoliosis (Elefteriou et al., 2009). Long bone dysplasia is seen in 3–4% of the NF1 patients (Friedman and Birch, 1997). Also decreased bone mineral density and shortness of stature are reported to occur frequently in NF1 patients (Elefteriou et al., 2009, Kuorilehto et al., 2005).

The NF1 patients have an increased risk for developing both benign and malignant tumors. Dermal neurofibromas are present in over 95% of adult NF1 patients and are a hallmark of the disease (Huson, 2008). Neurofibromas are benign peripheral nerve sheath tumors at cutaneous or subcutaneous locations. They are generally less than 3 cm in diameter, numerous, and are not reported to progress to malignancy. The number of cutaneous neurofibromas in adult NF1 patients varies from few to thousands (Peltonen and Pöyhönen, 2012). Neurofibromas are a major

cause of morbidity in adult NF1 patients mainly because of the number and visibility of these tumors (Page et al., 2006).

Cutaneous neurofibromas rarely occur before the age of ten years, however a child with NF1 can have a plexiform neurofibroma visible from birth at any location of the body (Mautner and Boltshauser, 2008). These tumors may also become apparent later in life. Plexiform neurofibromas are larger peripheral nerve sheath tumors that involve long nerve segments and can form a large disfiguring mass. They produce morbidity by causing disfigurement, pain, impairment of nerve function, or potentially a progression to malignancy (Evans et al., 2002, Kim et al., 2017). Whole-body MRI reveals plexiform tumors in over 50% of the NF1 patients, while about 25–30% of the patients have a visible or symptomatic plexiform neurofibroma (Nguyen et al., 2011).

NF1 is associated with benign neurofibromas but also with malignancies of several types. In particular, optic pathway gliomas (Listernick et al., 1997), and malignant peripheral nerve sheath tumors (MPNSTs) (Evans et al., 2002, Evans et al., 2011) are common malignant tumors observed in NF1 patients.

Interdisciplinary medical care is often needed because of the variety of characteristic clinical features associated with NF1. NF1 also shows major phenotypic variability of symptoms. The degree of severity is unpredictable among patients, even those carrying identical *NF1* germline mutations (Viskochil, 2002). An individual with mild clinical symptoms can have a child with a more severe phenotype or vice versa (Easton et al., 1993, Riccardi and Lewis, 1988).

2.1.2 *NF1* diagnosis

The diagnosis of NF1 is usually based on clinical findings. The diagnostic criteria of NF1 was agreed at the NIH 1987 NF Consensus Conference (Stumpf et al., 1988). At least two of the diagnostic criteria listed in Table 1 are needed for clinical diagnosis of NF1. Most important of these, café-au-lait macules, skinfold freckles, and neurofibromas should be visible on the skin. The diagnostic criteria has been shown to be specific and sensitive for adults with NF1 (Gutmann et al., 1997, Kluwe et al., 2004). However, clinicians often face situations where there are some NF1 symptoms but not sufficient for a clinical diagnosis. Diagnosing small children is often challenging, because about half of sporadic NF1 cases fail to fulfill the NIH Diagnostic Criteria by the age of 1 year (DeBella et al., 2000). Since NF1 is a multiorgan disease with frequent complications from various organ systems, the correct early diagnosis is essential. At present, genetic testing is

recommended to confirm NF1 diagnosis, particularly in children fulfilling only pigmentary features of the diagnostic criteria (Evans et al., 2017). The same holds true for adults with an atypical clinical presentation (Peltonen and Pöyhönen, 2012).

Table 1 NF1 diagnostic criteria. The NIH consensus development conference statement diagnostic criteria for NF1 are met in an individual who has two or more of the following:

- **Six or more café-au-lait macules of over 5 mm in greatest diameter in prepubertal individuals and over 15 mm in greatest diameter in postpubertal individuals.**
- **Two or more neurofibromas of any type or one plexiform neurofibroma.**
- **Freckling in the axillary or inguinal regions.**
- **Optic glioma.**
- **Two or more Lisch nodules (iris hamartomas).**
- **A distinctive osseous lesion such as sphenoid dysplasia or thinning of the long bone cortex with or without pseudarthrosis.**
- **A first-degree relative (parent, sibling, or offspring) with NF1 by the above criteria.**

2.1.3 Other NF1 subtypes

Segmental NF1

Segmental or mosaic NF1 is used to describe the patients with an *NF1* constitutional mutation present in only a small population of cells. Therefore, the disease features are limited to the affected area, which varies from a narrow strip to one-half of the body (Ruggieri and Huson, 2001). Segmental NF1 is caused by a postzygotic mutation of the *NF1* gene during embryonic development (Listernick et al., 2003, Ruggieri and Huson, 2001). The most common initial finding in these patients is unilateral presentation of pigmentary changes, and most of these patients seek medical care because of an unusual appearance of the skin (Huson, 2008, Lara-Corrales et al., 2017). Patients with segmental NF1 are less likely to have a severe disease and also a lower risk to have children with NF1 (Ruggieri and Huson, 2001). It should be noted that the segmental NF1 may explain the absence of detectable *NF1* mutation in blood.

Spinal NF1

Spinal neurofibromatosis is a rare form of NF1, which causes multiple spinal nerve root neurofibromas as a principal feature. The severe subtype is characterized by bilateral neurofibromas of all spinal roots. The patients have some pigmentary features of NF1 but an absence of dermal neurofibromas (Huson, 2008, Messiaen et al., 2003). Only around 100 cases with this form of NF1, both sporadic and familial, have been reported (Poyhonen et al., 1997, Ruggieri et al., 2015). It has been suggested that individuals with the severe familial subtype more frequently carry an *NF1* missense or splicing mutation (Kluwe et al., 2003, Messiaen et al., 2003).

2.2 The *NF1* gene

The *NF1* gene (OMIM 613113) is located on the long arm of chromosome 17 at position 11.2 (17q11.2). The gene spans 283 kb of genomic DNA in the interval of 31,094,927-31,377,677 (Genome Reference Consortium, GRCh38) on Chr17. The large size of the gene is consistent with the very high spontaneous mutation rate. The gene comprises as much as 57 constitutive and 4 alternatively spliced exons 9a, 10a-2, 23a, and 48a (Figure 2) (Upadhyaya, 2008, 2010). The GAP-related domain (GRD), encoded by exons 20-27a, is the most highly conserved region of the gene (Upadhyaya, 2010). The *NF1* gene has two large introns: 1 and 27b. The large 60 kb intron 27b contains three small unrelated genes *EVI2A*, *EVI2B* and *OMG*, which are transcribed from the opposite strand and thus in the reverse direction than the *NF1* gene (Upadhyaya, 2008, Viskochil et al., 1990). Each of these genes has two exons.

The human *NF1* gene was cloned simultaneously during the 1990s by Wallace et al. and Viskochil et al. (Viskochil et al., 1990, Wallace et al., 1990). The *NF1* gene is evolutionarily conserved. The remarkably high degree of *NF1* sequence conservation between human and mouse was first described by Bernards et al. (Bernards et al., 1993). They found more than 98% sequence identity between the neurofibromin amino acid sequences of these two species (Bernards et al., 1993).

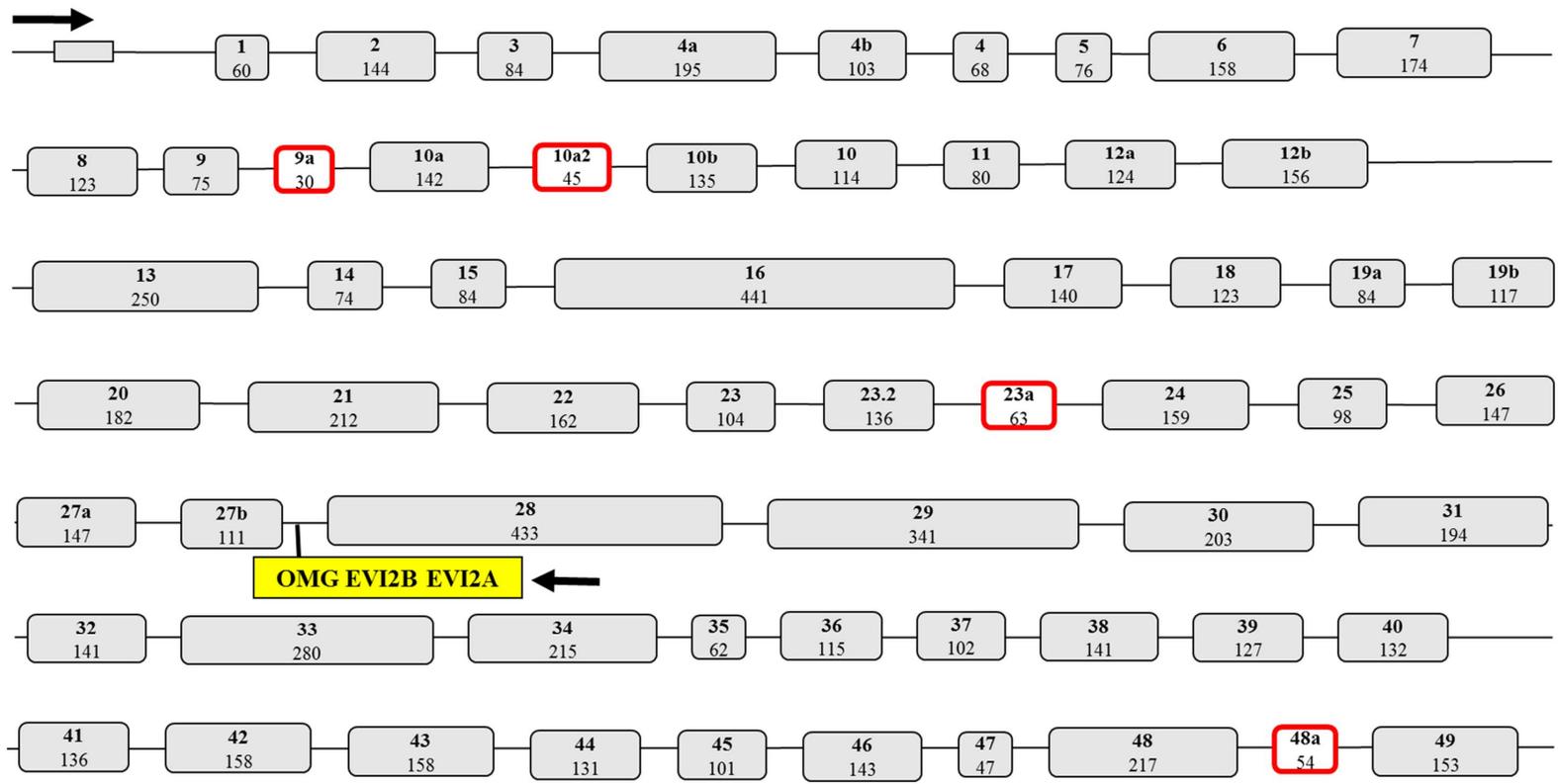


Figure 2 The structure of the *NFI* gene. The exons are numbered according to the accepted *NFI* gene nomenclature, and the size in base pairs is indicated below the exon number. The core promoter of the gene is shown as a rectangle box before exon 1, and the alternatively spliced exons are shown as red boxes. *OMG*, *EVI2B*, *EVI2A* are the three embedded genes in intron 27b, which are transcribed in the opposite direction. Modified from Upadhyaya et al., 2010.

NF1 pseudogenes

Several *NF1*-like sequences, i.e., pseudogenes have been found throughout the human genome at least in chromosomes 2, 12, 14, 15, 18, 21, and 22 (Kehrer-Sawatzki et al., 1997, Legius et al., 1992, Luijten et al., 2001, Purandare et al., 1995, Suzuki et al., 1994). Many of these pseudogenic sequences display significant homology of over 90% to the *NF1* sequence, however all of them have various inactivating nucleotide substitutions, insertions, or deletions (Upadhyaya, 2008). It is thought that the *NF1* pseudogenes initially arose during evolution by duplication and transposition of parts of the functional *NF1* gene (Luijten et al., 2001). The highly homologous pseudogene sequences can interfere with gDNA-based sequencing methods.

2.2.1 Neurofibromin

Tumor suppressor genes, such as *NF1*, encode proteins that are responsible for regulating cell division, and thus a loss-of-function mutation of such genes can contribute to tumorigenesis. The *NF1* gene gives rise to a 12-kilobase mRNA transcript encoding the neurofibromin protein, whose activity is impaired in patients with NF1. Neurofibromin is a giant 2,818 amino acid protein that is ubiquitously expressed at low concentrations in most tissues and during development but expressed at the highest levels in cells of the central nervous system and brain (Daston and Ratner, 1992, Daston et al., 1992, Wallace et al., 1990). The cellular distribution of neurofibromin varies by tissue and cell type. Neurofibromin was originally discovered as a cytosolic protein (DeClue et al., 1991) but, now, it is found to be associated with at least the plasma membrane, the endoplasmic reticulum, and to co-localize with mitochondria (Malhotra and Ratner, 1994, Nordlund et al., 1993, Roudebush et al., 1997).

Xu et al. found that *NF1* gene encodes a cytoplasmic GAP-like protein (Xu et al., 1990). The catalytic Ras-specific GTPase activating protein (RasGAP) domain resides in a central portion of neurofibromin, which is termed the GAP-related domain (GRD) (Xu et al., 1990). Neurofibromin is recognized as a tumor suppressor protein, and its best characterized function is in the down-regulation of functionally active Ras proteins or in the regulation of cellular levels of activated Ras proteins (Figure 3). Neurofibromin is reported to interact with the major Ras isoforms: H-, K-, and N-Ras (Ahmadian et al., 1997). Ras proteins exist in two forms in the cell with the inactive form bound to GDP, whereas a very small number of Ras proteins contains a bound GTP and is metabolically active. Only this activated form can upregulate many of the downstream effector proteins.

Neurofibromin downregulates the biological activity of the small guanine nucleotide binding of Ras by accelerating the conversion of active Ras-bound GTP into its inactive Ras-GDP state (Cichowski and Jacks, 2001, Johannessen et al., 2005, Upadhyaya, 2010).

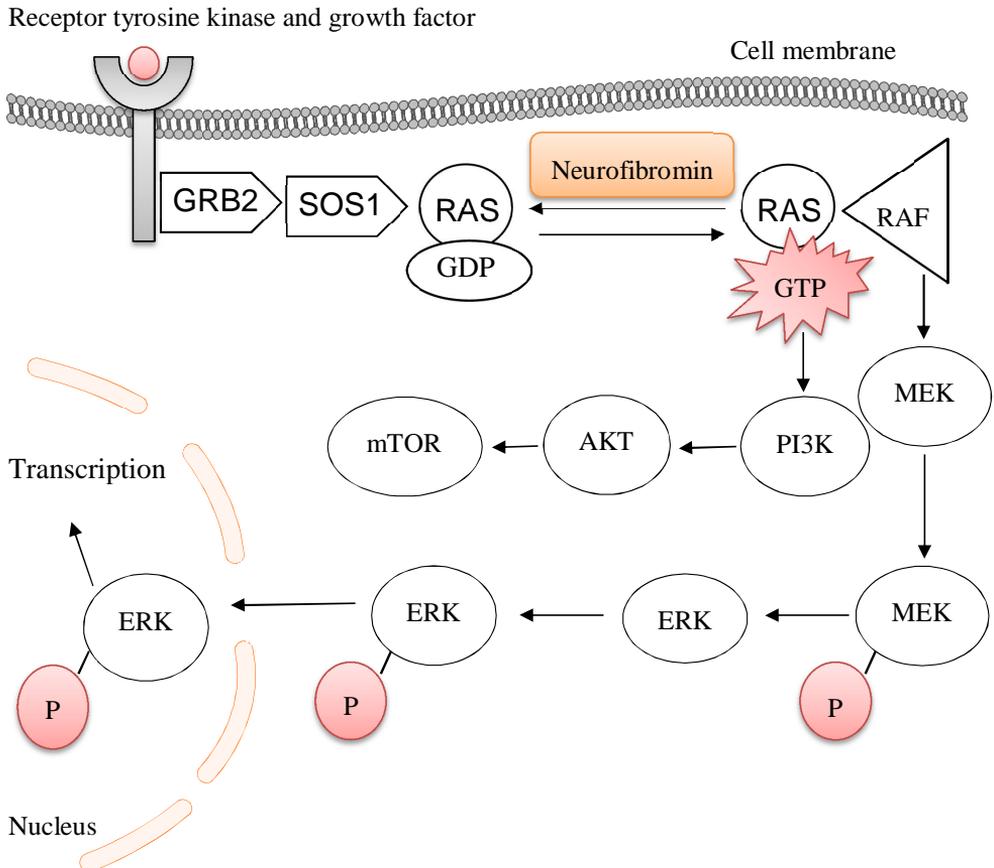


Figure 3 In the absence of neurofibromin and the negative regulation of RAS proteins, the GTP-bound RAS levels are increased. Signaling pathways downstream of RAS show enhanced activation. These pathways include the MEK-ERK signaling cascade downstream of RAF and the PI3K pathway. Modified from Brems et al., 2009.

Neurofibromin as a RasGAP is located at a central position of several signaling pathways (Figure 3). The role that neurofibromin plays in the regulation of the Ras/MAPK pathway has an effect on, for example, cellular growth, neural development, and cognition. It has also been shown that neurofibromin regulates

the mTOR pathway in a Ras/PI3K-dependent manner (Johannessen et al., 2005). The PI3K/AKT/mTOR signaling pathway protects cells from apoptosis, but in the absence of functional neurofibromin, it is activated, which results in an increase in cell proliferation and survival. Moreover, both the PI3K/AKT and Raf/MAPK pathways can activate mTOR signaling with mTOR pathway activation occurring in the absence of growth factors (Johannessen et al., 2005). Increased Ras activity may also be associated with an NF1-mediated learning deficiency due to long term potentiation impairment resulting from increased GABA-mediated inhibition (Shilyansky et al., 2010).

The RasGAP activity of neurofibromin has been widely studied in biochemical and structural detail, but there are also other domains in neurofibromin, less intensively studied. For example, Sec14 and PH domains bind to phospholipids, but the precise function of the Sec14-PH module remains unclear (D'Angelo et al., 2006, Welti et al., 2007). A number of other proteins than Ras have also been identified to interact with neurofibromin. Such are, for example, tubulin (Bollag et al., 1993), kinesin-1 (Hakimi et al., 2002), syndecan (Hsueh et al., 2001), caveolin-1 (Boyanapalli et al., 2006), and Spred1 (Stowe et al., 2012, Dunzendorfer-Matt et al., 2016, Hirata et al., 2016)

Alternative splicing of NF1 gene and neurofibromin isoforms

Neurofibromin is a 327 kDA protein and has several alternative isoforms through the inclusion of alternative spliced exons (Figure 2). The ubiquitously expressed type II isoform includes an alternatively spliced exon 23a, and thus an additional 21 amino acids inserted in the GTPase activating protein-related domain (Suzuki et al., 1991). An isoform abundantly expressed in muscle contains an alternatively spliced exon 48a and thus an additional 54-base pair in-frame insertion (Gutman et al., 1993). In addition, there is an isoform highly expressed in the central nervous system containing an additional exon 9a with an insertion of 30 base pairs (Danglot et al., 1995).

2.3 NF1 mutations

A mutation is a permanent change in DNA resulting from random accidents and errors in the storage and copying of genetic information.

Homozygous deletion of the *Nf1* gene is lethal in mice indicating that neurofibromin is essential for development (Jacks et al., 1994). In NF1 patients, the *NF1* gene is nonfunctional due to gene mutation. *NF1* is a tumor suppressor gene, accordingly heterozygous loss of function mutations of the *NF1* gene result

in tumorigenesis. The loss of neurofibromin functionality due to mutation in the *NF1* gene results in sustained levels of active Ras-GTP and thus to a prolonged activation of the Ras/Raf/MAPK signaling pathway. The overall results are increased cell proliferation and loss of growth control (Upadhyaya, 2010).

The mutation rate of the *NF1* gene (~1:10,000) is among the highest known for any gene in humans (Friedman 2017). This is partly explained by the large size of the gene, but the cause of this unusually high mutation rate is mostly unknown. About half of the NF1 cases result from *de novo* mutation, i.e., healthy parents have a child with NF1 (Huson et al., 1989). According to The Human Gene Mutation Database (HGMD), 2,689 different mutations of the *NF1* gene have been published so far. More than 2,800 different NF1 pathogenic variants have been identified at the University of Alabama at Birmingham (UAB) cohort, with only 31 unique pathogenic variants present in $\geq 0.5\%$ of all unrelated individuals (Koczkowska et al., 2018). There is no evidence of mutation hot spots across the gene. Mutations in the *NF1* gene are spread over the entire coding region and include a very diverse spectrum of mutation types, such as total gene deletions and intragenic copy number changes. Deletions/duplications involving one or several exons, frameshifts, nonsense mutations, splice mutations, missense mutations and in-frame deletions or duplications involving one or several codons have been described (Messiaen and Wimmer, 2008). Most of the *NF1* germline mutations are predicted to result in a truncated transcript and protein (Messiaen and Wimmer, 2008).

Regarding the genetic epidemiology of NF1, no differences among ethnic backgrounds have been published. Two single nucleotide polymorphisms (SNPs) in the *NF1* gene region have enrichment in the Finnish population and thus have a higher allele frequency (Odds ratio >2) (sisuproject.fi). These SNPs have however uncertain clinical significance.

NF1 genotype/phenotype correlations

There is no evidence from a family transmitting the *NF1* mutation from one generation to the next through someone who carries the mutation but does not have an NF1 phenotype. Based on the current knowledge, NF1 penetrance is virtually complete after childhood, which means that all individuals with an *NF1* gene mutation have some phenotypic traits of the syndrome (Viskochil, 2002, Huson et al., 1989).

The numerous attempts to find associations between a specific *NF1* gene mutation and a characteristic clinical phenotype have mostly been unsuccessful. There are examples how monozygotic twins share some NF1 features, but more distant affected relatives in the same family often exhibit a more variable clinical

phenotype (Easton et al., 1993, Szudek et al., 2002). This suggests that the type of the mutation of the *NF1* gene itself plays only a minor role in the clinical phenotype and severity or outcome of the disease in affected individuals. High levels of intra- and interfamilial clinical variability are observed in many NF1 pedigrees, even though all affected individuals carry identical *NF1* mutations. One confounding problem is that the phenotype is determined by the constitutional mutation, by the patient's age, the timing of second hit mutations in different cells and tissues, potential mosaicism in founder patients as well as modifying and environmental factors (Rojnueangnit et al., 2015).

