

## ARTICLE

# Neurofibromatosis type 1 molecular diagnosis: what can NGS do for you when you have a large gene with loss of function mutations?

Eric Pasmant<sup>\*,1,2,4</sup>, Béatrice Parfait<sup>1,2,4</sup>, Armelle Luscan<sup>1,2</sup>, Philippe Goussard<sup>1</sup>, Audrey Briand-Suleau<sup>1,2</sup>, Ingrid Laurendeau<sup>2</sup>, Corinne Fouveaut<sup>1</sup>, Chrystel Leroy<sup>1</sup>, Annelore Montadert<sup>1,2</sup>, Pierre Wolkenstein<sup>3</sup>, Michel Vidaud<sup>1,2</sup> and Dominique Vidaud<sup>1,2</sup>

Molecular diagnosis of neurofibromatosis type 1 (NF1) is challenging owing to the large size of the tumour suppressor gene *NF1*, and the lack of mutation hotspots. A somatic alteration of the wild-type *NF1* allele is observed in NF1-associated tumours. Genetic heterogeneity in *NF1* was confirmed in patients with *SPRED1* mutations. Here, we present a targeted next-generation sequencing (NGS) of *NF1* and *SPRED1* using a multiplex PCR approach (230 amplicons of ~150 bp) on a PGM sequencer. The chip capacity allowed mixing 48 bar-coded samples in a 4-day workflow. We validated the NGS approach by retrospectively testing 30 *NF1*-mutated samples, and then prospectively analysed 279 patients in routine diagnosis. On average, 98.5% of all targeted bases were covered by at least 20X and 96% by at least 100X. An *NF1* or *SPRED1* alteration was found in 246/279 (88%) and 10/279 (4%) patients, respectively. Genotyping throughput was increased over 10 times, as compared with Sanger, with ~90€ for consumables per sample. Interestingly, our targeted NGS approach also provided quantitative information based on sequencing depth allowing identification of multiexons deletion or duplication. We then addressed the *NF1* somatic mutation detection sensitivity in mosaic NF1 patients and tumours.

*European Journal of Human Genetics* (2015) 23, 596–601; doi:10.1038/ejhg.2014.145; published online 30 July 2014

## INTRODUCTION

Neurofibromatosis type 1 (NF1, MIM no. 162200) is a common autosomal dominant genetic disorder affecting ~1/3000 individuals irrespective of gender or ethnicity.<sup>1</sup> As many as ~50% of NF1 patients present as sporadic patients. The clinical diagnosis of NF1 is based on the clinical diagnostic criteria outlined in the National Institutes of Health (NIH) consensus development conference in 1987.<sup>2</sup> The commonest NF1-associated tumours are benign peripheral nerve sheath tumours called neurofibromas that may be either dermal or plexiform (PNFs). PNFs can undergo malignant transformation to malignant peripheral nerve sheath tumours (MPNSTs) in ~10% of patients.<sup>3</sup> MPNSTs often metastasize in an early phase of the disease and are associated with a poor prognosis.

NF1 is caused by dominant loss-of-function mutations of the tumour suppressor *NF1* (*Neurofibromin 1*; MIM no. 613113), located at 17q11.2. *NF1* comprises 57 constitutive exons, plus three alternative spliced exons. There are at least seven partial duplications of *NF1* on other chromosomes, showing evidence of transcription. The 8454 nucleotides open reading frame (NM\_000267.3) of *NF1* encodes neurofibromin, a 2818 amino-acid peptide that appears to be expressed ubiquitously. Neurofibromin is a tumour suppressor protein, reflecting its role as a key negative regulator of the cellular RAS-MAPK (mitogen-activated protein kinases) signalling pathway. Any *NF1* loss-of-function mutation is expected to increase cellular levels of the active

GTP-bound form of RAS, leading to uncontrolled cell growth and potentially tumorigenesis.