To date, only few clinically significant genotype/phenotype correlations in *NF1* have been identified. The first is the so called *NF1* microdeletion, which is observed in about 5-11% of the NF1 patients (Cnossen et al., 1997, Kluwe et al., 2004, Zhang et al., 2015). The second genotype/phenotype correlation reported was the presence of a 3-base pair deletion of the *NF1* gene (c.2970-2972 delAAT, p.Met992del) (Upadhyaya et al., 2007). This specific one amino acid deletion in exon 17 is associated with a milder phenotype in many patients. The patients with this mutation have multiple café au lait macules but lack of cutaneous or plexiform neurofibromas (Upadhyaya et al., 2007). Another genotype/phenotype correlation reported was a higher incidence of Noonan syndrome features including short stature and pulmonic stenosis in patients carrying *NF1* missense mutations affecting p.Arg1809 (Pinna et al., 2015, Rojnueangnit et al., 2015). These patients presented also with multiple café au lait macules but no externally visible plexiform neurofibromas or clear cutaneous neurofibromas. A recent clinically relevant genotype-phenotype correlation was identified in patients with missense mutations affecting *NF1* codons 844-846 (Koczkowska et al., 2018). Variants in this region seem to confer a more severe *NF1* phenotype and a high predisposition to developing malignancies.

An additional possible correlation involves the NF1 patients characterized by bilateral neurofibromas located at spinal nerve roots with only a few other disease features. Many of these patients have been reported to have an *NF1* germline missense or a splicing mutation (Kluwe et al., 2003, Messiaen and Wimmer, 2008, Ruggieri et al., 2015).

***NF1* microdeletions**

NF1 microdeletions refer to large deletions, which cover the entire *NF1* gene and a number of flanking genes (Cnossen et al., 1997, Kluwe et al., 2004). Microdeletions are the most frequent recurring mutations in *NF1*. The three established types of *NF1* microdeletions are different in terms of their size and breakpoint position. The type 1 *NF1* microdeletion is the most frequent encompassing 1.4 Mb. The type 2 microdeletion spanning 1.2 Mb and type 3

spanning 1.0 Mb are less frequent (Bengesser et al., 2010, De Raedt et al., 2004, Kehrer-Sawatzki et al., 2004, Petek et al., 2003).

Patients with a constitutional *NF1* microdeletion usually show a more severe NF1 phenotype. The patients are characterized by a large number of neurofibromas at a young age, dysmorphic facial features, a developmental delay, and possible intellectual disability (Tonsgard et al., 1997, Upadhyaya et al., 1998). *NF1* microdeletion patients may also have cardiac defects and growth and skeletal abnormalities. Patients with *NF1* microdeletion have a two to three times higher risk for MPNST than patients with other *NF1* mutations (De Raedt et al., 2003). The co-deletion of the *SUZ12* gene in the *NF1* microdeletion region is thought to be a risk factor for the malignant neoplasms (Kehrer-Sawatzki et al., 2017b).

2.4 *NF1* mutation analysis

During the 21st century, molecular diagnostics of NF1 has become possible and increasingly required. The *NF1* mutation analysis is challenging due to it being a large and complex gene, the lack of mutational hotspots, the occurrence of a very diverse spectrum of mutation types, and the presence of *NF1* pseudogenes (Luijten et al., 2001, Messiaen and Wimmer, 2008). However, mutation analysis of the *NF1* gene has proven valuable especially in young children who may only partially fulfill the clinical criteria. The same holds true for adults with atypical clinical presentation. Also prenatal testing and preimplantation genetic diagnosis are implications for *NF1* mutation analysis.

Several different techniques have been applied to look for mutations in this challenging gene. Fahsold and coworkers screened 500 unrelated NF1 patients for mutations in the *NF1* gene either by the protein truncation test, temperature-gradient gel electrophoresis of genomic PCR products, or by direct genomic sequencing of all individual exons (Fahsold et al., 2000). Mutation-detection efficiencies of these various screening methods were very similar: 47.1% for protein truncation test, 53.7% for temperature-gradient gel electrophoresis, and 54.9% for direct genomic sequencing. Ars et al. (Ars et al., 2000) used a whole *NF1* cDNA screening methodology to study 80 unrelated NF1 patients. This approach was able to detect mutations in 87% of the familial cases but only in 51% of the sporadic ones. Denaturing high performance liquid chromatography was used for *NF1* mutational analysis by Han et al. (Han et al., 2001). The sensitivity of this method was evaluated in a retrospective study of a cohort of 111 unrelated NF1 patients with known germline mutations and of these, 97% of the mutations

were detected. In a subsequent prospective analysis of 50 unrelated NF1 patients, germline mutations were identified in 68% of the patients.

At the moment, a comprehensive multi-step approach with an RNA-based center assay complemented with additional methods is used to analyze the entire *NF1* coding region (Messiaen and Wimmer, 2008). This method starts by dividing a blood sample from a patient into aliquots for gDNA extraction, for preparation of cell suspension, for interphase fluorescence *in situ* hybridization, and for short-term lymphocyte culture for RNA extraction. Following RNA extraction and cDNA synthesis, the entire coding region is amplified by PCR in three overlapping fragments. All fragments are analyzed by gel electrophoresis followed by direct cDNA sequencing. All alterations detected at the cDNA level are further characterized at the genomic level using specific primers. Total gene deletion is screened with microsatellite analysis and further analyzed by multiplex ligation depended probe amplification (MLPA) (Wimmer et al., 2006) and FISH analysis (Messiaen et al., 2000). Using this comprehensive method, the mutation detection rate is over 95% in non-founder NF1 patients fulfilling the NIH diagnostic criteria (Messiaen et al., 2000).

At present, also direct sequence analysis of all coding exons and exon/intron boundaries of *NF1* gene together with MLPA analysis is in clinical use (Croonen et al., 2012). With this method, a pathogenic mutation is found in ~60% of the index patients and in 80.9% of the index patients fulfilling the NF1 diagnostic criteria (van Minkelen et al., 2014).

2.4.1 Next-generation sequencing

Frederick Sanger developed the traditional DNA sequencing technique in 1977 (Sanger et al., 1977). The completion of the Human Genome Project in 2001 made a deeper understanding of the genome possible. However, the first sequencing of the human genome with the traditional Sanger sequencing technology required approximately 13 years and a cost of about 3 billion dollars (Lander et al., 2001, Venter et al., 2001). It was not until 2005 that the first next-generation sequencing (NGS) method was introduced (Margulies et al., 2005). The discovery of new sequencing methods made sequencing a lot more efficient; the parallelization of the sequencing reaction has greatly increased the total number of produced sequence reads per run (Buermans and den Dunnen, 2014). The NGS methods enable a simultaneous and rapid sequencing of millions of DNA molecules at a reduced cost. With NGS, a human genome can be sequenced within a week at a cost close to 1,000 dollars (<https://www.genome.gov/sequencingcostsdata/>). A

further advantage of NGS is the versatility of applications. In addition to DNA sequence determination, it is applicable for detection of amplifications, deletions, gene fusions, DNA methylation, and gene expression (Müllauer, 2017). Though, a successful NGS project requires expertise both in the wet lab as well as at the bioinformatics side to secure a high quality of data and data interpretation (Buermans and den Dunnen, 2014).

NGS has been applied to *NF1* mutation analysis in few validation studies. Sequencing the whole genome without selection of a specific region is not necessary for this purpose. Instead, targeted resequencing of the large *NF1* gene and possibly other disease-related genes has been used. Chou et al. had two NF1 samples and used DNA sequence capture and enrichment by a customized high-density microarray of the *NF1* gene region (280kb) followed by sequencing with the Roche/454 GS FLX system (Chou et al., 2010). Maruoka et al. used the *NF1* gene together with 108 causative genes for more common classical congenital malformation syndromes as a target region (Maruoka et al., 2014). The 108 genes included also other known RASopathy genes such as *SPRED1* and *HRAS*. The sequencing was carried out using the MiSeq (Illumina) system and was able to detect 92.1% (70/76) of the mutations when large deletions were excluded.

Cunha et al. used a hybridization capture-based next-generation sequencing performed on Ion Torrent PGM to screen coding and noncoding *NF1* regions (Cunha et al., 2016). In this study, 10/11 (91%) mutations were found. Similarly, a middle-throughput Ion Torrent PGM platform was used with a target of *NF1* gene plus flanking introns by Cali et al. to determine the mutations in patients with clinical diagnosis of NF1 (Cali et al., 2017). A total of 73 mutations were identified in 79 patients with this method and a further mutation detection rate of 80% was estimated.

Pasmant et al. used a targeted next-generation sequencing of *NF1* and *SPRED1* genes using a multiplex PCR approach in mutation analysis (Pasmant et al., 2015). Also, this study utilized the Ion Torrent PGM platform. The study included 30 validation samples and 279 patient samples, and *NF1* or *SPRED1* alteration was found in 246/279 (88%) and 10/279 (4%) of the patients.

The *NF1* gene is also included in some NGS gene panels, for example, in hereditary cancer testing (Castellanos et al., 2017), anaplastic thyroid cancer (Latteyer et al., 2016), and soft tissue sarcomas (Jour et al., 2014).

2.5 Epidemiology of NF1

2.5.1 NF1 prevalence and incidence

There is no population in which NF1 is known not to occur (Friedman, 1999). The NF1 incidence and prevalence have been estimated in many studies carried out with various methods in different continents since the 1950s. The frequency of NF1 varies from study to study, probably due to ascertainment differences and population- versus hospital-based patient groups. Previous population-based studies have been limited to regions of Gothenburg, Sweden (Samuelsson and Axelsson, 1981); southeast Wales (Huson et al., 1989); Dunedin, New Zealand (Fuller et al., 1989); northeast Italy (Clementi et al., 1990); and northern Finland (Poyhonen et al., 2000).

Crowe et al. (Crowe et al., 1956) was the first to estimate the frequency of NF1. The study was based on surveys on regional hospital admissions and state mental institutions in Michigan, USA and resulted in a prevalence estimation of 1/2,500-1/3,300. Sergeyev (Sergeyev, 1975) found the disease to be less prevalent by ascertaining patients via medical examination of Russian 16-year-old military pre-recruits. The study ended up in a prevalence estimate of 1/7,812. Military pre-recruits have also been used in other studies (Fazii et al., 1998, Garty et al., 1994, Ingordo et al., 1995). The highest prevalence estimate reported is as high as 1/960 by Garty et al. (Garty et al., 1994). This was reached by examining 17-year-old military pre-recruits in Israel. Lammert et al. studied NF1 prevalence in Germany by screening children for NF1 during routine medical examinations during elementary school enrollment (Lammert et al., 2005). They reported a prevalence estimate of 1/2,996 among 6-year-old children.

Many studies have utilized searches of medical records, letters to medical institutions, and physicians, together with examination of affected patients and their relatives (Fuller et al., 1989, Huson et al., 1989, Poyhonen et al., 2000, Samuelsson and Axelsson, 1981). Also searches on medical records at departments of genetics or genetic registers have been carried out (Clementi et al., 1990, Evans et al., 2010, McKeever et al., 2008). The NF1 prevalence is reported to be higher in young children than in adults (Clementi et al., 1990, Huson et al., 1989). This is probably due to early deaths of some NF1 patients.

Birth incidence of NF1 has been reported in only a few studies. Huson et al. identified 69 families with a total of 135 NF1 patients in southeast Wales (Huson

et al., 1989). In these families, penetrance of NF1 was 100% by the age of 5 years. The study suggested an incidence as frequent as 1/2,558. A genetic register-based approach was used in patient ascertainment in the Manchester region of northwest England (Evans et al., 2010). The best estimate for birth incidence of NF1 in this study was 1/2,699 with 979 NF1 patients from 539 families.

Lammert et al. assumed that the NF1 incidence at birth was the same as the prevalence at age 6 and resulted in an incidence estimate of 1/2,600-1/3,000 (Lammert et al., 2005). NF1 incidence has been previously assessed in northern Finland by Pöyhönen et al. with a peak incidence estimate of 1/2,703 (Poyhonen et al., 2000). This study was based on 197 NF1 patients in 119 families identified through hospital records followed by clinical assessment during the years 1989-1996.

The traditionally accepted proportion of *de novo* NF1 cases is 50%. In the Pöyhönen study, 96 of the 197 (49%) NF1 cases identified were considered to represent probable new mutations of the disease gene (Poyhonen et al., 2000). Similar results of a positive family history in about half of the cases have been published in several studies (Clementi et al., 1990, Evans et al., 2010, Samuelsson and Axelsson, 1981). However, occurrence estimation of new mutations is also subject to ascertainment bias. The later age at diagnosis in *de novo* NF1 cases means that the overall rate might be an underestimate. Evans et al. reported a median age at diagnosis being 4 years later for *de novo* than familial cases (Evans et al., 2010).

Most of the studies and the ascertainment methods used could have missed mildly affected NF1 patients, particularly children who have a new mutation or have no major disease complications. NF1 is also infrequently diagnosed at birth, because most of the symptoms develop later in life making estimates of birth incidence difficult. Also, Pöyhönen et al. reported that the mean age at diagnosis of NF1 has dropped from 20 years among patients born in the 1960s to less than six years for patients born in the 1980s (Pöyhönen, 2000). This was explained by a growing awareness of the disease and better knowledge about the diagnostic criteria. Accordingly, greater awareness of neurofibromatosis leads to earlier diagnosis.

2.5.2 Mortality in NF1

Information about NF1 mortality is limited, but all published studies have found a high rate of cancer-derived mortality. Previous studies about NF1-derived

mortality can be divided into studies exploiting death certificates searching for NF1 diagnosis and studies using patient cohorts.

Death certificate studies

There have been at least three NF1 mortality studies using death certificates in the search for NF1 patients. Imaizumi (Imaizumi, 1995) studied the death rate from neurofibromatosis by analyzing Japanese vital statistics during the period of 1968-1992. Only cases with neurofibromatosis recorded as a cause of death were used, and there was no separation between NF1 and NF2 patients.

Rasmussen et al. analyzed NF1 mortality using U.S. death certificates from 1983-1997 (Rasmussen et al., 2001). They identified 3,770 presumed NF1 cases among 32,722,122 deaths. This translates to 1 in 8,679 deaths and assuming the death rate and incidence remain constant, NF1 patients seem to be underascertained in this study due to not having NF1 recorded in every death certificate of NF1 patients. NF1 patients were 1.2 times more likely (proportionate mortality ratio, PMR: 1.21; 95% Confidence Interval: CI 1.14-1.28) to have a malignant neoplasm listed on their death certificate and 34 times more likely (PMR: 34.3; 95% CI 30.8-38.0) to have a malignant connective or soft-tissue neoplasm listed on their death certificate. The mean age at death in this study was 15.7 years lower than in the general population.

The third study utilized Italian death certificates from the period of 1995-2006 (Masocco et al., 2011). They found 632 deaths with an NF1 diagnosis on the death certificate, and an approximately 20-year lower mean age of death in NF1 patients compared to the general population. Supporting previous findings, they reported a high amount of deaths caused by malignant neoplasms in the brain (PMR: 4.2; 95% CI: 2.69-6.15) and in connective and other soft tissue (PMR 22.3; 95% CI 15.50-30.95).

These types of studies are dependent on the NF1 diagnosis being accurately mentioned in death certificates, and thus are biased towards recording only the known NF1-related causes of death like malignant neoplasms of connective tissue, soft tissue, and brain. Similarly, these studies are also prone to detect only severe NF1 cases.

Cohort studies

Sørensen et al. (Sørensen et al., 1986) utilized a Danish cohort of 212 NF1 patients, who had been identified from hospital admissions and from the search of affected family members during the years 1924-1944 by Borberg (Borberg, 1951). The cohort was followed-up over a 42-year period and analyzed for mortality, survival,

and cancer (Sørensen et al., 1986). The most common causes of death among the 113 deceased patients were typical of those of the general population: cancer, myocardial infarction, cerebrovascular accidents, and pneumonia accounting for 77% of deaths. Female probands had the lowest survival rates in the study.

A Swedish cohort of 70 adult NF1 patients (Samuelsson and Axelsson, 1981) from Göteborg, Sweden was followed up for 12 years (Zöller et al., 1995). The mean age at death was 61.6 years, which was 15 years earlier than in the general population. A population-based study from northwest England included 130 deaths in a cohort of 1,186 NF1 patients (Evans et al., 2011). The most common cause of death was MPNST with 26% of deaths being due to this condition. MPNST was reported significantly more than expected as a cause of death for both NF1 women (Standardized mortality ratio, SMR: 7,788.2; 95% CI: 4,355.7-12,846.2) and men (SMR: 3,819.6; 95% CI: 1,971.4-6,672.5). The median survival of NF1 patients was 71.5 years, which was 8 years less than the survival in the general population (Evans et al., 2011). A French study with a cohort of 1,226 NF1 patients and a median follow-up time of 6.8 years documented 67 deaths. An excess mortality was seen in the patient groups under 40 years with MPNST being the most common cause of death (60%) (Duong et al., 2011).

2.5.3 Syndromes with phenotypic overlap with NF1

Several syndromes have an overlapping phenotype with NF1 and differential diagnosis is sometimes needed.

Neurofibromatosis type 2

The clinical and genetic distinction between Neurofibromatosis type 2 (NF2) and NF1 was not fully recognized until the differences in clinical picture and genetics resulted in the definition of two distinct conditions. Before the National Institutes of Health Consensus Conference in 1987 (Stumpf et al., 1988), many studies contained both NF1 and NF2 patients. NF2, previously called bilateral acoustic/central neurofibromatosis, is also a dominantly inherited disease but caused by mutations in the *NF2* gene on chromosome 22 (Rouleau et al., 1993, Seizinger et al., 1987). The hallmarks of NF2 include bilateral vestibular schwannomas, other schwannomas, and intracranial meningiomas (Evans et al., 1992a, Mautner et al., 1996, Parry et al., 1994). NF2 is considerably less common than NF1. The reported incidences of NF2 have varied between 1/33,000 (Evans et al., 1992b) and 1/87,410 (Antinheimo et al., 2000). The mortality in patients with NF2 has been investigated in a few studies (Baser et al., 2002, Evans et al., 1992a, Parry et al., 1994). Evans et al reported a mean age of death at 36 years; 39

of the 40 deaths resulted from a complication of neurofibromatosis (Evans et al., 1992a). The age at diagnosis was the strongest single predictor of the risk of mortality in NF2 patients (Baser et al., 2002).

Schwannomatosis

Schwannomatosis is a genetic disease characterized by multiple schwannomas (MacCollin et al., 2005). Schwannomas are benign nerve sheath tumors that most commonly occur singularly in the general population. Multiple schwannomas in a single patient are most often seen in NF2 patients, but the lack of vestibular schwannomas distinguishes schwannomatosis from NF2 (MacCollin et al., 1996). Genetic studies indicate that germline mutations of either *SMARCB1* or *LZTR1* tumor suppressor genes are identified in 86% of familial and 40% of sporadic schwannomatosis patients (Kehrer-Sawatzki et al., 2017a, Plotkin et al., 2013).

CMMRD

Patients with constitutional mismatch repair deficiency syndrome (CMMRD) can have a cutaneous phenotype remarkably similar to NF1, and the affected individuals may meet the NIH diagnostic criteria for NF1. However, these patients have a different cancer spectrum (Tabori et al., 2017). CMMRD is a distinct childhood cancer predisposition syndrome with a very broad tumor spectrum. The tumors related to this syndrome can be divided into three main categories: hematological malignancies (i.e., non-Hodgkin's lymphomas); central nervous system tumors (i.e., high grade gliomas and glioblastomas); and carcinomas of the colon, rectum, and small intestine (Wimmer et al., 2014). CMMRD results from biallelic germline mutations in one of the four MMR genes (Wimmer et al., 2014). CMMRD is an important differential diagnosis in children who have café au lait macules and a malignancy that is not typically associated with NF1.

Legius syndrome

NF1 belongs to the group of RASopathies, which are genetic syndromes with disease-causing mutations in genes coding for proteins that play an important role in the RAS-MAPK pathway (Zenker, 2011). These syndromes share many clinical features but are caused by disease-causing mutations in different genes of the pathway. Legius syndrome, identified in 2007 (Brems et al., 2007), is one RASopathy. It is caused by heterozygous germline mutations in the *SPRED1* gene in chromosome 15 (Brems et al., 2007). The clinical features of Legius syndrome resemble and partially overlap with those of NF1. Multiple café-au-lait macules are the hallmark of both syndromes, but patients with Legius syndrome do not have tumors typical for NF1 such as neurofibromas or optic pathway gliomas (Brems et al., 2007, Brems and Legius, 2013, Denayer et al., 2010). Patients with Legius

syndrome may have skinfold freckles, but they do not have Lisch nodules (iris hamartomas), which belong to the NF1 diagnostic criteria (Brems et al., 2007). Distinguishing Legius syndrome from NF1 is sometimes impossible on the basis of clinical features alone especially in young children. This is because cutaneous neurofibromas and Lisch nodules, which characterize most patients with NF1, do not usually arise until later in childhood or adolescence. Because of the clinical overlap with NF1, genetic testing is necessary to establish the diagnosis. At least 153 different mutations of the *SPRED1* gene have been reported (<http://www.lovd.nl/SPRED1>). These include different mutation types such as point mutations and copy number alterations (Brems et al., 2007, Messiaen et al., 2009, Spencer et al., 2011).

2.6 Cancer

2.6.1 Cancer types

Neoplasms are classified as being benign or malignant with malignant being equivalent to cancer. Cancers are traditionally described according to their cell of origin or the tissue from which they arise. Most common cancers, accounting for about 80% of cases, arise in epithelial cells and are classified as carcinomas (Alberts et al., 2015). This could be explained by the fact that most cell proliferation and exposure to damaging agents in adults occur in epithelia. The type of the epithelium can also be added to the name, for example, in the case of breast adenocarcinoma, which arises from glandular epithelium of the breast tissue (Alberts et al., 2015). Correspondingly, mesenchymal cells, e.g., connective tissue and muscle cells give rise to sarcomas. Cancers originating from different cell types are in general very different diseases. However, two tumors with identical pathological classification and morphological features can have different mutation profiles, but then again, cancers arising from different organs can have a highly similar set of mutational profiles and might respond similarly to target therapies (Hoadley et al., 2014). Knowledge of the molecular defects that lead to cancer enables classification of cancers also according to the mutations that make the tumor cells cancerous (Wood et al., 2007).

2.6.2 Biology of cancer

Cancer originates from a single cell that starts to behave abnormally due to mutations in its genome. A single mutation is not enough to cause cancer. The

development of cancer requires that a substantial number of independent, rare genetic and epigenetic changes occur and accumulate in one cell (Alberts et al., 2015). In common solid tumors, such as those derived from the colon, breast, brain, or pancreas, an average of 33-66 genes have undergone a somatic mutation affecting the sequence of their protein products (Vogelstein et al., 2013). Driver mutations are fundamental factors in the development of cancer and critical for driving malignant transformation and supporting the hallmarks of neoplastic growth (Stratton et al., 2009). Passenger mutations are irrelevant to the development of the disease and confer no selective growth advantage for the cell but happen mostly because of the genetic instability of the cancer cell (Alberts et al., 2015, Stratton et al., 2009). A typical tumor contains two to eight driver gene mutations, while the remaining mutations are passenger mutations (Vogelstein et al., 2013).

The driver genes can be classified into 12 signaling pathways that regulate three core cellular processes: cell fate, cell survival, and genome maintenance (Vogelstein et al., 2013). The hallmarks of cancer have been proposed by Hanahan and Weinberg (Hanahan and Weinberg, 2000, 2011). Accordingly, cancer cells have acquired few major properties. Cancer cells have a mutation or epigenetic change that increases the rate at which cells proliferate or enables the proliferation when normal cells would stop (Hanahan and Weinberg, 2000). In other words, cancer cells reproduce in defiance of the normal restraints on cell growth and division. Dysregulation of pathways responsible for cell division and proliferation ultimately makes cells vulnerable to additional genetic alterations that contribute to cancer formation. Cancer cells also fail to go into apoptosis when normal cells would (Hanahan and Weinberg, 2000). Important features of the cancer cell are also the ability to induce angiogenesis, invade surrounding tissue, and form distant metastasis (Hanahan and Weinberg, 2000). Two hallmarks and two enabling traits of cancer were included later to the list of major properties: reprogramming energy metabolism, evading immune response in addition to genome instability, and tumor-promoting inflammation (Hanahan and Weinberg, 2011).

2.6.3 *NF1* gene alterations in cancer

Acquired somatic mutations of the *NF1* gene are found in a wide variety of malignant neoplasms in patients without the NF1 syndrome. Large-scale next-generation sequencing data on cancer genomes and exomes is publicly available. The Cancer Genome Atlas (TCGA) (<http://cancergenome.nih.gov/>) is a public funded project that aims to catalogue and discover major cancer-causing genomic alterations to create a comprehensive atlas of cancer genomic profiles (Tomczak

et al., 2015). TCGA has studied more than 10,000 specimens of 33 different tumor entities on the DNA, RNA, and epigenome level (Tomczak et al., 2015). In addition to TCGA, the International Cancer Genome Consortium (ICGC; <https://icgc.org/>) has been massively characterizing tumor genomes. The results of TCGA and ICGC are freely available to the public.

The cBioportal is an open-access resource for exploration of cancer genomic data sets providing access to data from over 160 cancer studies and currently over 30,000 samples (Cerami et al., 2012, Gao et al., 2013). According to the cBioportal data set, the highest frequency of somatic *NF1* mutations were found in desmoplastic melanomas (45%, 9/20 cases). The high frequency of *NF1* mutations in desmoplastic melanomas might indicate that neurofibromin plays an important role in the cancer biology of this rare type of melanoma (Wiesner et al., 2015). Moreover, a somatic mutation was found in 12-15% of all melanoma cases, thus *NF1* is the third most commonly mutated gene in melanoma (Cancer Genome Atlas Network 2015). A frequent *NF1* mutation was also seen in MPNSTs (40%, 6/15 cases). Sporadic *NF1* mutations and inactivation of neurofibromin are shown to play a critical role in the development of MPNSTs (Lee et al., 2014). Also adenocarcinoma and squamous cell carcinoma of the lung, glioblastoma, and some myeloid malignancies are among the cancers with most frequent *NF1* mutations (Philpott et al., 2017). In addition, the *NF1* gene is reported to be mutated frequently in sporadic breast cancer, and the *NF1* gene has been implicated as a breast cancer driver gene (Wallace et al., 2012). Based on TCGA data set, the *NF1* gene deletions and mutations were most frequent in basal- and *HER2*-enriched breast cancer subtypes (Wallace et al., 2012).

2.7 Cancer in NF1

2.7.1 *NF1* cancer risk

NF1 predisposes individuals to various benign and malignant tumors throughout life. Several studies have previously attempted to estimate the malignancy risk of NF1 patients. Zöller et al. reported a long-term follow-up of 212 Swedish patients with NF1 with a relative risk of malignancy of 4.0 (95% CI: 2.1–7.6) (Zöller et al., 1997). Walker et al. conducted a study with 448 individuals with NF1 with a total of 5,705 years of patient follow-up (Walker et al., 2006). They found a 2.7 (95% CI: 1.9–3.7) times higher overall risk of cancer compared to the general population. Cumulative risk of a malignancy by the age of 50 years was 20% (95%

CI: 14–29%). The most frequent types of cancer were those of connective tissue with a standardized incidence ratio (SIR) of 122 (95% CI: 61.0–219) and brain tumors with a SIR of 22.6 (95% CI: 9.06–46.5) (Walker et al., 2006). Seminog and Goldacre (Seminog and Goldacre, 2013) studied the cancer risk of both neurofibromatosis types 1 and 2 combined with a population-based record-linkage study in England and found a rate ratio (RR) for cancer of 4.3 (95% CI: 4.0–4.6). If only people with presumed NF1 were considered, the RR was 4.0 (95% CI 3.7–4.3) (Seminog and Goldacre, 2013).

Females with NF1 have been reported to have a higher risk of cancer in some studies (Airewele et al., 2001, Sørensen et al., 1986) but not in others (Schneider et al., 1986, Walker et al., 2006). Also the effect of the inheritance pattern on cancer risk has been studied. Inheriting the the *NF1* gene mutation from NF1 parent increased cancer risk in some studies (Schneider et al., 1986, Zvulunov et al., 1998) but not in others (Airewele et al., 2001, Huson et al., 1989).

2.7.2 MPNST

Tumors of the peripheral nerve belong to the hallmark complications of NF1. A malignant peripheral nerve sheath tumor (MPNST), previously referred to as neurofibrosarcoma, is an aggressive soft-tissue sarcoma with half of all cases developing in individuals with NF1 (Evans et al., 2002). It causes substantial morbidity and mortality among NF1 patients. Most MPNSTs of NF1 patients arise from a malignant transformation of plexiform neurofibroma (Friedrich et al., 2007). However, prediction of transformation is clinically and histologically challenging. MPNSTs can also develop from deep tissues with no previous plexiform tumor (Evans et al., 2002). Consequently, these cancers are difficult to detect.

The incidence of MPNSTs in NF1 patients is reported to be 2-19% (Ducatman et al., 1986, McGaughan et al., 1999). The cumulative risk of developing MPNSTs in NF1 patients has been reported to be 8-13% (Evans et al., 2002). The mean age at diagnosis of NF1-associated MPNSTs is reported to be significantly younger than in sporadic cases. The median age at diagnosis was 26 years for NF1 patients, while it was 62 years in patients with sporadic MPNST (Evans et al., 2002). The survival was stated to be statistically better in NF1 females than NF1 males (5 and 10-year survival 46% and 41.5% versus 22% and 8.2%; $p=0.05$) (Evans et al., 2002, Ingham et al., 2011). The survival in women was possibly better due to an earlier stage at the diagnosis (Ingham et al., 2011).

The leading cause of mortality in patients with NF1 is MPNST. Evans et al. observed 26% of deaths being due to this condition (Evans et al., 2011). Duong et al. detected high mortality in the patient groups under 40 years, and MPNST was the most common cause of death (Duong et al., 2011). The five-year survival from diagnosis was reported to be 21% for NF1 MPNST patients, while it was 42% for sporadic MPNST cases (Evans et al., 2002). MPNSTs often metastasize in an early phase and are associated with poor prognosis. This is because MPNST is an aggressive tumor with a high number of local recurrences and has an aggressive nature. Total resection with wide margins (≥ 10 cm) is the preferred primary treatment for larger tumors, but the localization, size, and metastases of the tumors often reduce the feasibility of surgery (Ferner and Gutmann, 2002, Friedrich et al., 2007). If the tumor is non-localized at the time of the diagnosis, the disease is extremely difficult to manage (Bates et al., 2014).

MPNST cells harbor complex genomic alterations. *NF1* loss is also seen in majority of sporadic MPNSTs, but it is not sufficient for MPNST transformation alone (Perry et al., 2001). Other mutations in driver genes, such as *TP53* and *CDKN2A*, are also required (Kim et al., 2017).

2.7.3 Brain tumors

Central nervous system tumors represent a significant portion of all the malignancies affecting NF1 patients. Rasmussen et al. has reported a five times higher (PMR 5.52; 95% CI 4.74–6.38) brain tumor risk (Rasmussen et al., 2001). The most common brain tumor in NF1 patients is optic pathway glioma (OPG). It is a World Health Organization Grade I pilocytic astrocytoma of the optic pathway seen in 15 to 20% of the NF1 patients (Listernick et al., 1994, Listernick et al., 1989). OPGs usually affect the patients in early childhood with a mean diagnosis age of 4.2 years (Listernick et al., 1994). OPGs account for significant morbidity especially in young children with NF1, because one-third to half of these tumors will cause symptoms such as abnormalities in vision (e.g., vision loss, proptosis) or precocious puberty (Balcer et al., 2001, de Blank et al., 2017, Listernick et al., 1994). Many OPGs never cause symptoms and not all patients with symptomatic OPG will require treatment (de Blank et al., 2017, King et al., 2003, Thiagalingam et al., 2004). Overall survival is usually good, but no clear prognostic factors have been identified (Trevissan et al., 2017). Chemotherapy is considered as standard first-line treatment, and radiation is used cautiously because of the risk of second malignancy as well as more frequent neurocognitive and endocrine toxicities in the young patients (Sharif et al., 2006, Sievert and Fisher, 2009). Molecular analysis of NF1-associated pilocytic astrocytomas has shown loss of

heterozygosity of the *NF1* region and loss of neurofibromin expression (Gutmann et al., 2000, Lau et al., 2000). Also mutations in *TP53* and *CDKN2A* genes have been reported (Gutmann et al., 2003).

Brainstem gliomas are less frequent and usually represent pilocytic astrocytomas. These are often incidental findings of neuroimaging studies (Campian and Gutmann, 2017). Patients with NF1 may also develop high grade brain tumors particularly with an onset later in life (Rosenfeld et al., 2010). Gutmann et al. reported a prevalence of brain tumors in patients with NF1 < 50 years of age to be more than 100 times greater than in the general population (Gutmann et al., 2002).

2.7.4 Other cancers with possible risk

Several other cancer types have been reported to have an increased incidence in NF1 patients. Gastrointestinal stromal tumors (GIST) most commonly occur sporadically, but an increased incidence for this tumor type has been suggested in NF1 patients (Maertens et al., 2006, Zöller et al., 1997). GISTs are mesenchymal tumors that arise from the interstitial cell of Cajal and are the most common nonepithelial tumors of the gastrointestinal tract. The NF1-associated GISTs lack the mutations in *KIT* and *PDGFRA* seen in sporadic GISTs. In contrast, the GISTs in NF1 patients carry a second hit mutation of the *NF1* gene (Maertens et al., 2006). NF1-associated GISTs also occur as multiple tumors, which is not typical in sporadic cases (Brems et al., 2009).

Pheochromocytoma is a catecholamine-producing tumor of the adrenal medulla. Hypertension may be caused by this tumor type. Pheochromocytomas have been reported to be clinically identified in 0.1-5.7% of the NF1 patients (Walther et al., 1999). Molecular genetic analyses have demonstrated a role of *NF1* inactivation also in the pathogenesis of pheochromocytomas (Bausch et al., 2007).

Juvenile myelomonocytic leukemia (JMML) has been reported to be overrepresented in NF1 children (Niemeyer et al., 1997, Stiller et al., 1994). JMML is a rare mixed myeloplastic/myeloproliferative disorder that represents approximately 2% of all pediatric hematopoietic malignancies in the general population. Somatic loss of *NF1* in JMML cells has been demonstrated in the majority of affected NF1 patients, which supports a role of *NF1* inactivation in the development of this cancer type in NF1 patients (Side et al., 1997).

Rhabdomyosarcomas, which are non-neurogenic sarcomas, have also been reported in association with NF1 (Matsui et al., 1993).

2.8 Breast cancer and NF1

2.8.1 Breast cancer

Breast cancer is the most common cancer in women, and breast cancer incidence in the general population is highest among 50-69 years old women. Breast cancer survival is generally good, with 97% of the women diagnosed with breast cancer being alive one year after a breast cancer diagnosis and 88% are alive 5 years after the diagnosis, respectively (NORDCAN, www.ancre.nu). Breast cancer is a heterogeneous group of tumors with variable prognosis mainly based on the age of the patient, tumor grade, tumor size, regional lymph nodes status, proliferation index, estrogen (ER) and progesterone (PR) receptor status, and the positivity of the HER2 (human epidermal growth factor receptor 2) (Finek et al., 2007). These are also the main factors contributing to the treatment choices of individual breast cancer cases (Coates et al., 2015).

2.8.2 Breast cancer risk factors

Several risk factors for breast cancer have been documented. The most important ones and also factors with the highest relative risk are related to genetic predisposition, such as family history of breast cancer as well as specific germline mutations associated with breast cancer (Sun et al., 2017, Singletary, 2003). Age is one of the best documented risk factors for breast cancer, and over 80% of all breast cancer cases occur in women aged 50 or older. Factors related to endogenous and exogenous estrogen exposure such as early menarche, late first full-term pregnancy, nulliparity, late menopause, and hormone replacement therapy have been linked to an increased incidence of breast cancer (Clemons and Goss, 2001, Hsieh et al., 1994, Kobayashi et al., 2012, Singletary, 2003). There is a positive correlation between high body mass index (BMI) and the risk of breast cancer in postmenopausal women (Tretli, 1989). Also modern lifestyle with excessive alcohol consumption and too much dietary fat intake can increase the risk of breast cancer (Sun et al., 2017). Increased breast tissue density has also been associated with increased risk (McCormack and dos Santos Silva, 2006, Pettersson et al., 2014).

Familial breast cancer

Family history of breast cancer is a strong risk factor for breast cancer (Singletary, 2003). A cohort study of over 113,000 women in the UK demonstrated that women with one first-degree relative with breast cancer have a 1.75-fold higher risk of developing this disease than women without any affected relatives (Brewer et al., 2017). Twin studies have also implicated a familial risk and a high rate of the heritability of cancer (Möller et al., 2016, Mucci et al., 2016).

BRCA1 and *BRCA2* (breast cancer 1 and 2) genes are the two most important and best-known genes in the hereditary predisposition to breast cancer (Miki et al., 1994, Wooster et al., 1995). Women with germline heterozygous mutations in *BRCA1* or *BRCA2* genes have a hereditary susceptibility to developing breast, ovarian, and other cancers (Brody and Biesecker, 1998). Germline mutations in *BRCA1* or *BRCA2* cause an average cumulative risk of 65% in the case of *BRCA1* or 39% in the case of *BRCA2* for breast cancer (Antoniou et al., 2003). Studies have shown that mutations in these genes occur in about 1% of all breast cancer cases (2000, Kurian, 2010). *BRCA1* and *BRCA2* together with the *TP53* gene are considered as high-risk genes, which cause a lifetime risk of 40-85% for breast cancer. Correspondingly, a mutation in a moderate risk gene such as *PALB2*, *ATM*, or *CHEK2* causes a lifetime risk of 20-40% for breast cancer (Laloo and Evans, 2012).

However, familial breast cancers represent only 5-7% of all breast cancer cases (Melchor and Benítez, 2013), and familial clustering of breast cancer remains largely unexplained. Genome-wide association studies (GWAS) are aimed at the identification of new breast cancer susceptibility variants. Recent studies found more than 50 new breast cancer risk loci associated with the risk for breast cancer (Milne et al., 2017, Michailidou et al., 2017). These and other previous similar studies in GWAS Catalog (<http://www.ebi.ac.uk/gwas/>) have failed to identify the *NFI* gene as a risk locus. This might be due to the huge number of different mutations of the large *NFI* gene, and only small number of unique pathogenic variants present in $\geq 0.5\%$ of all unrelated individuals (Koczkowska et al., 2018).

2.8.3 Molecular profiling of breast cancer

Breast cancers are classified into subgroups according to histopathological and molecular characteristics. Breast cancers form a heterogeneous group of tumors with a variable prognosis that is mainly based on the age of the patient, tumor grade, tumor size, regional lymph nodes status, proliferation index, ER and PR

status, and positivity of the HER2 amplification (Coates et al., 2015, Finek et al., 2007). The intrinsic breast cancer subtypes, established based on their gene expression profiles, are an additional way to classify breast cancers (Perou et al., 2000). At least four molecular classes shown below are distinguished: luminal A, luminal B, HER2-enriched, and basal-like cancers (Coates et al., 2015). Luminal A and B are both ER/PR-positive types, but luminal B is often also positive for HER2. Luminal B is also often a high-grade tumor with high Ki-67 expression. The HER2-subtype is typically negative for ER and PR but shows over-expression for HER2. Basal-like breast cancer is an aggressive triple negative breast cancer type (ER-, PR-, and HER2-) that expresses cytokeratins of the basal epithelial layer such as CK5/6 and CK14 (Coates et al., 2015). The breast cancers can be classified into subgroups following the recommendations below (Coates et al, 2015):

- **Luminal A= ER/PR+, HER2-, Ki-67 < 14%**
- **Luminal B= 1) ER+, HER2+ or 2) ER+, HER2-, and either Ki-67 >14% or PR+**
- **HER2 subtype= ER-, PR-, HER2+**
- **Triple-negative= ER-, PR-, HER2-**
- **Basal-like= ER-, PR-, HER2-, CK5+**

2.8.4 Breast cancer incidence in NF1

Increased incidence of breast cancer in NF1 women has been reported in few publications and numerous case reports. The first to suggest an increased risk of breast cancer in women with NF1 were Sharif et al. (Sharif et al., 2007). They retrospectively evaluated the risk of developing breast cancer in a cohort of 304 NF1 women. Sharif et al. noticed a five-fold risk of breast cancer in women <50 years in addition to an overall SIR of 3.5 (95% CI: 1.9-5.9) for breast cancer. Since then, there have been other supportive studies about increased breast cancer risk-related to NF1 (Madanikia et al., 2012, Seminog and Goldacre, 2015, Wang et al., 2012). Seminog et al. reported that the risk was most elevated (RR: 6.53; 95%: CI 2.6-13.5) in young women aged 30-39 (Seminog and Goldacre, 2015). Also some case studies about NF1 men with a breast cancer diagnosis have been published (Lakshmaiah et al., 2014, Mann et al., 2017).

Wang et al. studied the germline genomic profile of 14 NF1 women who developed breast cancer (Wang et al. 2017). They found *NF1* mutations but no deleterious mutations in known high or moderate penetrance cancer genes. Some rare or common variants in cancer-related genes were identified, but because of the small sample size, definitive conclusions about the association between the variants and NF1 breast cancer could not be reached in the study.

Screening for breast cancer has been suggested for NF1 women from the age of 40 years (Sharif et al., 2007). However, there are also concerns about the extra radiation dose related to mammography, because radiotherapy has been related to secondary malignancies in NF1 (Sharif et al., 2006). There have been suggestions about late breast cancer diagnosis because of NF1 (Evans, 2012). Challenges may be associated with detecting an advanced breast cancer because of numerous skin neurofibromas, which will possibly cause a substantial delay in cancer diagnosis (Da Silva et al., 2015).

Ras in Breast Cancer

Elevated Ras signaling has been reported to be associated with basal-like and HER2 subtype tumors in humans and have a significant negative association with ER signaling in breast cancer (Wright et al., 2015). However, high level Ras signaling was associated with relapse and death of ER-positive breast cancer patients (Wright et al., 2015), and the activation of the Ras pathway predicts a poor survival outcome of patients treated with tamoxifen (McGlynn et al., 2009).

3 AIMS OF THE STUDY

The purpose of the present study was to develop a new method for NF1 molecular diagnostics. This thesis study evaluated the epidemiology and cancer incidence of NF1 in Finland with a national NF1 cohort and retrospective register-based total population study. The aim of this study was also to advance the knowledge of characteristics of the breast cancers diagnosed in NF1 patients.

The specific aims of the study were:

1. To develop and validate a new method for *NF1* molecular diagnostics exploiting NGS.
2. To identify all NF1 patients in Finland and to study the NF1 incidence as well as to evaluate the overall- and cause-specific mortality of Finnish NF1 patients using SMR estimates.
3. To determine the cancer incidence among NF1 patients compared to the general population.
4. To study the breast cancer of NF1 patients and determine the characteristics of the NF1 breast cancers.

4 MATERIALS AND METHODS

4.1 DNA sampling (I)

Saliva samples from 16 unrelated NF1 patients were collected. All of these patients were of Finnish ancestry and fulfilled the NIH diagnostic criteria for NF1. Of these patients, 4 had an inherited mutation, while 12 patients did not have a known family history of NF1. The *NF1* mutations of 6 patients were previously known, and in addition, one patient had a previously detected *NF1* type 2 microdeletion.

Genomic DNA was isolated from saliva using a saliva sampling kit (Oragene, DNA Genotek, OG-575). Contaminating RNA was degraded using RNase treatment at 37°C for 30 min (RNase Cocktail Enzyme mix, Ambion, Applied Biosystems, AM2286). The DNA was further purified using a gDNA isolation kit (NucleoSpin® Tissue Macherey-Nagel GmbH & Co., 740952). The quality of the DNA was evaluated on agarose gels, the NanoDrop spectrophotometer (Thermo Scientific, ND-1000), and 2100 BioAnalyzer (Agilent, G2939AA). The gDNA was quantified using the Quant-iT Picogreen reagent (Life Technologies, P11496) and Plate Chameleon V fluorometer (Hidex, 425-106).

4.2 Sequence capture and sequencing (I)

Sample library preparation

A total of 16 indexed Rapid libraries were prepared according to the manufacturer's instructions starting from 500 ng of gDNA per sample (Rapid Library Preparation Method Manual, GS Junior Titanium Series, May 2010 and March 2012, Roche). The gDNA was fragmented by nebulization, and the fragments were end-repaired leaving a single A overhang in the 5' end of each strand. Indexed adaptors were ligated next to the fragment ends. The library preparation was completed by removing small DNA fragments with the Agencourt Ampure XP reagent (Ordior Inc, Helsinki, Finland, A63880). The final preparation contained less than 10% of fragments < 350 base pairs.

Sequence capture and sequencing

The sample libraries were amplified and pooled into two sets (A and B) prior to the sequence capture. The sample set A contained 10 samples including 3 controls with known mutations. Sample set B contained 6 samples including 4 controls. The total amount of DNA in both sample sets was 1 µg.

The *NFI* exome was enriched using a sequence capture with a custom probe design and a modified NimbleGen SeqCap EZ Choice Library protocol optimized for the enrichment of target areas < 100 kb (Roche NimbleGen SeqCap EZ Rapid Library Small Target Capture LR, December 2009, Roche Nimblegen Inc., Madison, WI, USA). The targeted regions included 58 *NFI* exons and an additional 50 bp of flanking upstream and downstream intronic sequences resulting in a total target size of 16 kb. The target region did not include alternatively spliced exons 10a2, 9b, or 48a. The average lengths of captured fragments were 757 base pairs and 727 base pairs in the sets of A and B, respectively.