For as yet, unknown reasons, the mutation rate at the *NF1* locus is one of the highest reported in any human disorder.<sup>4</sup> This observation is reflected in the finding that almost 50% of all NF1 patients exhibit a *de novo NF1* mutation. A huge number of different pathogenic *NF1* mutations have been reported (Leiden Open Variation Database, LOVD: [www.lovd.nl/NF1](http://www.lovd.nl/NF1)).<sup>5–11</sup> Over 1300 *NF1* mutations have been reported by The Human Gene Mutation Database (HGMD, Institute of Medical Genetics, Cardiff, <http://www.hgmd.org/>). Among them, 5–10% are large 17q11.2 deletions encompassing the entire *NF1* locus and neighbouring genes.<sup>12</sup> These large *NF1* locus deletions have been associated with a more severe phenotype including an elevated risk for MPNSTs.<sup>13–14</sup> For patients with intragenic *NF1* mutations (more than 90% of all NF1 cases), no clear-cut allele–phenotype correlations have been established so far<sup>15–18</sup> with the exception of a 3-bp inframe deletion (c.2970\_2972delAAT) in exon 22 of the *NF1* gene that has been associated with the absence of cutaneous neurofibromas.<sup>19</sup> Screen for *NF1* gene lesions routinely identify up to 95% of pathological mutations assumed to be present in patients presenting with typical NF1. This high sensitivity can be obtained with a multistep approach including a cDNA and/or DNA Sanger sequencing and copy number alteration (CNA) study.<sup>8–11</sup>

<sup>1</sup>Service de Biochimie et Génétique Moléculaire, Hôpital Cochin, Assistance Publique-Hôpitaux de Paris, Paris, France; <sup>2</sup>EA7331, Université Paris Descartes, Sorbonne Paris Cité, Faculté des Sciences Pharmaceutiques et Biologiques, Paris, France; <sup>3</sup>Département de Dermatologie, Centre de référence des neurofibromatoses, Hôpital Henri-Mondor, Université Paris Est Créteil (UPEC), Créteil, France

\*Correspondence: Dr E Pasmant, Service de Biochimie et Génétique Moléculaire, Hôpital Cochin, Assistance Publique-Hôpitaux de Paris, Bâtiment Jean DAUSSET—3ème étage, 27 rue du Faubourg Saint Jacques, Paris 75006, France. Tel: +33 1 01 58 41 16 01; Fax: +33 1 58 41 15 80; E-mail: eric.pasmant@gmail.com

<sup>4</sup>These authors contributed equally to this work.

Received 21 February 2014; revised 30 May 2014; accepted 25 June 2014; published online 30 July 2014

As *NF1* is a tumour suppressor gene, it is expected that tumorigenesis will occur when a second somatic mutation disrupts the remaining functional copy of this gene. In keeping with this expectation and consistent with Knudson's two hit hypothesis, a somatic loss of the second wild-type *NF1* allele can be found in *NF1*-associated tumours, including neurofibromas and MPNSTs.<sup>20</sup>

Genetic heterogeneity has been found in *NF1* by identification of a novel disorder, the Legius syndrome, caused by mutations in the *SPRED1* gene.<sup>21–23</sup> *SPRED1* encodes a negative regulator of the RAS-MAPK pathway. It is located on 15q13.2 and spans 104.4 kb of genomic sequence. It contains seven exons and the transcript is 7255 bp in length (NM\_152594.2) with an ORF coding a 444 amino-acid protein. A large spectrum of mutations over the entire gene has been described (LOVD: [www.lovd.nl/SPRED1](http://www.lovd.nl/SPRED1)).<sup>23</sup> However, *SPRED1* mutations are rare compared with *NF1* mutations.

Molecular diagnosis of *NF1* is challenging owing to the large size of the gene, the existence of multiple highly identical pseudogenes, the lack of mutational hotspots, and the complex mutational spectrum.<sup>11,24</sup> Targeted next-generation sequencing (NGS) can now be applied to the accurate and fast identification of *NF1* and *SPRED1* variants and CNAs, to establish the unambiguous diagnosis of *NF1* or Legius syndrome. In the present study, we validated a NGS approach for *NF1* and *SPRED1* variant screening and we then prospectively analysed 279 patients in *NF1* routine diagnosis. We also addressed the *NF1* somatic variant detection in mosaic *NF1* patients and tumours.