The emulsion PCR and sequencing were performed separately for both sets according to the manufacturer's instructions (emPCR Amplification Method Manual - Lib-L, and Sequencing Method Manual, GS Junior Titanium Series, May 2010 and March 2012, Roche) using the 2720 Thermal cycler (Life Technologies, Eugene, Oregon, USA, 435965) and the 454 GS Junior instrument (454 Life Sciences, Branford, CT, USA). Following the sequencing run, the reads were quality-filtered using the default settings of the GS Run Processor. Only reads that passed all the quality filters were used in the analysis. The number of reads passing the quality-filtering exceeded 100,000, which is the cut-off limit for a successful run (Sequencing Method Manual, GS Junior Titanium Series, May 2010 and March 2012, Roche). Furthermore, over 60% of the control reads had less than 5% errors in the first 400 base pairs, which is consistent with the criteria for a successful sequencing run.

Data analysis and identification of mutations

The bioinformatics tools used in data handling are listed in Table 2. The bioinformatics and data analysis are described in detail in the original Publication I.

Table 2 Bioinformatics tools

PROGRAM	APPLICATION	REFERENCE
SFF-EXTRACT SCRIPT (VERSION 0.3.0)	sff file format to fastq format	http://bioinf.comav.upv.es/sff_extract/index.html
BOWTIE 2 ALIGNMENT PROGRAM (VERSION 2.0.0-BETA7)	Mapping the reads	Langmead and Salzberg, 2012
SAMTOOLS (VERSION 0.1.18)	File handling	Li et al., 2009
PICARD (VERSION 1.78)	File handling	http://picard.sourceforge.net
BIOPYTHON (VERSION 1.60)	File handling	Cock et al., 2009
BEDTOOLS (VERSION 2.16.2)	Coverage calculation	Quinlan and Hall, 2010
R (VERSION 2.15.1)	Coverage calculation	Gentleman et al., 2004, Hadley, 2007
INTEGRATIVE GENOMICS VIEWER (IGV, VERSION 2.0.34 1623)	Alignment visualization	Robinson et al., 2011

Putative mutations in the region of the *NF1* gene were identified from the sequence alignments using a GATK UnifiedGenotyper [version 2.1–11 (McKenna et al., 2010)] with settings that allow for both single nucleotide polymorphisms (SNPs) and insertions/deletions to be called. The main parameters influencing the power of detecting heterozygous variants are the coverage and selection of the used sample variant frequency. The variants were further filtered by the following criteria: 1) minimum coverage of 20x or higher, 2) a sample variant frequency between 30-70%, 3) a variant in a targeted region, and 4) a variant present in only one sample within a sample set with a frequency of 30% or higher. The fourth criterion was based on the expectation that 16 unrelated NF1 patients were not

likely to have shared mutations. This criterion also excluded possible SNPs present in the Finnish population as well as a majority of false positive insertions/deletions in homopolymer sequences, which are a common type of error in 454 sequencing data (Loman et al., 2012, Margulies et al., 2005). The variants that passed all criteria were compared to the dbSNP database build 135 (Sherry et al., 2001), and the Finnish “The Sequencing Initiative Suomi” database (Lim et al., 2014).

Verification of discovered variants was carried out using Sanger sequencing. Fragments with putative mutations or potential mutations in homopolymer regions were amplified by PCR and sequenced with Applied Biosystems 3130xl Genetic Analyzer at the DNA Sequencing Service of Turku Centre for Biotechnology, Turku, Finland.

4.3 Data collection and patient ascertainment (II-IV)

Search for NF1 patients

The patients treated for neurofibromatosis during the years 1987-2011 were searched from the hospital registers of five University Hospitals and 15 Central Hospitals of Finland. The patients from the autonomous Åland Islands with around 28,000 residents are included in the study if they visited any of the University hospitals or 15 Central hospitals of mainland Finland. In addition to hospital records, the nation-wide Hospital Discharge Register (HILMO) containing information about all inpatient hospital admissions from 1967 onwards and two private organizations providing genetic counseling, The Family Federation of Finland (Väestöliitto) and Folkhälsan, were also included in the search. The search was carried out using the International Classification of Diseases (ICD) diagnosis codes. The ICD-9 classification has been in use in Finland during 1987-1995 and the ICD-10 codes since year 1996. The ICD-9 diagnosis code 2377A (neurofibromatosis), and ICD-10 diagnosis codes Q85.00 (neurofibromatosis type 1), Q85.0 (neurofibromatosis), Q85.09 (other neurofibromatosis), Q85 (phakomatoses, not elsewhere classified), and Q85.01 (neurofibromatosis type 2) were used in the search. After the initial search, a careful review of the medical records of the individuals with a focus on the exclusion of patients with other conditions and unclear cases was done. Only patients fulfilling the NIH diagnostic criteria were included in the NF1 study cohort.

Registry data collection

Since 1967, every Finn has received a unique personal identity code. The code includes the date of birth and gender and remains the same over their lifetime. This code is also being used in all national main registers in data storage and identification and thus can be used in collecting and cross-linking data from different registers.

The date of birth, date of death, and possible date of emigration of the NF patients were collected from the national Population Register Centre. The Causes of Death Register of the Statistics Finland provided the causes of death and death certificates, and all cancer data was obtained from the Finnish Cancer Registry.

4.4 Statistical analyses (II-IV)

The NF1 birth incidence was calculated as the number of new cases related to the total number of live births in Finland in the corresponding year. The numbers of live births in the total population were provided by the Statistics Finland (Official Statistics of Finland (OSF): Births http://tilastokeskus.fi/til/synt/index_en.html, accessed: 16.6.2014).

The NF1 patients were followed-up starting from the date of the first hospital visit due to NF1 between 1987–2011 and ending to death, emigration, or ending on 31 December 2012, whichever occurred first (II, III). The cut-off date for the follow-up in Study IV was 31 December 2013. Standardized incidence ratios (SIR) and mortality ratios (SMR) were calculated as the ratios of observed and expected cases, where the expected cases were obtained by multiplying the person-years with the corresponding population rate stratified by age, calendar-period, and gender. The 95% confidence intervals (CI) were based on the assumption that the number of observed cases followed a Poisson distribution.

The cumulative risk for cancer (i.e., probability of developing cancer by a certain age) was estimated applying competing risk methods that allow delayed entry (Putter et al., 2007). When estimating the cumulative risk of cancer, death was defined as the competing event that would prevent the subject from being diagnosed with cancer, and the cumulative incidence was estimated by cause-specific hazard method as described by Putter et al. (Putter et al., 2007).

The cancer-specific survival of NF1 patients was compared to that of matched controls from the Finnish Cancer Registry (Soininen et al., 2012). The controls were matched for cancer site, gender, diagnosis age (within 6 years or 4 years in the design where NF1 cancers were omitted), and diagnosis year (within 10 years

or 6 years, respectively). All available controls without NF1 were included. Controls were weighted such that their distributions of the cancer site, gender, age and time of the diagnosis were the same as those in the NF1 patients. Cumulative cancer-specific survival proportions with 95% confidence intervals were calculated using the weighted Kaplan-Meier method. The matched Cox regression model was used to test the statistical significance of the differences between groups (Cox, 1972).

Confidence intervals of SMRs in Study II were calculated with the Mid-P exact test using OpenEpi 2.3.1. statistical software (Open Source Epidemiologic Statistics for Public Health, www.OpenEpi.com, accessed 2014/06/17). Statistical analyses were conducted with statistical software R version 3.2.2 (<https://www.r-project.org/>) with popEpi package.

4.5 Breast cancer samples (IV)

Archival surgery specimens of all the available female NF1 breast cancers and clinical information related to the samples (including tumor size and lymph node status at the time of the surgery) were retrieved from the pathology units of the University and Central hospitals. The search was restricted to invasive breast carcinomas, and only one breast cancer was included in the case of two separate breast cancer diagnoses.

Immunohistochemistry and antibodies

Sections, 3 μm thick, were cut from the formalin-fixed paraffin-embedded breast cancer tissue samples and stained with hematoxylin and eosin (H&E) for histological analysis. The H&E-stained breast tumor slides were reviewed for histological type and grade of each tumor according to the WHO classification of tumors of the breast (Lakhani et al., 2012). The TNM classification of malignant tumors was used to grade the extent of spread of the breast carcinomas (Lakhani et al., 2012).

For ER, PR, CK5/6, CK14, Ki-67, and HER2 immunostainings, the automated immunostaining machine BenchMark XT (Roche/Ventana) and ultraView Universal DAB Detection Kit (Roche/Ventana), were applied. The used antibodies and immunohistochemistry methods are listed in Table 3. Cases for *HER2*-amplification testing were selected on the basis of immunohistochemistry, and gene amplification was thereafter confirmed by *in situ* hybridization (Table 4).

Table 3 Antibodies used in immunohistochemistry

ANTIBODY	ER	PR	HER2	KI-67	CK 5/6	CK14
TYPE	rabbit mono- clonal	rabbit mono- clonal	rabbit mono- clonal	rabbit mono- clonal	mouse mono- clonal	mouse mono- clonal
CLONE	SP1	1E2	4B5	30-9	D5&16B4	LL002
SOURCE	Ventana 790-4325	Ventana 790-4296	Ventana 790-2991	Ventana 790-4286	Ventana 760-4253	Leica Biosystems NCL-L- LL002
DILUTION	ready-to- use	ready-to- use	ready-to- use	ready-to- use	ready-to- use	1:50
ANTIGEN RETRIEVAL	mild CC1*	mild CC1*	mild CC1*	mild CC1*	standard CC1*	standard CC1*
INCUBATION	24min 37°C	28min 37°C	24min 37°C	12min 37°C	28min 37°C	28min 37 °C

*Roche/Ventana CC1 pretreatment buffer

Table 4 *Her2/Chr17* double *in situ* hybridization was performed with the BenchMark XT. All reagents were from Roche/Ventana. ISH Protease 3 (780-4149) for 8 min was used as a pretreatment step, and *HER2* hybridization was performed at 52°C for 6 h and *Chr17* hybridization at 44°C for 2 h.

REAGENT (CATALOG NUMBER)

INFORM HER2 DNA PROBE (780-4332)

INFORM CHROMOSOME 17 PROBE (780-4331)

ULTRAVIEW SISH DETECTION KIT (780-001)

ULTRAVIEW ALKALINE PHOSPHATASE RED ISH DETECTION
KIT (800-504)

Case-control analysis

A case-control analysis for tumor type and grade was used to study the characteristics of the NF1 breast cancers. For each NF1 breast cancer, five age- and sex-matched controls without NF1 were randomly sampled from the data of Auri Biobank, Turku, Finland (<https://www.auriabiopankki.fi/?lang=en>). The clinical- and tissue-based tumor parameters were compared between the NF1 group and the control group.

Similarly, the 5-year breast cancer survival was studied with a five-fold randomly chosen control group without NF1 retrieved from the Auri Biobank. These

controls were matched with the NF1 study group for age, breast cancer diagnosis year, gender, and estrogen receptor status (+/-). Cumulative survival proportions were estimated using the Kaplan-Meier method, and the NF1 and control groups were compared using matched Cox's regression model. The data satisfied the proportional hazards assumption of Cox's regression.

In order to evaluate the occurrence and relevance of alterations in the *NF1* gene in breast cancers in the general population, data generated by the TCGA Research Network (<http://cancergenome.nih.gov/>) was used. The TCGA breast cancers with a normal or mutated/deleted *NF1* were compared using a chi-squared test for tumor characteristics and the Cox proportional hazards model for survival outcome analysis.

Statistical analyses were conducted using the statistical software R versions 3.2.2–3.3.0 (The R Foundation; <https://www.r-project.org/>) with lme4 (version 1.1-11), ordinal (version 2015.6-28), popEpi (version 0.2.1), and survival (version 2.38-3) packages.

4.6 Ethics (I-IV)

All studies were conducted according to the Declaration of Helsinki principles. Study I was carried out with the approval of the Ethics Committee of the Hospital District of Southwest Finland and the research permission of Turku University Hospital. Saliva samples were collected with informed consents of all participating patients. Studies II-IV were carried out with the approval of the Ethical Committee of the Southwest Finland Hospital District and the Ministry of Social Affairs and Health, Finland. A research permit was acquired from all hospitals that provided information. The breast cancer tissue samples in Study IV were collected and used with the approval of the National Supervisory Authority for Welfare and Health.

5 RESULTS

5.1 Mutation analysis of the *NF1* gene (I)

Our approach to *NF1* mutation analysis combined sequence capture methodology with high-throughput sequencing. Two separate sample sets, A and B, were prepared and sequenced to search for the ideal number of samples per sequencing run. Both sample sets contained control samples with known mutations to validate the mutation analysis method.

5.1.1 Sequencing

The average coverage of targeted regions in the sequencing runs was 41x and 74x for the sets A and B, respectively. Exon 1 was covered poorly in both sample sets, with average coverages of only 3x and 6x. Approximately 32-35% of the reads were mapped to the *NF1* gene on chromosome 17. The chromosomes with the most off-target reads were locations of known *NF1* pseudogenes.

5.1.2 Identification of mutations

Data analysis of all the possible variants resulted in identification of a total of 63 variants as potential *NF1* mutations in the sample sets of A and B. Seven variants, which were listed in dbSNP database, were evaluated individually; and their pathogenicity was excluded. In addition, two out of the seven SNPs were included in the Finnish SNP database. The remaining 39 and 17 variants in sets A and B were assessed individually with respect to homopolymer-related sequencing errors, which are typical for the sequencing platform used. A total of 10 homopolymer-related regions with a potential mutation were selected for the Sanger-sequencing, but all of these proved to represent false positives.

Ten mutations were identified and confirmed as putative disease-causing *NF1* mutations (Table 5). These included six substitutions, an insertion, and three small deletions. Five previously unknown mutations of patients S47, E66, E71, E396, and S97 were confirmed by Sanger sequencing. One previously known mutation in a control sample (patient E39) was excluded in the filtering due to low coverage. However, visual inspection of this area revealed a mutation in two out of nine reads. The known microdeletion of a control sample could not be detected reliably. Thus, mutations of four patients remained unsolved. To learn why these were not

revealed with our method, the four DNA samples were sent to a diagnostic laboratory in Erasmus Medical Center, Rotterdam, The Netherlands, which sequenced all *NFI* exons plus 30 bp intronic sequence and carried out MLPA analysis. These analyses revealed one additional mutation in patient S49 (c.844C>T, p.Gln282*) in *NFI* exon 6. In our experiment, this area of sample S49 had low coverage of only 11 reads, and the mutation was visible in one read, and thus could not raise the suspicion of a pathogenic mutation. Three mutations remained undiscovered by our protocol and by a diagnostic laboratory in Rotterdam.

Table 5 Discovered mutations in study I. Modified from Uusitalo et al. 2014.

SAMPLE	NF1 MUTATION FOUND (CDNA MUTATION CODE NM_000267.3)	TOTAL DEPTH	VARIANT FREQUENCY	PROTEIN OR MRNA LEVEL CHANGE	REGION	PREVIOUSLY DESCRIBED	CONTROL SAMPLE
SAMPLE SET A							
E46	c.7368dupC	79	0.54	frameshift	Exon 41	no	Yes
E13	c.1541_1542delAG	25	0.36	frameshift	Exon 10c	(Robinson et al., 1996)	Yes
S65	c.4537C>T	37	0.51	p. Arg1513*	Exon 27a	(Side et al., 1997)	Yes
S47	c.4922G>A	54	0.52	p. Trp1641*	Exon 28	(Brinckmann et al., 2007)	No
E66	c.2851-1G>A	34	0.47	(splicing)	Intron 16	no	No
E71	c.499_502delTGTT	37	0.51	frameshift	Exon 4b	(Osborn and Upadhyaya, 1999)	No
E396	c.3911T>G	34	0.68	p. Leu1304*	Exon 23.1	No	No
E579	No mutation found	–	–	–	–	No	No
S96	No mutation found	–	–	–	–	No	No
S594	No mutation found	–	–	–	–	No	No
SAMPLE SET B							
E27	c. 910C>T	102	0.46	p.Arg304*	Exon 7	(Upadhyaya et al., 2008)	Yes
S2122	c. 4914_4917delCTCT	152	0.43	p.Lys1640fs.	Exon 28	(Side et al., 1997)	Yes
E39A	No mutation found (c.5710G>T)	–	–	–	–	(Laycock-van Spyk et al., 2011)	Yes
		(9)	(0.22)	(p.Glu1904*)	(Exon 30)		
E38	Type 2 NF1 microdeletion	–	–	–	–	–	Yes
S97	c.1797G>A	25	0.44	p.Trp599*	Exon 12a	(Ars et al., 2000)	No
S49A	No mutation found (c.844C>T)	–	–	–	–	–	No
		(11)	(0.09)	(p. Gln282*)	(Exon 6)	(Gasparini et al., 1996)	

5.2 Formation of the population-based NF1 cohort and NF1 incidence in Finland (II-IV)

5.2.1 Search for NF1 patients (II)

The population-based NF1 cohort was formed by searching all hospital visits of patients with associated diagnosis of NF1 (Figure 4). All University and Central Hospitals in Finland together with the hospital discharge register and two small private organizations were searched for patients with any of the following diagnosis codes registered to the electronic hospital registers: ICD-10: Q85.00, Q85.0, Q85.09, Q85.01, Q85, or ICD-9: 2377A. The initial search yielded 2,335 individuals. Next, the patient records from all medical specialties and autopsy reports were manually reviewed. After a careful review of the medical records and after excluding other conditions and unclear cases, a total of 1,471 patients with a verified NF1 diagnosis were identified (Figure 4). NF1 diagnoses were based on NIH clinical diagnostic criteria (Stumpf et al., 1988) and/or mutation analyses. Excluded cases represented patients with suspected but not verified cases of neurofibromatosis, patients with neurofibromatosis type 2, or Legius syndrome, patients with only one neurofibroma or spinal neurofibromas without other NF1 diagnostic findings; segmental neurofibromatosis; other syndromes listed under Q85; or cases with typing errors in recording diagnosis codes. All persons with uncertain diagnosis were excluded. Of the 774 excluded patients, 209 had most probably NF1, but the diagnosis could not be verified from the current documents.

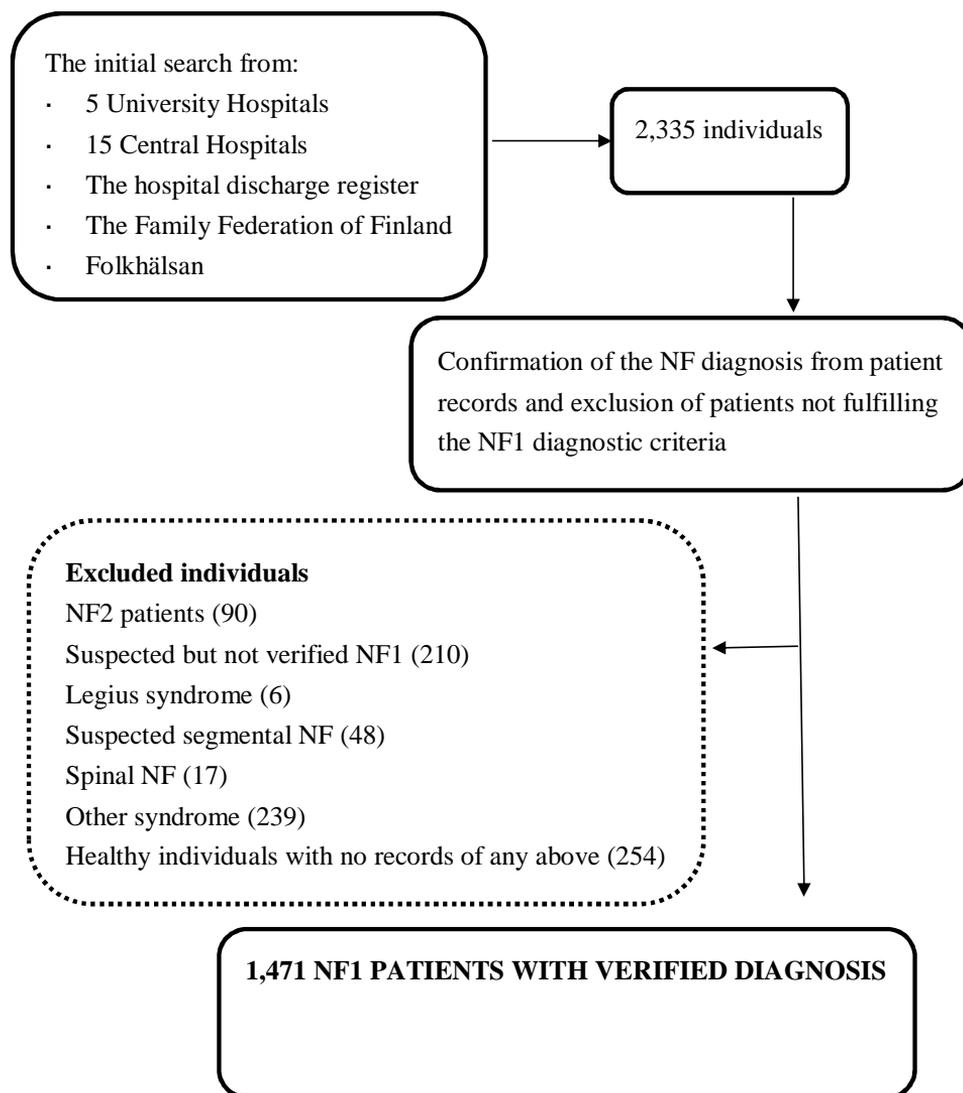


Figure 4 Collection of the NF1 patient cohort.

The NF1 patients were followed-up in different studies as shown in Table 6. The NF1 cohort inclusion criteria included a verified NF1 diagnosis and a hospital visit due to NF1 during the follow-up period. Despite similar cut-off dates, the number of the person years varies due to continued patient data collection resulting in more accurate cohort entry dates and thus a longer follow-up time. The age distribution of the 1,404 NF1 patients at the time of the cohort entry in Study III is listed in Table 7 and patients by the cohort entry year in Figure 5.

Table 6 The follow-up of the NF1 patients in different studies.

STUDY	N OF PATIENTS	FOLLOW-UP FROM	CUT OF DATE	N OF PERSON YEARS	N OF DEATHS
II	1393	1987–2011	31 Dec 2012	18,663	214
III	1404	1987–2011	31 Dec 2012	19,076	217
IV	1404	1987–2011	31 Dec 2013	20,248	232

Table 7 The age distribution of the 1,404 NF1 patients at the time of the cohort entry.

AGE (YEARS)	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-
NUMBER OF PATIENTS	523	207	157	176	139	115	60	22	5

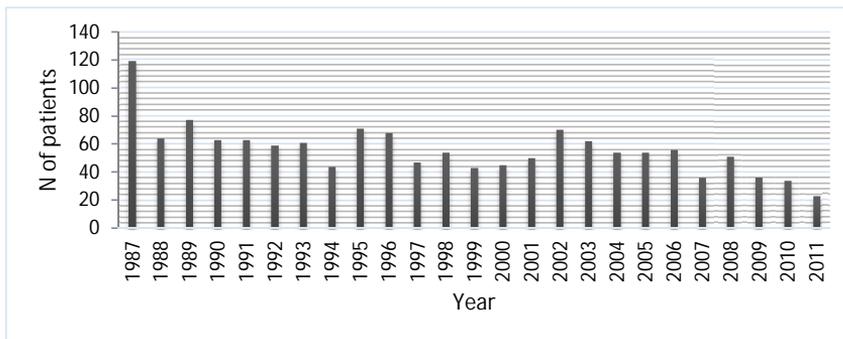


Figure 5 1,404 patients by cohort entry year (study III).

5.2.2 Higher NF1 incidence than previously reported (II)

The NF1 patients were born in 1905-2011 with a median year of birth of 1977. The median age of the patients was 33.3 years and the average 37.0 years at the end of the follow-up. To evaluate how common NF1 was, the birth incidence was calculated. The birth incidence according to birth year is shown in Figure 6. The highest incidence was observed in 1996 and was as high as 1/1,786. The peak incidence of a three-year period for NF1 was 1/1,871 noted in 1994-1996.

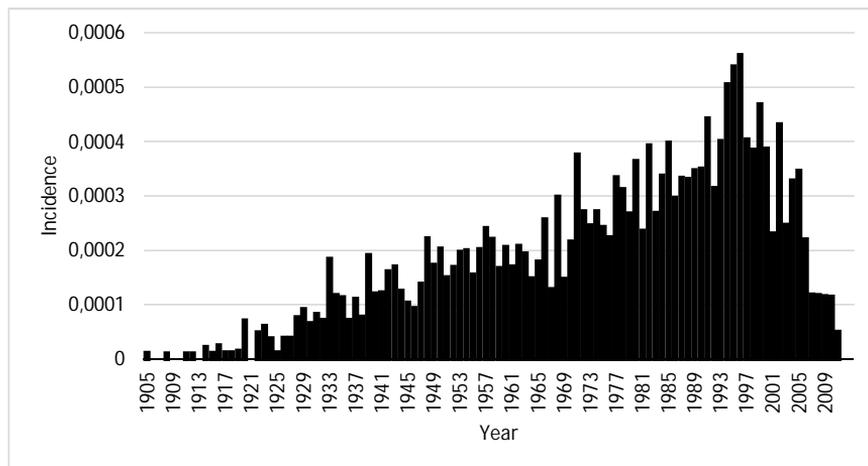


Figure 6 The observed birth incidence of NF1 according to the birth year. Modified from Uusitalo et al. 2015.

5.3 Mortality (II-IV)

The Causes of Death Register record linkage returned 248 deaths of the NF1 patients. Neurofibromatosis was mentioned in 77/248 (31%) of death certificates as an underlying cause of death, immediate cause of death, or contributing factor.

The overall SMR for NF1 patients was 3.22 (95% CI: 2.81-3.68) being significantly increased. The SMR for men was 2.66 (95% CI: 2.19-3.22) and for women 4.02 (95% CI: 3.32-4.83). In the age group <50 years, 86 deaths occurred, while the expected number was 13.64 resulting in a SMR of 6.31 (95% CI: 5.08-7.75). It should be noted that the SMR for women under 50 years was 10.44 (95% CI: 7.76-13.76), while SMR for men was 4.27 (95% CI 3.08-5.78). In the age group of >50 years, SMR was 2.43 (95% CI: 2.03-2.88) emphasizing that most of the excess mortality took place under 50 years of age.

Cancer-related deaths and cardiovascular deaths accounted for 58% and 13% of the excess mortality, respectively. SMR for pulmonary causes of death was 3.79 (95% CI: 2.11-6.32) accounting for 6% of the excess mortality. External causes were not increased compared to general population.

Cancer mortality

A total of 217 NF1 patients died during the follow-up period of 1987-2012, and 107 (49%) of these deaths were due to cancer. The SMR for cancer was 6.49 (95% CI: 5.32-7.85), and it was highest at young adulthood between ages 15-30 years.

Cancers with ICD-10 codes C47, C48-49, and C70-72 (cancers of autonomic nervous system, soft tissue, brain, and CNS) are here referred to as NF1-specific cancers. These are rare cancers in the general population but are traditionally considered typical for NF1. The highest SMRs were observed in these cancers in our study. The SMRs were 1,858 (95% CI: 1,294-2,583) for the autonomic nervous system, 87.2 (95% CI: 47.7-146) for soft tissue, and 30.2 (95% CI: 19.1-45.2) for brain and central nervous system tumors. However, when these cancer types were excluded, the SMR was 2.25 (95% CI: 1.57-3.13). This was explained by, for example, breast cancer and thyroid cancer deaths.

5.4 Cancer outcomes (III)

The cumulative risk for cancer in NF1 by the age of 30 years was 25.1% and 38.8% by the age of 50 years. The respective percentages in the Finnish population were 0.8% by 30 years and 3.9% by 50 years. In NF1, the cumulative risk for cancer by 50 years in women was 45.2% and in men 32.0%. The lifetime risk of cancer in NF1 was estimated to be 59.6%, while in the general population it is 30.8%.

A total of 244 cancers was observed, while the expected number was 48.5, resulting in a SIR of 5.03 (95% CI: 4.42-5.71). The highest numbers of cancer cases were observed in cancers with ICD-10 codes C47, C48-49, and C70-72 (cancers of autonomic nervous system, soft tissue, brain and central nervous system). These cancers are referred to as NF1-specific cancers. The SIR for autonomic nervous system tumors was 1,490 (95% CI: 1,091-1,987). A total of 46 autonomic nervous system tumors were observed in NF1 patients, while the expected number was 0.03. For soft tissue tumors, the SIR was 43.5 (95% CI: 25.4-69.7) and the SIR for brain and central nervous system tumors was 37.5 (95% CI: 30.2-46.0). More specifically, the SIR for MPNST was 2,056 with 58 observed cases in the whole NF1 cohort. The extremely high SIR results from a low expected number (0.03) of this rare cancer type in general population. When NF1-specific

cancers were excluded, the SIR for cancer remained elevated both for men (1.96) and women (1.98). Other cancers that had elevated SIRs were, for example, breast cancer and GIST.

Cancer survival

The 5-year survival of NF1 patients with cancer was significantly worse ($p=0.02$) compared to the survival of comparable cancer patients in the general population. The 5-year survival of NF1 cancer patients was 60.8% (95% CI: 54.1-68.4) versus 68.8% (95% CI: 62.6-75.6) in matched controls. Also, the 5-year survival of NF1 women was significantly worse, 64.3% (95% CI: 55.6-74.3%) when compared to that of controls being 77.2% (95% CI: 69.8-85.5%, $p=0.03$). The difference between NF1 and control men in survival was not significant in this respect: 56.2% (95% CI: 46.2-68.4%) in the NF1 group versus 57.8% (95% CI: 48.1-69.4%) in the control group ($p=0.21$). When NF1-specific cancers were excluded, the survival of cancer patients with NF1 was decreased when compared to the controls, 54.0% (95% CI: 43.1-67.8%) versus 67.5% (95% CI 57.5-79.3%, $p=0.01$).

5.5 Breast cancer in NF1 (IV)

Cross-referencing the NF1 cohort with the Finnish Cancer Registry brought out a total of 49 breast cancers diagnosed in 45 females and one male with NF1. Two of the female patients and the one male patient had two primary breast cancers diagnosed 8.3-19.6 years apart. The youngest NF1 patient was 28 and the oldest was 84 years at the time of the breast cancer diagnosis (median 49 years).

A total of 31 breast cancers were observed in the female NF1 cohort during the follow-up, which resulted in a SIR of 2.82 (95% CI 1.92-4.00). The SIR for breast cancer was highest in the youngest age groups: for women <40 years it was 14.25 (95% CI 6.51-27.04), while in women aged 40–50 years it was 2.60 (95% CI 0.95-5.65). The SIR of breast cancer was not significantly increased among women aged more than 50.

The cumulative risk for breast cancer in NF1 women by age of 40 years was 4.7% (95% CI: 1.6-7.6) and by age 50 years 7.8% (95% CI: 3.9-11.5). The respective numbers in the Finnish population were 0.45% and 2.1%. The estimated lifetime risk for breast cancer in NF1 women was 18.0% (95% CI: 11.4-24.1) compared to 9.7% in the general population. Women with NF1 had a 4.7% (95% CI: 1.5-7.9) risk of being diagnosed with breast cancer during years 30-39 of age, while the corresponding risk of the general population is only 0.34%.

Formalin-fixed paraffin-embedded breast cancer tissue samples from 26 NF1 breast cancer patients diagnosed in 1992–2011 were available for analysis. Patients were 29 to 84 years old at the time of the breast cancer surgery (median age 53 years). Of the breast tumors, 88.5% were ductal carcinomas and 7.7% lobular carcinomas. The breast cancer tumors were profiled using immunolabeling for ER, PR, Ki-67, CK5/6 and CK14, and *in situ* hybridization for *HER2*. The breast cancer subtype was luminal A in 15.4% of the NF1 samples, luminal B in 34.6%, *HER2*-amplified in 30.8%, and triple negative in 19.2%. The basal-like phenotype was seen in 15.4%. Grade III was the most frequent grade both in the youngest and in the oldest age groups.

Comparison of breast cancer characteristics between NF1 and controls

The breast cancer characteristics were compared between the NF1 patients and matched controls from the Auria Biobank. The results showed that the NF1 breast cancers were more often ER-negative (53.8% versus 20.9%, $p=0.001$) and PR-negative (65.4% versus 21.7%, $p<0.001$). *HER2*-amplification was overrepresented in the NF1 group (30.8% versus 9.6%, $p=0.006$). The NF1 breast cancers were also larger in size ($p=0.019$) by TNM-classification and of a higher grade ($p=0.050$).

Breast cancer survival

When breast cancer survival was analyzed alone, the results showed that the 5-year survival was lower in patients with NF1, 67.9% (95% CI: 50.6-91.0) versus 87.8% (95% CI: 75.0-100) in the control group ($p=0.004$). This could partly explain the low 5-year survival observed in women described in Study III.

Eleven (42.3%) out of the 26 NF1 patients died during the follow-up. Survival was similar in the cases where tissue samples could not be retrieved. Because the NF1 breast cancers were more often ER-negative, survival analysis was carried out again using a control group from the Auria Biobank matched for age, breast cancer diagnosis year, gender, and also ER status. Again, the overall survival was worse in the NF1 group: 5-year survival was 68.1% (95% CI: 52.0-89.1) in the NF1 group and 82.0% (95% CI 75.5-88.9) in the control group ($p=0.053$).

Relevance of NF1 alterations in breast cancers of the TCGA dataset

Mutations or deletions of the *NF1* gene were observed in 33% of the TCGA breast cancers. Patients with an *NF1* deletion or mutation were slightly younger than those without a mutation (56.8 years vs. 59.2 years, $p=0.013$). *NF1* mutations and deletions were associated with decreased survival: 5-year survival was 86.5% (95% CI: 81.7-91.6) for the breast cancer patients with a normal *NF1* gene status

but only 77.1% (95% CI 70.5-84.3) for those with an *NFI* mutation or deletion (95% CI 1.14-3.00, $p=0.014$). Comparisons of the receptor status showed that cancers with an *NFI* mutation or deletion were significantly more often ER-and PR-negative and *HER2*-amplified than cancers with a normal *NFI* gene ($p<0.001$ for all).

Restricting the analysis to *NFI* mutations only resulted in a similar trend in survival (hazard ratio 3.2, 95% CI: 1.1 to 9.0, $p=0.031$) and receptor status. However, the differences in ER and *HER2* status were not significant due to the low number of tumors in the TCGA dataset harboring *NFI* mutations.

6 DISCUSSION

6.1 New *NFI* mutation analysis method exploiting NGS (I)

The mutation analysis of the large *NFI* gene is known to be challenging not only because of the size of the gene but also due to the lack of mutational hotspots, the occurrence of a very diverse spectrum of mutation types (Messiaen and Wimmer, 2008), and the presence of *NFI* pseudogenes. Blood, which is currently mainly used as a source of DNA/RNA (Croonen et al., 2012, Messiaen and Wimmer, 2008), requires invasive sampling methods. Our approach used noninvasive and patient-friendly saliva samples as a source of DNA. Our method was one of the first studies and a proof-of-principle that the sequence capture methodology combined with high-throughput sequencing is applicable to *NFI* mutation analysis. The DNA extracted from saliva was high-quality. The saliva samples can even be collected by the patients at home, and because of the stability of the samples, they can easily be shipped to the laboratory without any need for cold storage.

We used the Roche/454 GS Junior sequencing platform, which was the first next-generation DNA sequencer on the market and the best option at the beginning of our study. The Roche/454 GS platform has been used also in another small scale *NFI* study (Chou et al., 2010). The average read length produced by the sequencer is 400 bases, which is long compared to other systems. Most of the recent *NFI* mutation analysis studies (Calì et al., 2017, Cunha et al., 2016, Pasmant et al., 2015) have been based on the Ion Torrent PGM sequencing platform. This sequencing platform is faster and has lower costs per produced Mb, which supposedly makes it a tempting option. The development of sequencing technologies is fast at the moment, and most likely there will be even better options for mutation analysis in the future.

A putative mutation was discovered in ten samples out of the total of 16 in our study. One patient had a previously detected *NFI* microdeletion. In this sample, the percentage of reads mapping to the *NFI* gene was lower compared to other samples. Thus, we can speculate that the difference between the proportions of reads mapped to the *NFI* gene as opposed to the pseudogenes could be distinct enough to raise suspicion of a microdeletion, which could be verified with, for example, MLPA.

We were not able to detect every *NFI* mutation in our analysis. Two patients had low coverage at the mutation site, and thus the results could not have raised suspicion of a pathogenic mutation. Three mutations remained undiscovered by

our protocol but also by an established international diagnostic laboratory. These patients had a clinical diagnosis of NF1 but may represent mosaic NF1 or patients with deep intronic *NF1* mutations. In somatic mosaicism, the *NF1* mutation may not be found in blood or saliva samples (Tinschert et al., 2000).

The most sensitive *NF1* mutation analysis method at the moment is still the multistep pathogenic variant detection protocol that combines analysis of genomic DNA and cDNA (mRNA) together with testing for whole-gene or exon copy number changes (Messiaen and Wimmer, 2008). This approach identifies more than 95% of *NF1* pathogenic variants in individuals fulfilling the NIH diagnostic criteria (Messiaen et al., 2000). The reported methods based on genomic DNA only, like our method, have a somewhat lower diagnostic sensitivity with 60-90% of pathogenic variants detected (Calì et al., 2017, Maruoka et al., 2014, Pasmant et al., 2015, van Minkelen et al., 2014).

Although the NF1 diagnosis is still often based only on clinical symptoms, the mutation analysis is valuable in many situations. With the help of mutation analysis, the NF1 diagnosis can be made early and follow-up started. This is especially the case in young children who may only partially fulfill the clinical criteria. Adults with atypical clinical presentation of NF1 are another patient group that will benefit from molecular diagnostics. Also, prenatal testing and preimplantation genetic diagnosis are possible if the parental mutation is known.

6.2 Total population-based NF1 cohort (II-IV)

The total population-based NF1 cohort in Studies II-IV was formed by searching inpatient and outpatient hospital visits of patients with a diagnosis or an associated diagnosis for NF1. Thus, all levels except the primary health care were covered. The patient search was done using ICD-9 and ICD-10 diagnosis codes, and all medical specialties were included. The ICD-9 codes have been in use from 1987 to 1995 and ICD-10 from 1996 onwards. Thus, the first possible hospital visit of an NF1 patient with the diagnosis codes in question registered in hospital records may have taken place in 1987. After the initial search, the NF1 diagnoses were manually verified from patient records and death certificates. Only patients fulfilling the NIH diagnostic criteria were included, and other individuals were excluded from the study. Excluded cases represented healthy family members, patients with only one neurofibroma or cases with clerical errors in recording diagnosis codes.

A hospital admission search with an NF1 diagnosis code has been used also in some previous studies as a patient ascertainment method (Seminog and Goldacre,

2013, 2015). However, in these particular studies, the NF1 diagnosis has not been confirmed with other methods. Based on our results, hospital searches using diagnosis codes result in data that must be reviewed by other methods such as examining the medical records to verify the correct diagnosis of the patients. Furthermore, most of the studies before the diagnostic criteria of NF1 and NF2 were agreed at the NIH 1987 NF Consensus Conference contained mixed samples of both NF1 and NF2 patients.

The NF1 patients in our study were followed-up starting from the first hospital visit for NF1 and ended with death, emigration, or by reaching the end of the follow-up period. We were only able to find patients who were alive and had a hospital visit for NF1 from 1987 onwards. Patients who have had their NF1 diagnosis before 1987, and no NF1-related hospital visits since have been missed in our patient search.

The Finnish NF1 cohort is hospital-based. The patient search from secondary (Central Hospitals) and tertiary (University Hospitals) referral centers may have led to an under ascertainment of patients having mild NF1 or NF1 with otherwise good health. However, the Finnish health care system is equally available to everyone regardless of income, and the network of secondary referral centers covers the whole country. Thus, the ascertainment should not be dramatically influenced by the place of residence or the income of the patients.

Our study, based on the Finnish population of 5.5 million, is the first total-population NF1 study and has a high level of ascertainment. Previous population-based studies have been limited to regions of Gothenburg, Sweden (population 440,082) (Samuelsson and Axelsson, 1981); southeast Wales (population 668,100) (Huson et al., 1989); Dunedin, New Zealand (population 113,700) (Fuller et al., 1989); northeast Italy (population 2,375,304) (Clementi et al., 1990); and northern Finland (population 733,037) (Poyhonen et al., 2000).

6.3 NF1 birth incidence in Finland (II)

The NF1 incidence has been estimated in few studies, but the published frequencies vary from study to study probably because of the differences in ascertainment methods. The generally accepted NF1 incidence has been around 1/2,600-1/3,000 (Evans et al., 2010, Huson et al., 1989, Lammert et al., 2005, Poyhonen et al., 2000). The NF1 prevalence is studied more but with highly varied estimates. Prevalence, which is the actual number of cases alive with a disease at a given moment, is affected by mortality. In NF1, the prevalence declines by

increasing age, because mortality is high already at a young age (Kallionpää et al., 2017).

In our study, the high incidence of 1/1,786 observed in 1996 was based on 34 NF1 patients born that year. The observed NF1 incidence of 1/1,871 in a three-year period during 1994-1996 was based on 101 NF1 patients. The observed high NF1 incidence in the 1990s may be related to enhanced efficacy in making the diagnosis, which is likely due to the general awareness of NF1 and the timing of our data collection, which included hospital visits since 1987. Established in 1987, the NIH diagnostic criteria have facilitated diagnosing NF1 in patients born since 1980s (Stumpf et al., 1988). It has been reported that the mean age at diagnosis of NF1 has dropped from 20 years among patients born in the 1960s to less than six years for patients born in the 1980s in Finland (Poyhonen, 2000).

The patients born during the mid-1990s were old enough to fulfill the NIH criteria by the time of our patient search, and the children born in the 1990s have been diagnosed and followed-up by specialists and are thus better covered in this study. The incidence of ~1/1,900 represents a minimum incidence, because only verified diagnoses were included, and the cohort may still not include patients with very mild or undiagnosed NF1. The real birth incidence of NF1 must be closer to 1/2,000 than to the generally accepted 1/3,000.

Although we observed an increasing number of verified NF1 diagnoses per live births from 1905 to 2000, there is no reason to assume that the birth incidence of NF1 would display major increasing or decreasing tendencies. Because the cases had to be alive in 1987 to be included in our search, immortality bias affects the estimates in the older age groups. However, our results are not affected by such a selection in the younger ones. The low observed incidence among those born between 1905 and 1975 is likely to result partly from a high mortality already under 50 years of age. It is also possible that some patients born before the 1970s, who have got their diagnosis before 1987, have not needed treatment for NF1 in any medical specialty from 1987 onward. Thus, they would not have been in the hospital records that we searched. The youngest age groups born in the 2000s may not yet fulfill the diagnostic criteria, and many of these are included in the group of suspected NF1.

6.4 Interpretation of the NF1 cancer and mortality study results (II-III)

The cancer incidence in NF1 was evaluated in subprojects III and IV with a retrospective register-based total population study. The cancer risk of the Finnish

population is similar to other Western countries showing no major genetic predisposition that could cause bias. We observed a five-fold increase in cancer incidence, which is the highest reported so far for NF1 with the patients under 50 years having the highest excess cancer risk.

The Causes of Death Register run yielded 248 NF1 deaths, of which 214 occurred in 1987-2012. Neurofibromatosis was mentioned in 77/248 (31%) of death certificates as an underlying cause of death, immediate cause of death, or contributing factor. If our study approach had been based on patient ascertainment through death certificates, then only this small proportion of the patients would have been found. Evans et al. reported a rather similar percentage (36%) of NF1 recorded on the death certificates of the NF1 patients (Evans et al., 2011). This must have been also the case in the U.S. study (Rasmussen et al., 2001). They found 3,770 NF1 deaths from a total of 3,272,122, which is only 1 in 8,679 deaths. Death certificate studies have many opportunities, but in the case of NF1, are likely to miss the majority of the patients.

MPNST is a hallmark complication of NF1. It is an aggressive soft tissue sarcoma with half of all cases developing in individuals with NF1 (Evans et al., 2002). We observed a total of 58 cases of MPNST and as many as 41 deaths due to this condition during the follow-up. The lifetime risk of MPNST was estimated to be 15.8% in our study, which is in-line with previous studies (Evans et al., 2002). The excess risk for central nervous system tumors was also expected. These were also a common cause of death in NF1 patients.

In contrast to findings in previous studies (Niemeyer et al., 1997, Stiller et al., 1994), we could not detect increased incidence of JMML. The number of individual cancer cases and follow-up time may have been too limited to allow definite conclusions of this rare cancer type.

6.5 Breast cancer (III-IV)

Increased incidence of breast cancer in NF1 women has been reported in some publications and numerous case reports. Our results show that the breast cancer incidence in NF1 women aged <40 years is increased with a cumulative risk of 4.7% while being 0.45% in the control population. Screening for breast cancer has been suggested for NF1 women from the age of 40 years (Evans, 2012, Sharif et al., 2007). Our results do support this, but more evidence is needed about the effectiveness and safety of different screening methods in NF1 patients.

The main finding of our breast cancer study was that NF1 breast cancers are often associated with poor survival and unfavorable prognostic factors, such as ER- and PR-negativity and *HER2*-amplification. The study shows, however, that the poor prognosis of the NF1-related breast cancer is not explained by histopathological subtype only. The poor prognosis of NF1 breast cancer patients may partly be due to an advanced stage at the time of diagnosis. The breast cancers of the NF1 patients were larger and of higher grade at the time of surgery, although the slightly worse lymph node status of NF1 patients was not statistically significant. It has been speculated that subcutaneous neurofibromas may interfere with the diagnostics of breast cancer (Da Silva et al., 2015, Evans, 2012), and this might cause a delay in NF1 breast cancer diagnosis.