## MATERIALS AND METHODS

### Patients and samples

Thirty *NF1* patient blood DNA samples for the validation study were retrospectively selected from the French *NF1* database (French clinical research programme PHRC *NF1*, coordinator: Professor Pierre Wolkenstein, Henri Mondor Hospital, Créteil, France). The French *NF1* database constituted through this program included a collection of 565 families, consisting of 1697 individuals among whom 1083 fulfilled NIH diagnostic criteria for *NF1*.<sup>11,17,25</sup> The study was approved by the local ethical committee and all participants gave their written informed consents.

A prospective study panel of 279 blood DNA samples from *NF1* clinically diagnosed index cases was analysed from routine genetic counselling. All blood samples drawn for DNA extraction were collected after informed consent.

DNA samples of 15 MPNSTs were retrospectively analysed. MPNSTs were obtained during surgery from 15 *NF1* patients of the French *NF1* cohort at Henri Mondor hospital (Créteil, France) and were snap-frozen. MPNSTs all arose from plexiform neurofibromas and showed very weak S100 immunostaining (data not shown).

### DNA extraction

DNA was isolated from peripheral blood mononuclear cells or MPNSTs using standard proteinase K digestion followed by phenol–chloroform extraction. DNA concentrations were quantified by using a Quant-iT dsDNA HS assay kit and a Qubit 2.0 Fluorometer (Life Technologies, Saint-Aubin, France).

### *NF1* and *SPRED1* NGS sequencing

Experiments were performed on the NGS facility of the Cochin hospital, Paris (Assistance Publique-Hôpitaux de Paris, France). The custom primer panel targeting the *NF1* and *SPRED1* genes was designed using the AmpliSeq Designer (reference IAD35072\_003, Life Technologies). The targeted region included the entire *NF1*- and *SPRED1*-coding exons, their intron boundaries (25 bp), and the 5' and 3' untranslated regions (UTRs). The targeted region (20 595 bp) was amplified by 230 amplicons (length between 125 and 175 bp) distributed in two primer pools: 20 ng of genomic DNA are amplified to generate the library using the Ion AmpliSeq Library Kit 2.0 (Life Technologies). Theoretic coverage was 90.8%. The entire *SPRED1*-coding sequence was theoretically covered (100% of the coding sequence) and 85.5% of the *SPRED1*

NM\_152594.2 reference sequence (1049 bp were missed in the 3' UTR) were covered. The *NF1* gene (reference sequence NM\_001042492.2) was theoretically covered at 93.6%: 110/9232 bp (98.8%) of the *NF1*-coding sequences were theoretically missed and 303 bp were missed in the *NF1* 3' UTR. NGS library preparation for sequencing of the two genes was performed using the Ion AmpliSeq Library Kit 2.0 (Life Technologies), according to the manufacturer's instructions (Ion AmpliSeq Library Preparation, Publication Part Number MAN0006735, Revision 5.0, July 2013, Life Technologies). The amplified libraries were purified using Agencourt AMPure XP beads (Beckman Coulter, Brea, CA, USA). Before library pooling and sequencing sample preparation, amplified libraries were validated and quantified using the 2100 Bioanalyzer microfluidic platform (Agilent Technologies, Santa Clara, CA, USA). Emulsion PCR was performed using the Ion OneTouch Instrument (Life Technologies). Enrichment of the template-positive Ion OneTouch 200 ion sphere particles (ISPs, containing clonally amplified DNA) PCR was performed using the Ion OneTouch ES (Life Technologies), according to the manufacturer's procedures. An ISP quality control was then performed using a QubitR 2.0 Fluorometer. The ISP Quality Control assay on the Qubit 2.0 Fluorometer labelled the ISPs 200 with two different fluorophores: Alexa Fluor 488 and Alexa Fluor 647. The probe-labeled Alexa Fluor 488 annealed to all of the ISPs present, whereas the probe-labeled Alexa Fluor 647 to only the ISPs with extended templates. The ratio of the Alexa Fluor 488 fluorescence (all ISPs present) to the Alexa Fluor 647 fluorescence (templated ISPs) yielded the % templated ISPs. The template-positive ISPs were loaded on Ion 316 chips and sequenced with an Ion Personal Genome Machine (PGM) System (Life Technologies). The chip capacity allowed mixing 48-bar-coded samples. A typical 48 sample run workflow was achieved in 4 days from the DNA dosage to the data collection, by one technician. Data Ion Torrent reads were collected by the Ion Torrent Suite software v3.6.2, which also sorted the data according to the barcodes. Data collected on the PGM were collated and reanalysed using the Torrent Suite 3.6.2 using FASTQ files from the Ion Torrent Browser. The software also scores the quality of the reads by assignment of Q17 and Q20 scores according to the Ion Torrent's quality scoring computation.