The poor outcome of breast cancer patients with NF1 could be explained by uninhibited Ras signaling due to a loss of neurofibromin functionality and thus increased PI3K and Raf/MAPK/ERK activity. Increased PI3K activity has been related to low survival and resistance to hormone treatment in ER-negative breast cancer (Yang et al., 2016), and elevated Ras/MAPK/ERK activity has been related to metastasis and low survival in both ER-positive (Wright et al., 2015) and ER-negative breast cancer (Giltane and Balko, 2014). In the absence of inhibition by neurofibromin, enhanced and prolonged Ras stimulation by cell membrane tyrosine kinase receptors, such as EGFR and HER2, can activate several transcription factors, such as c-myc and ETS1. This may strongly facilitate breast cancer development and progression in NF1 patients (Rad and Tee, 2016).

Another explanation for low breast cancer survival could be that the inactivation of *NF1* in breast cancer causes resistance to drug therapy. Silencing of the *NF1* gene has been shown to confer a tamoxifen-resistant phenotype in one study (Mendes-Pereira et al., 2012).

6.6 Strengths and limitations (I-IV)

The *NF1* mutation analysis method in our study was a proof-of-principle that the new sequencing methods together with sequence capture methodology can be applied to *NF1* gene mutation analysis. The method still requires more validation and modifications to be used in clinical purposes. The saliva samples used were a good and noninvasive source of DNA. Some of the samples were even collected by the patients at home or at other locations outside hospital. Although, blood could have been used as a starting point as well and could have been required if the need for DNA amount would have been great.

The false-negatives are a central problem of every mutation analysis method, this was also the case in our approach. The occurrence of a wide spectrum of mutation types and the lack of mutational hotspots makes the molecular analysis of the *NF1* gene exceptionally difficult. Some mutation types would require a RNA-based method to be discovered. The sequencing platform used in our study did cause a number of homopolymer-related sequencing errors that had to be studied individually with other methods. Another difficulty that was encountered was the *NF1* pseudogene sequences in the human genome. These lowered the coverage of the *NF1* gene region considerably but could be handled in data analysis. Specifically, the relatively long reads produced by the sequencing platform in question facilitated the correct mapping of the pseudogene reads.

Studies II-IV utilized a total population-based NF1 cohort. The cohort represents the most comprehensive and completely ascertained NF1 cohort to date. Although, the patients with a very mild form of NF1, and patients who have been diagnosed with NF1 but have not been referred to secondary or tertiary hospitals may have been missed in our study. The NF1 diagnoses were verified from medical records and autopsy reports to identify cases fulfilling the NIH clinical diagnostic criteria. The NF1 patients were born between 1905 and 2011 with a median year of birth of 1977. The median age of the patients was 33.3 years, and the average age was 37.0 years at the end of the follow-up. The lack of older NF1 patients in the NF1 cohort may partly be explained by early deaths of the patients. Our patient ascertainment was based on inpatient and outpatient hospital visits during the years 1987-2012, which favors the inclusion of patients born after 1980's. The NF1 patients born before 1987 may not have had a proper diagnosis or hospital visits because of NF1 during our follow-up period. Small children on the other hand may not have developed enough symptoms needed for the clinical diagnosis during our follow-up period. In our study, 209 patients were suspected to have NF1 but were excluded from the cohort. The majority of these were children who did not fulfill the NIH criteria by the completion of the study. If the study was continued, part of these individuals could probably be included in the study cohort. We were not able to collect data on NF1 diagnosis dates, but based on a small amount of data available for those born in 1980s, the diagnosis age varied from less than one year to over 30 years. A total of 8% (9/113) of these patients got their diagnosis after 20 years of age. Thus, there might still be many NF1 patients born in the 1990s without a correct diagnosis. Presumably some patients will not get the NF1 diagnosis until after having a child with NF1.

Earlier studies have not estimated the NF1 incidence in a total population. Our study was the first whole population study with high levels of ascertainment and the largest patient cohort reported so far. There is no reason to assume that the observed high incidence in our study would be because NF1 is more common in

Finland. Instead, the observed high incidence may reflect enhanced diagnostics of NF1 and efficient ascertainment of our study. Healthcare in Finland is publicly funded and available to all; and thus, the patient search exploiting hospital visits is efficient.

Studies II-IV were conducted on a national scale using multiple registries. The Finnish registration of deaths and emigration is known to be comprehensive, and the cancer data was based on near complete cancer registration systems used in Finland (Gissler and Haukka, 2004, Leinonen et al., 2017, Teppo et al., 1994). Data stored in large national registries provides great opportunities for research, but registry data contains some limitations. For example, some clinically relevant information, such as detailed cancer histology, are not completely accurately registered into the Cancer Registry at the moment and must be collected from different sources (Teppo et al., 1994). Also, information on, for example, common cancer risk factors cannot be retrieved from the current registries.

The number of individual cancer cases and/or follow-up time in the cancer study may have been too limited to allow definite conclusions in many individual cancer types. However, the overall results of the study were convincing and reliable.

The NF1 breast cancer study (IV) was based on a small sample size, but the results were definite. A total of 31 breast cancers were diagnosed in NF1 patients during the follow-up period. Of these, we were able to retrieve a tumor sample from 26 cases. We reported that NF1-related breast cancer has poor prognosis, which is not solely explained by occurrence at young age or by histopathological type. There is no reason to presume that the 26 would not represent the whole group. The survival was similar in the cases where tissue samples could not be retrieved.

6.7 Suggestions for future research

The *NF1* mutation analysis method developed in our study did work but could be further improved. The design of the sequence capture would have allowed also other genes to be included to the target area. For example, genes that cause syndromes that have overlapping phenotypes to NF1, such as *SPRED1* gene, could be useful to analyze simultaneously with *NF1*. This could be exploited in the future. We could not detect every mutation in our study. The best way to increase both sensitivity and specificity of the method would be to increase coverage by sequencing a smaller number of samples per run (De Leeneer et al., 2011). Although the *NF1* mutation analysis method described in our study is not validated for clinical application, it was one of the first NF1 studies with NGS and paves the way for new and cost-effective approaches in *NF1* mutation analysis.

Our follow-up of the NF1 cohort was started earliest from 1987 and ended by 2013. The follow-up of the Finnish NF1 cohort could be continued, and the cancer incidence and mortality studied again in, for example, ten years. This could reveal more information about rare cancer types with too few cases or uncertain relation to NF1 in our study.

The Ras/MAPK pathway plays an important role in cancer biology in general. It is also a major target for anti-cancer drugs. However, the presence of an *NF1* mutation may confer resistance to several therapeutic agents. At the moment, *NF1*-associated drug resistance has been suggested at least to RAF and EGFR inhibitors in melanoma, lung cancers, and neuroblastoma (Hölzel et al., 2010, Maertens et al., 2013, Network, 2014). This is an extremely important subject that has an impact on cancer survival of the patients and should be studied more in the future.

The conclusion that NF1-related breast cancer has poor prognosis, which is not solely explained by the occurrence at a young age or by histopathological type, should enhance the follow-up practices for NF1 patients. The strong association between breast cancer and NF1 should also impact breast cancer research in general, because *NF1* mutations are frequent in breast cancer in the general population. The same holds true for cancer in general. The more cancers of the NF1 patients are studied the more information we get about sporadic cancers with *NF1* mutations.

7 CONCLUSIONS

NF1 is a multisystem disease causing a varying combination of symptoms. The present study was performed to have a better understanding about the syndrome. We used a national NF1 cohort and retrospective register-based total population study to evaluate the epidemiology and cancer incidence of NF1 in Finland. The intention of this study was also to advance knowledge about breast cancer that is related to NF1 and to develop a new method for NF1 molecular diagnostics.

The following conclusions were made on the basis of the results of the present study:

1. Sequence capture methodology combined with high-throughput sequencing is applicable to *NF1* mutation analysis.
2. NF1 incidence is higher than previously reported being 1/2,000. NF1 also causes significant mortality, which is largely explained by malignant neoplasms at a relatively young age.
3. NF1 cancer incidence is higher than previously reported. We observed a five-fold increase in cancer incidence in the NF1 population, which is the highest cancer incidence reported so far. Cancers of the NF1 patients have worse prognosis than the corresponding cancers in the general population. Active surveillance of the patients is needed to detect the tumors early and to improve the prognosis.
4. The risk for breast cancer in NF1 is elevated, particularly under age 40. NF1-related breast cancer has poor prognosis, which is not solely explained by occurrence at young age or by histopathological type. Awareness of the NF1-related breast cancer risk should be raised.

ACKNOWLEDGEMENTS

This thesis work was carried out at the Department of Cell Biology and Anatomy, Institute of Biomedicine and Department of Dermatology and Venereology, University of Turku. I wish to thank all the present and former heads of the Departments and Institute: Professors Kalervo Väänänen, Juha Peltonen, Sari Mäkelä, and Veli-Matti Kähäri for providing the facilities and supportive working environment for the current study.

I am truly grateful for my supervisors Professor Juha Peltonen and Docent Sirkku Peltonen for introducing me to the world of science. I truly appreciate all the time you have dedicated to this thesis work. Thank you for your support and guidance during these years.

Professor Jaakko Kaprio and Docent Sirpa Kivirikko are gratefully acknowledged for carefully reviewing this thesis and providing valuable criticism. I would also like to thank my supportive steering group: Professors Klaus Elenius and Leena Haataja. The annual meetings gave me motivation and enthusiasm to carry on.

All my co-authors Annika Brandt, Olli Carpen, Anna Hammis, Riikka Huovinen, Pirkko Härkönen, Eeva-Mari Jouhilahti, Roope Kallionpää, Anna Koffert, Pauliina Kronqvist, Samu Kurki, Jussi Leppävirta, Ville-Veikko Mäkelä, Elina Palonen, Janne Pitkäniemi, Eero Pukkala, Minna Pöyhönen, Matti Rantanen, Vincent Riccardi, Juhani Soini, Sakari Suominen, Tero Vahlberg, Jussi Vahtera, and Heli Ylä-Outinen are warmly thanked for their contribution to this project. Especially Roope and Jussi, I could not have managed without your help.

I would like to thank all the present and former NF1 team members for your support: Vesa Aaltonen, Maria Alanne, Eetu Heervä, Miso Immonen, Eeva-Mari Jouhilahti, Roope Kallionpää, Anna Koffert, Jussi Leppävirta, Paula Pennanen, Vivian Visnapuu and Heli Ylä-Outinen. Paula is especially thanked for being such a great travelling companion, it was fun to explore places with you. Eeva-Mari and Maria, I'm grateful for all the unforgettable moments in and outside the office, thank you for being good friends. Eeva-Mari is especially warmly thanked for all the long walks and discussions during our first year as new mothers, your support was invaluable. Eeva-Mari, Sirkku and Vesa are also thanked for providing clinical pictures for this thesis.

I would also like to thank the Turku Doctoral Programme of Molecular Medicine (TuDMM), and its director Kati Elima for supporting my participation in scientific conferences and for organizing interesting scientific and social events. I also wish to thank the TUDMM board members for the time on the board.

I wish to thank all the former and present colleagues at the Institute of Biomedicine, especially Milja Arponen, Ameya Bendre, Matthieu Bourger, Natalija Eigeliene, Teresa Elo, Terhi Heino, Teuvo Hentunen, Mirkka Hirvonen, Kaisa Ivaska-Papaioannou, Tiina Kähkönen, Salla Laine, Tiina Laitala-Leinonen, Niina Lopenen, Jussi Mäkilä, Jorma Määttä, Jonas Nyman, Jorma Paranko, Lauri Polari, Kati Tarkkonen, Johanna Tuomela, Liina Uusitalo-Kylmälä, Anu Wiklund, and Emrah Yatkin. Thank you for the friendly atmosphere and all the scientific and non-scientific discussions during these years.

I am grateful for Iris Dunder, Soili Huhta, Elina Tammi, Petra Heikkilä, Outi Irjala, and Pia Tahvanainen for excellent administrative and secretary help. Taina Malinen, Soili Jussila, Anneli Kurkela, Miso Immonen, Pirkko Rauhamäki, Krista Hänninen, Ljudmilla Shumskaja, and Sinikka Collanus are thanked for their technical assistance. Professor Robert M. Badeau, Ph.D. of Aura Professional English Consulting, Ltd. (www.auraenglish.com) is kindly acknowledged for providing this thesis's language checking.

This thesis work was financially supported by the Turku University Foundation, Turku University Hospital EVO funding, the Finnish Cancer Society, the Cancer Foundation of Southwestern Finland, the Ida Montini Foundation, and the Finnish Cultural Foundation. All the patients who provided samples to the study are warmly thanked.

My friends, especially Heidi, Anni, Leena, Eerika and Ida are warmly thanked for always being there for me. Thank you for all the memorable moments together, your friendship is priceless.

I owe my deepest gratitude to my family, especially to my parents, my siblings and my family-in-law. Thank you for your love and support. Elisa and Aulis are especially thanked for taking care of Ebba so that I could complete this project.

Finally, I wish to thank my husband Eero. Your love, companionship, positivity, support, and sense of humor truly fulfil my life! You and our daughter Ebba bring me happiness and remind me of the true meaning in life every day.

Turku, February 2018

Elina Uusitalo

REFERENCES

- Ahmadian MR, Hoffmann U, Goody RS, Wittinghofer A. Individual rate constants for the interaction of Ras proteins with GTPase-activating proteins determined by fluorescence spectroscopy. *Biochemistry* 1997;36(15):4535-41.
- Airewele GE, Sigurdson AJ, Wiley KJ, Frieden BE, Caldarera LW, Riccardi VM, et al. Neoplasms in neurofibromatosis 1 are related to gender but not to family history of cancer. *Genet Epidemiol* 2001;20(1):75-86.
- Alberts B, Johnson A, Lewis J, Morgan D, Raff M, Roberts K, et al. Cancer. In: *Molecular Biology of the Cell*. Sixth edition ed: Garland Science; 2015. p. 1091-144.
- Antinheimo J, Sankila R, Carpén O, Pukkala E, Sainio M, Jääskeläinen J. Population-based analysis of sporadic and type 2 neurofibromatosis-associated meningiomas and schwannomas. *Neurology* 2000;54(1):71-6.
- Antoniou A, Pharoah PD, Narod S, Risch HA, Eyfjord JE, Hopper JL, et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet* 2003;72(5):1117-30.
- Ars E, Serra E, García J, Kruyer H, Gaona A, Lázaro C, et al. Mutations affecting mRNA splicing are the most common molecular defects in patients with neurofibromatosis type 1. *Hum Mol Genet* 2000;9(2):237-47.
- Balcer LJ, Liu GT, Heller G, Bilaniuk L, Volpe NJ, Galetta SL, et al. Visual loss in children with neurofibromatosis type 1 and optic pathway gliomas: relation to tumor location by magnetic resonance imaging. *Am J Ophthalmol* 2001;131(4):442-5.
- Baser ME, Friedman JM, Aeschliman D, Joe H, Wallace AJ, Ramsden RT, et al. Predictors of the risk of mortality in neurofibromatosis 2. *Am J Hum Genet* 2002;71(4):715-23.
- Bates JE, Peterson CR, Dhakal S, Giampoli EJ, Constone LS. Malignant peripheral nerve sheath tumors (MPNST): a SEER analysis of incidence across the age spectrum and therapeutic interventions in the pediatric population. *Pediatr Blood Cancer* 2014;61(11):1955-60.
- Bausch B, Borozdin W, Mautner VF, Hoffmann MM, Boehm D, Robledo M, et al. Germline NF1 mutational spectra and loss-of-heterozygosity analyses in patients with pheochromocytoma and neurofibromatosis type 1. *J Clin Endocrinol Metab* 2007;92(7):2784-92.
- Bengesser K, Cooper DN, Steinmann K, Kluwe L, Chuzhanova NA,

- Wimmer K, et al. A novel third type of recurrent NF1 microdeletion mediated by nonallelic homologous recombination between LRRC37B-containing low-copy repeats in 17q11.2. *Hum Mutat* 2010;31(6):742-51.
- Bernards A, Snijders AJ, Hannigan GE, Murthy AE, Gusella JF. Mouse neurofibromatosis type 1 cDNA sequence reveals high degree of conservation of both coding and non-coding mRNA segments. *Hum Mol Genet* 1993;2(6):645-50.
- Bollag G, McCormick F, Clark R. Characterization of full-length neurofibromin: tubulin inhibits Ras GAP activity. *EMBO J* 1993;12(5):1923-7.
- Borberg A. Clinical and genetic investigations into tuberous sclerosis and Recklinghausen's neurofibromatosis; contribution to elucidation of interrelationship and eugenics of the syndromes. *Acta Psychiatr Neurol Scand Suppl* 1951;71:1-239.
- Boyanapalli M, Lahoud OB, Messiaen L, Kim B, Anderle de Sylor MS, Duckett SJ, et al. Neurofibromin binds to caveolin-1 and regulates ras, FAK, and Akt. *Biochem Biophys Res Commun* 2006; 340(4):1200-8.
- Brems H, Chmara M, Sahbatou M, Denayer E, Taniguchi K, Kato R, et al. Germline loss-of-function mutations in SPRED1 cause a neurofibromatosis 1-like phenotype. *Nat Genet* 2007; 39(9):1120-6.
- Brems H, Beert E, de Ravel T, Legius E. Mechanisms in the pathogenesis of malignant tumours in neurofibromatosis type 1. *Lancet Oncol* 2009; 10(5):508-15.
- Brems H, Legius E. Legius syndrome, an Update. Molecular pathology of mutations in SPRED1. *Keio J Med* 2013;62(4):107-12.
- Brewer HR, Jones ME, Schoemaker MJ, Ashworth A, Swerdlow AJ. Family history and risk of breast cancer: an analysis accounting for family structure. *Breast Cancer Res Treat* 2017;165(1):193-200
- Brinckmann A, Mischung C, Bässmann I, Kühnisch J, Schuelke M, Tinschert S, et al. Detection of novel NF1 mutations and rapid mutation prescreening with Pyrosequencing. *Electrophoresis* 2007;28(23): 4295-301.
- Brody LC, Biesecker BB. Breast cancer susceptibility genes. BRCA1 and BRCA2. *Medicine* (Baltimore) 1998; 77(3):208-26.
- Buermans HP, den Dunnen JT. Next generation sequencing technology: Advances and applications. *Biochim Biophys Acta* 2014;1842(10):1932-41.
- Cali F, Chiavetta V, Ruggeri G, Piccione M, Selicorni A, Palazzo D, et al. Mutation spectrum of NF1 gene in Italian patients with neurofibromatosis type 1 using Ion Torrent PGM™ platform. *Eur J Med Genet* 2017;60(2):93-9.
- Campian J, Gutmann DH. CNS Tumors in Neurofibromatosis.

- J Clin Oncol* 2017;35(21):2378-85.
- Cancer Genome Atlas Network. Genomic Classification of Cutaneous Melanoma. *Cell* 2015; 161(7): 1681-96.
- Castellanos E, Gel B, Rosas I, Tornero E, Santín S, Pluvinet R, et al. A comprehensive custom panel design for routine hereditary cancer testing: preserving control, improving diagnostics and revealing a complex variation landscape. *Sci Rep* 2017;7: 39348.
- Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2012;2(5):401-4.
- Chou LS, Liu CS, Boese B, Zhang X, Mao R. DNA sequence capture and enrichment by microarray followed by next-generation sequencing for targeted resequencing: neurofibromatosis type 1 gene as a model. *Clin Chem* 2010;56(1):62-72.
- Cichowski K, Jacks T. NF1 tumor suppressor gene function: narrowing the GAP. *Cell* 2001;104(4):593-604.
- Clementi M, Barbujani G, Turolla L, Tenconi R. Neurofibromatosis-1: a maximum likelihood estimation of mutation rate. *Hum Genet* 1990;84(2):116-8.
- Clemons M, Goss P. Estrogen and the risk of breast cancer. *N Engl J Med* 2001;344(4):276-85.
- Cnossen MH, van der Est MN, Breuning MH, van Asperen CJ, Breslau-Siderius EJ, van der Ploeg AT, et al. Deletions spanning the neurofibromatosis type 1 gene: implications for genotype-phenotype correlations in neurofibromatosis type 1? *Hum Mutat* 1997;9(5):458-64.
- Coates AS, Winer EP, Goldhirsch A, Gelber RD, Gnant M, Piccart-Gebhart M, et al. Tailoring therapies-improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. *Ann Oncol* 2015;26(8):1533-46.
- Cock PJ, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 2009;25(11): 1422-3.
- Cox DR. Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)* 1972. p. 187-220.
- Croonen EA, Yntema HG, van Minkelen R, van den Ouweland AM, van der Burgt I. Patient with a neurofibromatosis type 1 mutation but a clinical diagnosis of Noonan syndrome. *Clin Dysmorphol* 2012;21(4):212-4.
- Crowe FW, Schull WJ, Neel JV. A Clinical, Pathological and Genetic Study of Multiple Neurofibromatosis. Charles C Thomas, 1956.
- Cunha KS, Oliveira NS, Fausto AK, de Souza CC, Gros A, Bandres T, et al. Hybridization

- Capture-Based Next-Generation Sequencing to Evaluate Coding Sequence and Deep Intronic Mutations in the NF1 Gene. *Genes* (Basel) 2016;7(12).
- D'Angelo I, Welti S, Bonneau F, Scheffzek K. A novel bipartite phospholipid-binding module in the neurofibromatosis type 1 protein. *EMBO Rep* 2006;7(2):174-9.
- Da Silva AV, Rodrigues FR, Pureza M, Lopes VG, Cunha KS. Breast cancer and neurofibromatosis type 1: a diagnostic challenge in patients with a high number of neurofibromas. *BMC Cancer* 2015;15:183.
- Danglot G, Régnier V, Fauvet D, Vassal G, Kujas M, Bernheim A. Neurofibromatosis 1 (NF1) mRNAs expressed in the central nervous system are differentially spliced in the 5' part of the gene. *Hum Mol Genet* 1995;4(5):915-20.
- Daston MM, Ratner N. Neurofibromin, a predominantly neuronal GTPase activating protein in the adult, is ubiquitously expressed during development. *Dev Dyn* 1992;195(3):216-26.
- Daston MM, Scrabble H, Nordlund M, Sturbaum AK, Nissen LM, Ratner N. The protein product of the neurofibromatosis type 1 gene is expressed at highest abundance in neurons, Schwann cells, and oligodendrocytes. *Neuron* 1992;8(3):415-28.
- de Blank PMK, Fisher MJ, Liu GT, Gutmann DH, Listernick R, Ferner RE, et al. Optic Pathway Gliomas in Neurofibromatosis Type 1: An Update: Surveillance, Treatment Indications, and Biomarkers of Vision. *J Neuroophthalmol* 2017;37 Suppl 1:S23-S32.
- De Leeneer K, De Schrijver J, Clement L, Baetens M, Lefever S, De Keulenaer S, et al. Practical tools to implement massive parallel pyrosequencing of PCR products in next generation molecular diagnostics. *PLoS One* 2011;6(9):e25531.
- De Raedt T, Brems H, Wolkenstein P, Vidaud D, Pilotti S, Perrone F, et al. Elevated risk for MPNST in NF1 microdeletion patients. *Am J Hum Genet* 2003;72(5):1288-92.
- De Raedt T, Brems H, Lopez-Corra C, Vermeesch JR, Marynen P, Legius E. Genomic organization and evolution of the NF1 microdeletion region. *Genomics* 2004;84(2):346-60.
- DeBella K, Szudek J, Friedman J. Use of the national institutes of health criteria for diagnosis of neurofibromatosis 1 in children. *Pediatrics* 2000;105 (3 Pt 1):608-14.
- DeClue JE, Cohen BD, Lowy DR. Identification and characterization of the neurofibromatosis type 1 protein product. *Proc Natl Acad Sci USA* 1991;88(22):9914-8.
- Denayer E, Devriendt K, de Ravel T, Van Buggenhout G, Smeets E, Francois I, et al. Tumor spectrum in children with Noonan syndrome and SOS1 or RAF1 mutations. *Genes*

- Chromosomes Cancer* 2010; 49(3):242-52.
- Ducatman BS, Scheithauer BW, Piepgras DG, Reiman HM, Ilstrup DM. Malignant peripheral nerve sheath tumors. A clinicopathologic study of 120 cases. *Cancer* 1986;57(10):2006-21.
- Dunzendorfer-Matt T, Mercado EL, Maly K, McCormick F, Scheffzek K. The neurofibromin recruitment factor Spred1 binds to the GAP related domain without affecting Ras inactivation. *Proc Natl Acad Sci U S A*. 2016;113(27):7497-502
- Duong TA, Sbidian E, Valeyrie-Allanore L, Vialette C, Ferkal S, Hadj-Rabia S, et al. Mortality associated with neurofibromatosis 1: a cohort study of 1895 patients in 1980-2006 in France. *Orphanet J Rare Dis* 2011;6:18.
- Easton D, Ponder M, Huson S, Ponder B. An analysis of variation in expression of neurofibromatosis (NF) type 1 (NF1): evidence for modifying genes. *Am J Hum Genet* 1993; 53(2):305-13.
- Elefteriou F, Kolanczyk M, Schindeler A, Viskochil DH, Hock JM, Schorry EK, et al. Skeletal abnormalities in neurofibromatosis type 1: approaches to therapeutic options. *Am J Med Genet A* 2009;149A(10):2327-38.
- Evans DG, Huson SM, Donnai D, Neary W, Blair V, Newton V, et al. A clinical study of type 2 neurofibromatosis. *Q J Med* 1992a;84(304):603-18.
- Evans DG, Huson SM, Donnai D, Neary W, Blair V, Teare D, et al. A genetic study of type 2 neurofibromatosis in the United Kingdom. I. Prevalence, mutation rate, fitness, and confirmation of maternal transmission effect on severity. *J Med Genet* 1992b;29(12):841-6.
- Evans D, Baser M, McGaughran J, Sharif S, Howard E, Moran A. Malignant peripheral nerve sheath tumours in neurofibromatosis 1. *J Med Genet* 2002;39(5):311-4.
- Evans DG, Howard E, Giblin C, Clancy T, Spencer H, Huson SM, et al. Birth incidence and prevalence of tumor-prone syndromes: estimates from a UK family genetic register service. *Am J Med Genet A* 2010;152A(2):327-32.
- Evans DG, O'Hara C, Wilding A, Ingham SL, Howard E, Dawson J, et al. Mortality in neurofibromatosis 1: in North West England: an assessment of actuarial survival in a region of the UK since 1989. *Eur J Hum Genet* 2011;19(11): 1187-91.
- Evans DG. Are we ready for targeted early breast cancer detection strategies in women with NF1 aged 30-49 years? *Am J Med Genet A* 2012;158A(12):3054-5.
- Evans DGR, Salvador H, Chang VY, Erez A, Voss SD, Schneider KW, et al. Cancer and Central Nervous System Tumor Surveillance in Pediatric Neurofibromatosis 1. *Clin Cancer Res* 2017;23(12):e46-e53.

- Fahsold R, Hoffmeyer S, Mischung C, Gille C, Ehlers C, Kücküceylan N, et al. Minor lesion mutational spectrum of the entire NF1 gene does not explain its high mutability but points to a functional domain upstream of the GAP-related domain. *Am J Hum Genet* 2000;66(3):790-818.
- Fazii P, Passamonti M, Lappa G, Di Mascio R, Ballone E. Prevalence of neurofibromatosis type 1 among the young adult population of central Italy. *Arch Dermatol* 1998;134(10):1303-4.
- Ferner R, Gutmann D. International consensus statement on malignant peripheral nerve sheath tumors in neurofibromatosis. *Cancer Res* 2002; 62(5):1573-7.
- Finek J, Holubec L, Topolcan O, Elgrova L, Skalova A, Pecen L. The importance of prognostic factors in premenopausal women with breast cancer. *Anticancer Res* 2007;27(4A):1893-6.
- Friedman J, Birch P. Type 1 neurofibromatosis: a descriptive analysis of the disorder in 1,728 patients. *Am J Med Genet* 1997;70(2):138-43.
- Friedman JM. Epidemiology of neurofibromatosis type 1. *Am J Med Genet* 1999;89(1):1-6.
- Friedman JM. Neurofibromatosis 1. NCBI Bookshelf. A service of the National Library of Medicine, National Institutes of Health. Adam MP, Ardinger HH, Pagon RA, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2017.
- Friedrich R, Hartmann M, Mautner V. Malignant peripheral nerve sheath tumors (MPNST) in NF1-affected children. *Anticancer Res* 2007;27(4A): 1957-60.
- Fuller L, Cox B, Gardner R. Prevalence of von Recklinghausen neurofibromatosis in Dunedin, New Zealand. *Neurofibromatosis* 1989;2(5-6):278-83.
- Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 2013; 6(269):p11.
- Garg S, Lehtonen A, Huson SM, Emsley R, Trump D, Evans DG, et al. Autism and other psychiatric comorbidity in neurofibromatosis type 1: evidence from a population-based study. *Dev Med Child Neurol* 2013;55(2):139-45.
- Garty BZ, Laor A, Danon YL. Neurofibromatosis type 1 in Israel: survey of young adults. *J Med Genet* 1994;31(11): 853-7.
- Gasparini P, D'Agruma L, Pio de Cillis G, Balestrazzi P, Mingarelli R, Zelante L. Scanning the first part of the neurofibromatosis type 1 gene by RNA-SSCP: identification of three novel mutations and of two new polymorphisms. *Hum Genet* 1996;97(4):492-5.
- Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for

- computational biology and bioinformatics. *Genome Biol* 2004;5(10):R80.
- Giltneane JM, Balko JM. Rationale for targeting the Ras/MAPK pathway in triple-negative breast cancer. *Discov Med* 2014;17(95):275-83.
- Gissler M, Haukka J. Finnish health and social welfare registers in epidemiological research. *Norsk Epidemiologi* 2004; 14(1):113-20.
- Gutmann DH, Andersen LB, Cole JL, Swaroop M, Collins FS. An alternatively-spliced mRNA in the carboxy terminus of the neurofibromatosis type 1 (NF1) gene is expressed in muscle. *Hum Mol Genet* 1993;2(7):989-92.
- Gutmann DH, Aylsworth A, Carey JC, Korf B, Marks J, Pyeritz RE, et al. The diagnostic evaluation and multi-disciplinary management of neurofibromatosis 1 and neurofibromatosis 2. *JAMA* 1997;278(1):51-7.
- Gutmann DH, Donahoe J, Brown T, James CD, Perry A. Loss of neurofibromatosis 1 (NF1) gene expression in NF1-associated pilocytic astrocytomas. *Neuropathol Appl Neurobiol* 2000;26(4): 361-7.
- Gutmann DH, Rasmussen SA, Wolkenstein P, MacCollin MM, Guha A, Inskip PD, et al. Gliomas presenting after age 10 in individuals with neurofibromatosis type 1 (NF1). *Neurology* 2002;59(5): 759-61.
- Gutmann DH, James CD, Poyhonen M, Louis DN, Ferner R, Guha A, et al. Molecular analysis of astrocytomas presenting after age 10 in individuals with NF1. *Neurology* 2003;61(10): 1397-400.
- Gutmann DH, Ferner RE, Listernick RH, Korf BR, Wolters PL, Johnson KJ. Neurofibromatosis type 1. *Nat Rev Dis Primers* 2017;3:17004.
- Hadley W. Reshaping Data with the reshape Package. *Journal of Statistical Software* 2007;21 (12):1-20.
- Hakimi MA, Speicher DW, Shiekhhattar R. The motor protein kinesin-1 links neurofibromin and merlin in a common cellular pathway of neurofibromatosis. *J Biol Chem* 2002;277(40):36909-12.
- Han SS, Cooper DN, Upadhyaya MN. Evaluation of denaturing high performance liquid chromatography (DHPLC) for the mutational analysis of the neurofibromatosis type 1 (NF1) gene. *Hum Genet* 2001;109(5):487-97.
- Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000;100(1):57-70.
- Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011;144(5): 646-74.
- Hirata Y, Brems H, Suzuki M, Kanamori M, Okada M, Morita R, et al. Interaction between a Domain of the Negative Regulator of the Ras-ERK Pathway, SPRED1 Protein, and the GTPase-activating Protein-related Domain of Neurofibromin Is Implicated in Legius Syndrome and Neuro-

- fibromatosis Type 1. *J Biol Chem*. 2016;291(7):3124-34.
- Hoadley KA, Yau C, Wolf DM, Cherniack AD, Tamborero D, Ng S, et al. Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. *Cell* 2014;158(4):929-44.
- Hsieh C, Pavia M, Lambe M, Lan SJ, Colditz GA, Ekbohm A, et al. Dual effect of parity on breast cancer risk. *Eur J Cancer* 1994;30A(7):969-73.
- Hsueh YP, Roberts AM, Volta M, Sheng M, Roberts RG. Bipartite interaction between neurofibromatosis type I protein (neurofibromin) and syndecan transmembrane heparan sulfate proteoglycans. *J Neurosci* 2001;21(11):3764-70.
- Huson S, Compston D, Clark P, Harper P. 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(11):704-11.
- Huson S. The Neurofibromatoses: Classification, Clinical Features and Genetic Counselling. In: Kaufmann D, editor. Neurofibromatoses. 16th edition. Karger; 2008. p. 1-20.
- Hölzel M, Huang S, Koster J, Ora I, Lakeman A, Caron H, et al. NF1 is a tumor suppressor in neuroblastoma that determines retinoic acid response and disease outcome. *Cell* 2010;142(2):218-29.
- Imaizumi Y. Mortality of neurofibromatosis in Japan, 1968-1992. *J Dermatol* 1995; 22(3):191-5.
- Ingham S, Huson SM, Moran A, Wylie J, Leahy M, Evans DG. Malignant peripheral nerve sheath tumours in NF1: improved survival in women and in recent years. *Eur J Cancer* 2011;47(18):2723-8.
- Ingordo V, D'Andria G, Mendicini S, Grecucci M, Baglivo A. Segmental neurofibromatosis: is it uncommon or underdiagnosed? *Arch Dermatol* 1995;131(8):959-60.
- Jacks T, Shih TS, Schmitt EM, Bronson RT, Bernards A, Weinberg RA. Tumour predisposition in mice heterozygous for a targeted mutation in Nf1. *Nat Genet* 1994;7(3):353-61.
- Johannessen CM, Reczek EE, James MF, Brems H, Legius E, Cichowski K. The NF1 tumor suppressor critically regulates TSC2 and mTOR. *Proc Natl Acad Sci U S A* 2005;102(24): 8573-8.
- Jouhilahti EM, Peltonen S, Heape AM, Peltonen J. The pathoetiology of neurofibromatosis 1. *Am J Pathol* 2011;178(5): 1932-9.
- Jour G, Scarborough JD, Jones RL, Loggers E, Pollack SM, Pritchard CC, et al. Molecular profiling of soft tissue sarcomas using next-generation sequencing: a pilot study toward precision therapeutics. *Hum Pathol* 2014;45(8):1563-71.
- Kallionpää RA, Uusitalo E, Leppävirta J, Pöyhönen M,

- Peltonen S, Peltonen J. Prevalence of neurofibromatosis type 1 in the Finnish population. *Genet Med* 2017 doi: 10.1038/gim.2017.215.
- Kanada KN, Merin MR, Munden A, Friedlander SF. A prospective study of cutaneous findings in newborns in the United States: correlation with race, ethnicity, and gestational status using updated classification and nomenclature. *J Pediatr* 2012; 161(2):240-5.
- Kehrer-Sawatzki H, Schwickardt T, Assum G, Rocchi M, Krone W. A third neurofibromatosis type 1 (NF1) pseudogene at chromosome 15q11.2. *Hum Genet* 1997;100(5-6):595-600.
- Kehrer-Sawatzki H, Kluwe L, Sandig C, Kohn M, Wimmer K, Krammer U, et al. High frequency of mosaicism among patients with neurofibromatosis type 1 (NF1) with microdeletions caused by somatic recombination of the JAZ1 gene. *Am J Hum Genet* 2004; 75(3):410-23.
- Kehrer-Sawatzki H, Farschtschi S, Mautner VF, Cooper DN. The molecular pathogenesis of schwannomatosis, a paradigm for the co-involvement of multiple tumour suppressor genes in tumorigenesis. *Hum Genet* 2017a;136(2):129-48.
- Kehrer-Sawatzki H, Mautner VF, Cooper DN. Emerging genotype-phenotype relationships in patients with large NF1 deletions. *Hum Genet* 2017b;136(4):349-76.
- Kim A, Stewart DR, Reilly KM, Viskochil D, Miettinen MM, Widemann BC. Malignant Peripheral Nerve Sheath Tumors State of the Science: Leveraging Clinical and Biological Insights into Effective Therapies. *Sarcoma* 2017;2017:7429697.
- King A, Listerick R, Charrow J, Piersall L, Gutmann DH. Optic pathway gliomas in neurofibromatosis type 1: the effect of presenting symptoms on outcome. *Am J Med Genet A* 2003;122A(2):95-9.
- Kluwe L, Tatagiba M, Fünsterer C, Mautner VF. NF1 mutations and clinical spectrum in patients with spinal neurofibromas. *J Med Genet* 2003;40(5):368-71.
- Kluwe L, Siebert R, Gesk S, Friedrich RE, Tinschert S, Kehrer-Sawatzki H, et al. Screening 500 unselected neurofibromatosis 1 patients for deletions of the NF1 gene. *Hum Mutat* 2004;23(2):111-6.
- Kobayashi S, Sugiura H, Ando Y, Shiraki N, Yanagi T, Yamashita H, et al. Reproductive history and breast cancer risk. *Breast Cancer* 2012;19(4):302-8.
- Koczkowska M, Chen Y, Callens T, Gomes A, Sharp A, Johnson S, et al. Genotype-Phenotype Correlation in NF1: Evidence for a More Severe Phenotype Associated with Missense Mutations Affecting NF1 Codons 844-848. *Am J Hum Genet* 2018;102(1):69-87.
- Krab LC, Aarsen FK, de Goede-Bolder A, Catsman-Berrevoets CE, Arts WF, Moll

- HA, et al. Impact of neurofibromatosis type 1 on school performance. *J Child Neurol* 2008;23(9):1002-10.
- Kuorilehto T, Pöyhönen M, Bloigu R, Heikkinen J, Väänänen K, Peltonen J. Decreased bone mineral density and content in neurofibromatosis type 1: lowest local values are located in the load-carrying parts of the body. *Osteoporos Int* 2005; 16(8):928-36.
- Kurian AW. BRCA1 and BRCA2 mutations across race and ethnicity: distribution and clinical implications. *Curr Opin Obstet Gynecol* 2010; 22(1):72-8.
- Lakhani S, Ellis I, Schnitt S, Tan P, van de Vijver M. WHO Classification of Tumors of the Breast. 4th edition, International Agency for Research on Cancer, 2012.
- Lakshmaiah KC, Kumar AN, Purohit S, Viveka BK, Rajan KR, Zameer MAL, et al. Neurofibromatosis type I with breast cancer: not only for women! *Hered Cancer Clin Pract* 2014;12(1):5.
- Laloo F, Evans DG. Familial breast cancer. *Clin Genet* 2012;82 (2):105-14.
- Lammert M, Friedman J, Kluwe L, Mautner V. Prevalence of neurofibromatosis 1 in German children at elementary school enrollment. *Arch Dermatol* 2005;141(1): 71-4.
- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature* 2001;409(6822):860-921.
- Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;9(4):357-9.
- Lara-Corrales I, Moazzami M, García-Romero MT, Pope E, Parkin P, Shugar A, et al. Mosaic Neurofibromatosis Type 1 in Children: A Single-Institution Experience. *J Cutan Med Surg* 2017; 1203475417708163.
- Latteyer S, Tiedje V, König K, Ting S, Heukamp LC, Meder L, et al. Targeted next-generation sequencing for TP53, RAS, BRAF, ALK and NF1 mutations in anaplastic thyroid cancer. *Endocrine* 2016;54(3): 733-41.
- Lau N, Feldkamp MM, Roncari L, Loehr AH, Shannon P, Gutmann DH, et al. Loss of neurofibromin is associated with activation of RAS/MAPK and PI3-K/AKT signaling in a neurofibromatosis 1 astrocytoma. *J Neuropathol Exp Neurol* 2000;59(9):759-67.
- Laycock-van Spyk S, Thomas N, Cooper DN, Upadhyaya M. Neurofibromatosis type 1-associated tumours: their somatic mutational spectrum and pathogenesis. *Hum Genomics* 2011;5(6):623-90.
- Lee W, Teckie S, Wiesner T, Ran L, Prieto Granada CN, Lin M, et al. PRC2 is recurrently inactivated through EED or SUZ12 loss in malignant peripheral nerve sheath tumors. *Nat Genet* 2014; 46(11):1227-32.

- Legius E, Marchuk DA, Hall BK, Andersen LB, Wallace MR, Collins FS, et al. NF1-related locus on chromosome 15. *Genomics* 1992;13(4):1316-8.
- Lehtonen A, Howie E, Trump D, Huson SM. Behaviour in children with neurofibromatosis type 1: cognition, executive function, attention, emotion, and social competence. *Dev Med Child Neurol* 2013;55(2):111-25.
- Leinonen MK, Miettinen J, Heikkinen S, Pitkaniemi J, Malila N. Quality measures of the population-based Finnish Cancer Registry indicate sound data quality for solid malignant tumours. *Eur J Cancer* 2017;77:31-9.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009;25(16):2078-9.
- Lim ET, Würtz P, Havulinna AS, Palta P, Tukiainen T, Rehnström K, et al. Distribution and medical impact of loss-of-function variants in the Finnish founder population. *PLoS Genet* 2014;10(7):e1004494.
- Listernick R, Charrow J, Greenwald MJ, Esterly NB. Optic gliomas in children with neurofibromatosis type 1. *J Pediatr* 1989;114(5):788-92.
- Listernick R, Charrow J, Greenwald M, Mets M. Natural history of optic pathway tumors in children with neurofibromatosis type 1: a longitudinal study. *J Pediatr* 1994;125(1):63-6.
- Listernick R, Louis D, Packer R, Gutmann D. Optic pathway gliomas in children with neurofibromatosis 1: consensus statement from the NF1 Optic Pathway Glioma Task Force. *Ann Neurol* 1997;41(2):143-9.
- Listernick R, Mancini AJ, Charrow J. Segmental neurofibromatosis in childhood. *Am J Med Genet A* 2003;121A(2):132-5.
- Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, et al. Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol* 2012;30(5):434-9.
- Luijten M, Redeker S, Minoshima S, Shimizu N, Westerveld A, Hulsebos TJ. Duplication and transposition of the NF1 pseudogene regions on chromosomes 2, 14, and 22. *Hum Genet* 2001;109(1):109-16.
- MacCollin M, Woodfin W, Kronn D, Short MP. Schwannomatosis: a clinical and pathologic study. *Neurology* 1996;46(4):1072-9.
- MacCollin M, Chiocca E, Evans D, Friedman J, Horvitz R, Jaramillo D, et al. Diagnostic criteria for schwannomatosis. *Neurology* 2005;64(11):1838-45.
- Madanikia SA, Bergner A, Ye X, Blakeley JO. Increased risk of breast cancer in women with NF1. *Am J Med Genet A* 2012;158A(12):3056-60.
- Maertens O, Prenen H, Debiec-Rychter M, Wozniak A, Sciort R, Pauwels P, et al. Molecular pathogenesis of multiple gastrointestinal stromal

- tumors in NF1 patients. *Hum Mol Genet* 2006;15(6):1015-23.
- Maertens O, Johnson B, Hollstein P, Frederick DT, Cooper ZA, Messiaen L, et al. Elucidating distinct roles for NF1 in melanomagenesis. *Cancer Discov* 2013;3(3):338-49.
- Malhotra R, Ratner N. Localization of neurofibromin to keratinocytes and melanocytes in developing rat and human skin. *J Invest Dermatol* 1994;102(5):812-8.
- Mann N, Ma T, Dalton A. Neurofibromatosis type 1 and male breast cancer: emerging risk factor? *J Surg Case Rep* 2017;2017(7):rjw138.
- Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bembem LA, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 2005;437(7057):376-80.
- Maruoka R, Takenouchi T, Torii C, Shimizu A, Misu K, Higasa K, et al. The use of next-generation sequencing in molecular diagnosis of neurofibromatosis type 1: a validation study. *Genet Test Mol Biomarkers* 2014;18(11):722-35.
- Masocco M, Kodra Y, Vichi M, Conti S, Kanieff M, Pace M, et al. Mortality associated with neurofibromatosis type 1: a study based on Italian death certificates (1995-2006). *Orphanet J Rare Dis* 2011;6:11.
- Matsui I, Tanimura M, Kobayashi N, Sawada T, Nagahara N, Akatsuka J. Neurofibromatosis type 1 and childhood cancer. *Cancer* 1993;72(9):2746-54.
- Mautner VF, Lindenau M, Baser ME, Hazim W, Tatagiba M, Haase W, et al. The neuroimaging and clinical spectrum of neurofibromatosis 2. *Neurosurgery* 1996;38(5): 880-5; discussion 5-6.
- Mautner V-F, Boltshauser E. Treatment and Management of Neurofibromatosis 1. In: Kaufmann D, editor. Neurofibromatoses. 16th edition, Karger; 2008. p. 21-31.
- McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2006;15(6):1159-69.
- McGaughan JM, Harris DI, Donnai D, Teare D, MacLeod R, Westerbeek R, et al. A clinical study of type 1 neurofibromatosis in north west England. *J Med Genet* 1999;36(3):197-203.
- McGlynn LM, Kirkegaard T, Edwards J, Tovey S, Cameron D, Twelves C, et al. Ras/Raf-1/MAPK pathway mediates response to tamoxifen but not chemotherapy in breast cancer patients. *Clin Cancer Res* 2009;15(4):1487-95.
- McKeever K, Shepherd CW, Crawford H, Morrison PJ. An epidemiological, clinical and genetic survey of neurofibromatosis type 1 in children under sixteen years of age. *Ulster Med J* 2008;77(3):160-3.

- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;20(9):1297-303.
- Melchor L, Benítez J. The complex genetic landscape of familial breast cancer. *Hum Genet* 2013;132(8):845-63.
- Mendes-Pereira AM, Sims D, Dexter T, Fenwick K, Assiotis I, Kozarewa I, et al. Genome-wide functional screen identifies a compendium of genes affecting sensitivity to tamoxifen. *Proc Natl Acad Sci U S A* 2012;109(8):2730-5.
- Messiaen L, Callens T, Mortier G, Beysen D, Vandenbroucke I, Van Roy N, et al. Exhaustive mutation analysis of the NF1 gene allows identification of 95% of mutations and reveals a high frequency of unusual splicing defects. *Hum Mutat* 2000;15(6):541-55.
- Messiaen L, Riccardi V, Peltonen J, Maertens O, Callens T, Karvonen SL, et al. Independent NF1 mutations in two large families with spinal neurofibromatosis. *J Med Genet* 2003;40(2):122-6.
- Messiaen LM, Wimmer K. NF1 Mutational Spectrum. In: Kaufmann D., editor. Neurofibromatoses. Monographs in Human Genetics. 16th edition, Karger; 2008. p. 63-77.
- Messiaen L, Yao S, Brems H, Callens T, Sathienkijanchai A, Denayer E, et al. Clinical and mutational spectrum of neurofibromatosis type 1-like syndrome. *JAMA* 2009; 302(19):2111-8
- Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* 1994; 266(5182):66-71.
- Michailidou K, Lindström S, Dennis J, Beesley J, Hui S, Kar S, et al. Association analysis identifies 65 new breast cancer risk loci. *Nature* 2017; 551(7678):92-4.
- Milne RL, Kuchenbaecker KB, Michailidou K, Beesley J, Kar S, Lindström S, et al. Identification of ten variants associated with risk of estrogen-receptor-negative breast cancer. *Nat Genet* 2017; 49(12):1767-78.
- Mucci LA, Hjelmborg JB, Harris JR, Czene K, Havelick DJ, Scheike T, et al. Familial Risk and Heritability of Cancer Among Twins in Nordic Countries. *JAMA* 2016;315 (1):68-76.
- Müllauer L. Milestones in pathology-from histology to molecular biology. *Memo* 2017;10(1): 42-5.
- Möller S, Mucci LA, Harris JR, Scheike T, Holst K, Halekoh U, Adami HO, et al. The Heritability of Breast Cancer among Women in the Nordic Twin Study of Cancer. *Cancer Epidemiol Biomarkers Prev* 2016;25(1):145-50.
- Network CGAR. Comprehensive molecular profiling of lung

- adenocarcinoma. *Nature* 2014;511(7511):543-50.
- Nguyen R, Kluwe L, Fuensterer C, Kentsch M, Friedrich RE, Mautner VF. Plexiform neurofibromas in children with neurofibromatosis type 1: frequency and associated clinical deficits. *J Pediatr* 2011;159(4):652-5.e2.
- Niemeyer CM, Arico M, Basso G, Biondi A, Cantu Rajnoldi A, Creutzig U, et al. Chronic myelomonocytic leukemia in childhood: a retrospective analysis of 110 cases. European Working Group on Myelodysplastic Syndromes in Childhood (EWOG-MDS). *Blood* 1997;89(10):3534-43.
- Nordlund M, Gu X, Shipley MT, Ratner N. Neurofibromin is enriched in the endoplasmic reticulum of CNS neurons. *J Neurosci* 1993;13(4):1588-600.
- Osborn MJ, Upadhyaya M. Evaluation of the protein truncation test and mutation detection in the NF1 gene: mutational analysis of 15 known and 40 unknown mutations. *Hum Genet* 1999; 105(4):327-32.
- Page PZ, Page GP, Ecosse E, Korf BR, Lepage A, Wolkenstein P. Impact of neurofibromatosis 1 on Quality of Life: a cross-sectional study of 176 American cases. *Am J Med Genet A* 2006;140(18):1893-8.
- Parry DM, Eldridge R, Kaiser-Kupfer MI, Bouzas EA, Pikus A, Patronas N. Neurofibromatosis 2 (NF2): clinical characteristics of 63 affected individuals and clinical evidence for heterogeneity. *Am J Med Genet* 1994;52(4):450-61.
- Pasmant E, Parfait B, Luscan A, Goussard P, Briand-Suleau A, Laurendeau I, et al. Neurofibromatosis type 1 molecular diagnosis: what can NGS do for you when you have a large gene with loss of function mutations? *Eur J Hum Genet* 2015;23(5):596-601.
- Peltonen S, Pöyhönen M. Clinical Diagnosis and Atypical Forms of NF1. In: Upadhyaya M, Cooper DN, editors. Neurofibromatosis type 1 Molecular and Cellular Biology. Springer; 2012. p. 17-30.
- Perou CM, Sørli T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature* 2000;406(6797):747-52.
- Perry A, Roth KA, Banerjee R, Fuller CE, Gutmann DH. NF1 deletions in S-100 protein-positive and negative cells of sporadic and neurofibromatosis 1 (NF1)-associated plexiform neurofibromas and malignant peripheral nerve sheath tumors. *Am J Pathol* 2001; 159(1):57-61.
- Petek E, Jenne DE, Smolle J, Binder B, Lasinger W, Windpassinger C, et al. Mitotic recombination mediated by the JJAZF1 (KIAA0160) gene causing somatic mosaicism and a new type of constitutional NF1 microdeletion in two children of a mosaic female with only

- few manifestations. *J Med Genet* 2003;40(7):520-5.
- Pettersson A, Graff RE, Ursin G, Santos Silva ID, McCormack V, Baglietto L, et al. Mammographic density phenotypes and risk of breast cancer: a meta-analysis. *J Natl Cancer Inst* 2014;106(5).
- Philpott C, Tovell H, Frayling IM, Cooper DN, Upadhyaya M. The NF1 somatic mutational landscape in sporadic human cancers. *Hum Genomics* 2017; 11(1):13.
- Pinna V, Lanari V, Daniele P, Consoli F, Agolini E, Margiotti K, et al. p.Arg1809Cys substitution in neurofibromin is associated with a distinctive NF1 phenotype without neurofibromas. *Eur J Hum Genet* 2015;23(8):1068-71.
- Plotkin SR, Blakeley JO, Evans DG, Hanemann CO, Hulsebos TJ, Hunter-Schaedle K, et al. Update from the 2011 International Schwannomatosis Workshop: From genetics to diagnostic criteria. *Am J Med Genet A* 2013;161A(3):405-16.
- Poyhonen M, Leisti E, Kytölä S, Leisti J. Hereditary spinal neurofibromatosis: a rare form of NF1? *J Med Genet* 1997; 34(3):184-7.
- Poyhonen M. A clinical assessment of neurofibromatosis type 1 (NF1) and segmental NF in Northern Finland. *J Med Genet* 2000;37(12):E43.
- Poyhonen M, Kytölä S, Leisti J. Epidemiology of neurofibromatosis type 1 (NF1) in northern Finland. *J Med Genet* 2000;37(8):632-6.
- Purandare SM, Huntsman Breidenbach H, Li Y, Zhu XL, Sawada S, Neil SM, et al. Identification of neurofibromatosis 1 (NF1) homologous loci by direct sequencing, fluorescence in situ hybridization, and PCR amplification of somatic cell hybrids. *Genomics* 1995; 30(3):476-85.
- Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med* 2007; 26(11):2389-430.
- Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010;26(6): 841-2.
- Rad E, Tee AR. Neurofibromatosis type 1: Fundamental insights into cell signalling and cancer. *Semin Cell Dev Biol* 2016;52:39-46.
- Rasmussen SA, Yang Q, Friedman JM. Mortality in neurofibromatosis 1: an analysis using U.S. death certificates. *Am J Hum Genet* 2001;68(5):1110-8.
- Riccardi VM, Lewis RA. Penetrance of von Recklinghausen neurofibromatosis: a distinction between predecessors and descendants. *Am J Hum Genet* 1988;42(2): 284-9.
- Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol* 2011;29(1):24-6.
- Robinson PN, Buske A, Neumann R, Tinschert S, Nürnberg P.

- Recurrent 2-bp deletion in exon 10c of the NF1 gene in two cases of von Recklinghausen neurofibromatosis. *Hum Mutat* 1996; 7(1):85-8.
- Rojnueangnit K, Xie J, Gomes A, Sharp A, Callens T, Chen Y, et al. High Incidence of Noonan Syndrome Features Including Short Stature and Pulmonic Stenosis in Patients carrying NF1 Missense Mutations Affecting p.Arg1809: Genotype-Phenotype Correlation. *Hum Mutat* 2015; 36(11):1052-63.
- Rosenfeld A, Listerick R, Charrow J, Goldman S. Neurofibromatosis type 1 and high-grade tumors of the central nervous system. *Childs Nerv Syst* 2010;26(5):663-7.
- Roudebush M, Slabe T, Sundaram V, Hoppel CL, Golubic M, Stacey DW. Neurofibromin colocalizes with mitochondria in cultured cells. *Exp Cell Res* 1997;236(1):161-72.
- Rouleau GA, Merel P, Lutchman M, Sanson M, Zucman J, Marineau C, et al. Alteration in a new gene encoding a putative membrane-organizing protein causes neuro-fibromatosis type 2. *Nature* 1993;363(6429):515-21.
- Ruggieri M, Huson SM. The clinical and diagnostic implications of mosaicism in the neuro-fibromatoses. *Neurology* 2001;56(11):1433-43.
- Ruggieri M, Polizzi A, Spalice A, Salpietro V, Caltabiano R, D'Orazi V, et al. The natural history of spinal neurofibromatosis: a critical review of clinical and genetic features. *Clin Genet* 2015; 87(5):401-10.
- Samuelsson B, Axelsson R. Neurofibromatosis. A clinical and genetic study of 96 cases in Gothenburg, Sweden. *Acta Derm Venereol Suppl (Stockh)* 1981;95:67-71.
- Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 1977;74(12):5463-7.
- Schneider M, Obringer AC, Zackai E, Meadows AT. Childhood neurofibromatosis: risk factors for malignant disease. *Cancer Genet Cytogenet* 1986;21(4): 347-54.
- Seizinger BR, Rouleau GA, Ozelius LJ, Lane AH, Faryniarz AG, Chao MV, et al. Genetic linkage of von Recklinghausen neurofibromatosis to the nerve growth factor receptor gene. *Cell* 1987;49(5):589-94.
- Seminog OO, Goldacre MJ. Risk of benign tumours of nervous system, and of malignant neoplasms, in people with neurofibromatosis: population-based record-linkage study. *Br J Cancer* 2013;108(1):193-8.
- Seminog OO, Goldacre MJ. Age-specific risk of breast cancer in women with neurofibromatosis type 1. *Br J Cancer* 2015;112(9):1546-8.
- Sergeyev AS. On the mutation rate of neurofibromatosis. *Human-genetik* 1975;28(2):129-38.
- Shah KN. The diagnostic and clinical significance of café-au-lait

- macules. *Pediatr Clin North Am* 2010;57(5):1131-53.
- Sharif S, Ferner R, Birch JM, Gillespie JE, Gattamaneni HR, Baser ME, et al. Second primary tumors in neurofibromatosis 1 patients treated for optic glioma: substantial risks after radiotherapy. *J Clin Oncol* 2006;24(16):2570-5.
- Sharif S, Moran A, Huson S, Iddenden R, Shenton A, Howard E, et al. Women with neurofibromatosis 1 are at a moderately increased risk of developing breast cancer and should be considered for early screening. *J Med Genet* 2007;44(8):481-4.
- Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 2001;29(1):308-11.
- Shilyansky C, Lee YS, Silva AJ. Molecular and cellular mechanisms of learning disabilities: a focus on NF1. *Annu Rev Neurosci* 2010;33:221-43.
- Side L, Taylor B, Cayouette M, Conner E, Thompson P, Luce M, et al. Homozygous inactivation of the NF1 gene in bone marrow cells from children with neurofibromatosis type 1 and malignant myeloid disorders. *N Engl J Med* 1997;336(24):1713-20.
- Sievert AJ, Fisher MJ. Pediatric low-grade gliomas. *J Child Neurol* 2009;24(11):1397-408.
- Singletary SE. Rating the risk factors for breast cancer. *Ann Surg* 2003;237(4):474-82.
- Soininen L, Pokhrel A, Dyba T, Pukkala E, Hakulinen T. Survival of Sami cancer patients. *Int J Circumpolar Health* 2012;71:18959.
- Spencer E, Davis J, Mikhail F, Fu C, Vijzelaar R, Zackai E, et al. Identification of SPRED1 deletions using RT-PCR, multiplex ligation-dependent probe amplification and quantitative PCR. *Am J Med Genet* 2011;155:1352-9.
- Stiller CA, Chessells JM, Fitchett M. Neurofibromatosis and childhood leukaemia/lymphoma: a population-based UKCCSG study. *Br J Cancer* 1994;70(5):969-72.
- Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature* 2009;458(7239):719-24.
- Stumpf D, Alksne J, Annegers J, Brown S, Conneally P, Housman D, et al. Neurofibromatosis. Conference statement. National Institutes of Health Consensus Development Conference. *Archives of Neurology* 1988;45(5):575-8.
- Stowe IB, Mercado EL, Stowe TR, Bell EL, Oses-Prieto JA, Hernández H., et al. A shared molecular mechanism underlies the human rasopathies Legius syndrome and neurofibromatosis-1. *Genes Dev* 2012;26(13):1421-6.
- Sun YS, Zhao Z, Yang ZN, Xu F, Lu HJ, Zhu ZY, et al. Risk Factors and Preventions of Breast Cancer. *Int J Biol Sci* 2017;13(11):1387-1397.

- Suzuki Y, Suzuki H, Kayama T, Yoshimoto T, Shibahara S. Brain tumors predominantly express the neurofibromatosis type 1 gene transcripts containing the 63 base insert in the region coding for GTPase activating protein-related domain. *Biochem Biophys Res Commun* 1991;181(3):955-61.
- Suzuki H, Ozawa N, Taga C, Kano T, Hattori M, Sakaki Y. Genomic analysis of a NF1-related pseudogene on human chromosome 21. *Gene* 1994; 147(2):277-80.
- Szudek J, Joe H, Friedman JM. Analysis of intrafamilial phenotypic variation in neurofibromatosis 1 (NF1). *Genet Epidemiol* 2002;23(2): 150-64.
- Sørensen SA, Mulvihill JJ, Nielsen A. Long-term follow-up of von Recklinghausen neurofibromatosis. Survival and malignant neoplasms. *N Engl J Med* 1986;314(16): 1010-5.
- Tabori U, Hansford JR, Achatz MI, Kratz CP, Plon SE, Frebourg T, et al. Clinical Management and Tumor Surveillance Recommendations of Inherited Mismatch Repair Deficiency in Childhood. *Clin Cancer Res* 2017;23(11):e32-e7.
- Teppo L, Pukkala E, Lehtonen M. Data quality and quality control of a population-based cancer registry. Experience in Finland. *Acta Oncol* 1994; 33(4):365-9.
- Thiagalingam S, Flaherty M, Billson F, North K. Neurofibromatosis type 1 and optic pathway gliomas: follow-up of 54 patients. *Ophthalmology* 2004;111(3):568-77.
- Tinschert S, Naumann I, Stegmann E, Buske A, Kaufmann D, Thiel G, et al. Segmental neurofibromatosis is caused by somatic mutation of the neurofibromatosis type 1 (NF1) gene. *Eur J Hum Genet* 2000;8(6):455-9.
- Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)* 2015;19(1A):A68-77.
- Tonsgard JH, Yelavarthi KK, Cushner S, Short MP, Lindgren V. Do NF1 gene deletions result in a characteristic phenotype? *Am J Med Genet* 1997;73(1):80-6.
- Tretli S. Height and weight in relation to breast cancer morbidity and mortality. A prospective study of 570,000 women in Norway. *Int J Cancer* 1989;44(1):23-30.
- Trevisson E, Cassina M, Opocher E, Vicenzi V, Lucchetta M, Parrozzani R, et al. Natural history of optic pathway gliomas in a cohort of unselected patients affected by Neurofibromatosis 1. *J Neuro-oncol* 2017.
- Upadhyaya M, Ruggieri M, Maynard J, Osborn M, Hartog C, Mudd S, et al. Gross deletions of the neurofibromatosis type 1 (NF1) gene are predominantly of maternal origin and commonly associated with a learning disability, dysmorphic features and

- developmental delay. *Hum Genet* 1998;102(5):591-7.
- Upadhyaya M, Huson SM, Davies M, Thomas N, Chuzhanova N, Giovannini S, 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. *Am J Hum Genet* 2007;80(1):140-51.
- Upadhyaya M, Spurlock G, Monem B, Thomas N, Friedrich RE, Kluwe L, et al. Germline and somatic NF1 gene mutations in plexiform neurofibromas. *Hum Mutat* 2008;29(8):E103-11.
- Upadhyaya M. NF1 gene structure and NF1 Genotype/Phenotype Correlations. In: Kaufmann D, editor. *Neurofibromatoses*. 16th edition, Karger; 2008. p. 46-62.
- Upadhyaya M. Neurofibromatosis type 1: diagnosis and recent advances. *Expert Opin Med Diagn* 2010;4(4):307-22.
- van der Vaart T, Rietman AB, Plasschaert E, Legius E, Elgersma Y, Moll HA, et al. Behavioral and cognitive outcomes for clinical trials in children with neurofibromatosis type 1. *Neurology* 2016;86(2):154-60.
- van Minkelen R, van Bever Y, Kromosoeto JN, Withagen-Hermans CJ, Nieuwlaat A, Halley DJ, et al. A clinical and genetic overview of 18 years neurofibromatosis type 1 molecular diagnostics in the Netherlands. *Clin Genet* 2014; 85(4):318-27.
- Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. *Science* 2001; 291(5507):1304-51.
- Viskochil D, Buchberg AM, Xu G, Cawthon RM, Stevens J, Wolff RK, et al. Deletions and a translocation interrupt a cloned gene at the neurofibromatosis type 1 locus. *Cell* 1990;62(1):187-92.
- Viskochil D. Genetics of neurofibromatosis 1 and the NF1 gene. *J Child Neurol* 2002;17(8):562-70; discussion 71-2, 646-51.
- Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer genome landscapes. *Science* 2013;339(6127):1546-58.
- Walker L, Thompson D, Easton D, Ponder B, Ponder M, Frayling I, et al. A prospective study of neurofibromatosis type 1 cancer incidence in the UK. *Br J Cancer* 2006;95(2):233-8.
- Wallace MD, Pfefferle AD, Shen L, McNairn AJ, Cerami EG, Fallon BL, et al. Comparative oncogenomics implicates the neurofibromin 1 gene (NF1) as a breast cancer driver. *Genetics* 2012;192(2):385-96.
- Wallace MR, Marchuk DA, Andersen LB, Letcher R, Odeh HM, Saulino AM, et al. Type 1 neurofibromatosis gene: identification of a large transcript disrupted in three NF1 patients. *Science* 1990; 249(4965):181-6.
- Walther MM, Herring J, Enquist E, Keiser HR, Linehan WM. von

- Recklinghausen's disease and pheochromocytomas. *J Urol* 1999;162(5):1582-6.
- Wang X, Levin AM, Smolinski SE, Vigneau FD, Levin NK, Tainsky MA. Breast cancer and other neoplasms in women with neurofibromatosis type 1: a retrospective review of cases in the Detroit metropolitan area. *Am J Med Genet A* 2012; 158A(12):3061-4.
- Welti S, Fraterman S, D'Angelo I, Wilm M, Scheffzek K. The sec14 homology module of neurofibromin binds cellular glycerophospholipids: mass spectrometry and structure of a lipid complex. *J Mol Biol* 2007;366(2):551-62.
- Wiesner T, Kiuru M, Scott SN, Arcila M, Halpern AC, Hollmann T, et al. NF1 Mutations Are Common in Desmoplastic Melanoma. *Am J Surg Pathol* 2015;39(10):1357-62.
- Wimmer K, Yao S, Claes K, Kehrer-Sawatzki H, Tinschert S, De Raedt T, et al. Spectrum of single- and multiexon NF1 copy number changes in a cohort of 1,100 unselected NF1 patients. *Genes Chromosomes Cancer* 2006; 45(3):265-76.
- Wimmer K, Kratz CP, Vasen HF, Caron O, Colas C, Entz-Werle N, et al. Diagnostic criteria for constitutional mismatch repair deficiency syndrome: suggestions of the European consortium 'care for CMMRD' (C4CMMRD). *J Med Genet* 2014;51(6):355-65.
- Wood LD, Parsons DW, Jones S, Lin J, Sjöblom T, Leary RJ, et al. The genomic landscapes of human breast and colorectal cancers. *Science* 2007;318 (5853):1108-13.
- Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, et al. Identification of the breast cancer susceptibility gene BRCA2. *Nature* 1995;378 (6559):789-92.
- Wright KL, Adams JR, Liu JC, Loch AJ, Wong RG, Jo CE, et al. Ras Signaling Is a Key Determinant for Metastatic Dissemination and Poor Survival of Luminal Breast Cancer Patients. *Cancer Res* 2015;75(22):4960-72.
- Xu GF, O'Connell P, Viskochil D, Cawthon R, Robertson M, Culver M, et al. The neurofibromatosis type 1 gene encodes a protein related to GAP. *Cell* 1990;62(3):599-608.
- Yang SX, Polley E, Lipkowitz S. New insights on PI3K/AKT pathway alterations and clinical outcomes in breast cancer. *Cancer Treat Rev* 2016;45:87-96.
- Zenker M. Clinical manifestations of mutations in RAS and related intracellular signal transduction factors. *Curr Opin Pediatr* 2011;23(4):443-51.
- Zhang J, Tong H, Fu X, Zhang Y, Liu J, Cheng R, et al. Molecular Characterization of NF1 and Neurofibromatosis Type 1 Genotype-Phenotype Correlations in a Chinese Population. *Sci Rep* 2015;5: 11291.
- Zvulunov A, Weitz R, Metzker A. Neurofibromatosis type 1 in childhood: evaluation of clinical and epidemiologic

- features as predictive factors for severity. *Clin Pediatr (Phila)* 1998;37(5):295-9.
- Zöller M, Rembeck B, Akesson HO, Angervall L. Life expectancy, mortality and prognostic factors in neurofibromatosis type 1. A twelve-year follow-up of an epidemiological study in Göteborg, Sweden. *Acta Derm Venereol* 1995;75(2):136-40.
- Zöller M, Rembeck B, Odén A, Samuelsson M, Angervall L. Malignant and benign tumors in patients with neurofibromatosis type 1 in a defined Swedish population. *Cancer* 1997;79(11):2125-31.

Annales Universitatis Turkuensis



Turun yliopisto
University of Turku

ISBN 978-951-29-7155-8 (PRINT)
ISBN 978-951-29-7156-5 (PDF)
ISSN 0355-9483 (PRINT) | ISSN 2343-3213 (PDF)