### NGS bioinformatics analysis: variant identification

Sequence alignment was performed with the Ion Torrent Suite on the Ion Torrent Browser (Life Technologies). Single-nucleotide polymorphisms (SNPs) and short insertions and/or deletions calling was performed using the Variant Caller plugin on the Ion Torrent Browser and DNA sequences visualized using the Integrative Genomics Viewer (IGV, version 2.3) from Broad Institute (Cambridge, MA, USA). The NextGENe software v2.3.3 (Softgenetics, State College, PA, USA) was also used for calling of SNPs and short insertions and/or deletions, and their visualization and annotation. In brief, major calling parameters were chosen as follows: minimum allele frequency (MAF)  $\geq 10\%$  for both SNPs and short insertions and/or deletions, minimum sequencing depth  $\geq 6X$  for SNPs and  $\geq 15X$  for short insertions and/or deletions, and minimum sequencing depth on either strand  $\geq 5X$  for short insertions and/or deletions.

The 10% MAF cutoff filter allowed avoiding any false-negative result. Selected variants using this filter were then analysed in our bioinformatic workflow including two main parameters for variant calling: the variant score (from the NextGENe software, Softgenetics) and the recurrence of the variant. The NextGENe software (Softgenetics) calculated a variant score to provide an empirical estimation of the likelihood that a given variation was not an artifact of sequencing or alignment error (NextGENe Users Manual v2.3, Softgenetics). This score depended mostly on the coverage of the variant. As reads near the 5' end are more accurate than reads at the 3' end, mismatches that were found at the beginning of a read count were weighted more heavily than mismatches that were found in the 3' end of the read. The score was modulated by (i) the ratio of the number of forward reads to the number of reverse reads at the variant location, (ii) variant frequencies in the forward and reverse directions, and (iii) the number of mismatches that occurred in a minimum percentage of reads in the 10-bp region that is found on either of side of the variant. For a run of 48 samples, variant calling was performed for each sample separately. All variants were then collected and aggregated. As no mutation hotspot was described in the *NF1* gene, presence of a variant in more than 2/48 samples

(from 48 unrelated patients) likely indicates a sequencing systematic error or polymorphism.

Presence of several *NF1* partial pseudogenes showing a high sequence identity with the functional *NF1* gene was taken into account in the bioinformatics workflow. During the sequence alignment with the Torrent Suite (Life Technologies), reads mapping at two different positions in the genome were discarded and not included in the following analysis. Sequence alignment was performed against the whole genome to avoid forcing alignment of pseudogene amplicons on the functional *NF1* sequence. Variant calling (using the Variant Caller plugin from Life Technologies and the NextGENe software from Softgenetics) was then performed on the sole functional *NF1* sequence.

#### NGS bioinformatics analysis: single and multiexon deletion/duplication identification

In this method, quantitative values were obtained from the number of reads for each amplicon of each sample, extracted using the *Coverage Analysis* plugin on the Ion Torrent Browser v3.6.2 (Life Technologies). Read number for each separated *NF1* and *SPRED1* amplicon was normalized by dividing each amplicon read number by the total of amplicon read numbers of a control gene from the same sample. For *NF1* CNA analysis, *SPRED1* was considered as the control gene, and reciprocally for the *SPRED1* CNAs. Normalised read numbers obtained for each amplicon of a sample were then divided by the average normalised read number of control samples for the corresponding amplicon. Copy number ratios of  $<0.7$  and  $>1.3$  were considered deleted and duplicated, respectively. DNA samples with *NF1* or *SPRED1* CNAs were submitted to MPLA-based gene dosage for confirmation.

#### Variant confirmation using Sanger sequencing

*NF1* or *SPRED1* variants identified by the NGS approach were confirmed using Sanger DNA sequencing analysis performed on the corresponding exon only, as previously described.<sup>11</sup> Sequences were aligned with the Seqscape analysis software v2.5 (Applied Biosystems, Saint-Aubin, France) and were compared with the corresponding genomic DNA reference sequence NC\_000017.10 (nt 29 422 328–29 701 173 on Human February 2009 (GRCh37/hg19) Assembly). Exons were named according to the NCBI nomenclature (exons numbered 1–58).

#### Single and multi-exon deletion/duplication confirmation using multiplex ligation-dependent probe amplification (MLPA) analysis

Single and multi-exon deletion/duplication screening was performed with MLPA analysis using the SALSA MLPA kits P081/P082 *NF1* and SALSA MLPA P295 *SPRED1* as recommended in the manufacturer's protocol (MRC Holland, Amsterdam, the Netherlands) and as previously described.<sup>11</sup> Briefly, four control samples and each *NF1* patient sample (each containing 100 ng of genomic DNA) were used for overnight hybridization with the probe mixes. After ligation and amplification performed with FAM-labelled primers, PCR products were analysed on an ABI Prism 3130 automatic DNA sequencer (Life Technologies).

#### Assay sensitivity for homopolymeric stretch sequencing

Error rates in homopolymer stretches have been described with PGM sequencers.<sup>26</sup> We retrospectively blindly tested 12 samples with known previous identified mutations in a homopolymeric stretch consisting of a seven-cytosine tract in *NF1* exon 18. Ten samples with duplications of one cytosine (c.2033dupC, p.Ile679fs) and two samples with deletions of one cytosine (c.2033delC, p.Pro678fs) in this homopolymeric stretch were tested.

#### Assay sensitivity for mosaic detection: mutation DNA dilution

To test the ability of NGS to detect mosaic mutation in blood, mock mosaic samples were prepared by serial dilutions of a heterozygous *NF1* mutant DNA (point mutation: c.5017G>T in exon 37) with a normal DNA (each 100 ng/ml). Six serial two-fold dilutions were prepared by diluting mutant DNA into wild-type DNA, corresponding to 25, 12.5, 6.2, 3.1, 1.6, and 0.8% theoretical relative mutation abundances. Serially diluted DNA samples were sequenced using

both NGS and Sanger approaches. NGS sequencing of the six samples on an Ion 316 chip generated a mean 390 Mb per sample. A mean sequencing depth of 1845X was obtained. Sanger sequencing was performed as previously described.<sup>11</sup>

## RESULTS

### *NF1* and *SPRED1* targeted NGS-sequencing design coverage

The targeted region included the entire *NF1*- and *SPRED1*-coding exons, their intron boundaries (25 bp), and the two genes 5' and 3' UTRs. The targeted region (20 595 bp) was amplified by 230 amplicons (length between 125 and 175 bp) distributed in two primer pools.

### Sequencing statistics

For a typical run of 48 samples, ~700 Mb were generated, corresponding to  $\sim 5.2 \times 10^6$  reads. The mean read length was ~135 bp. On average, for every sample, 99% of high-quality sequencing reads (98% of bases) mapped to the reference genome. This resulted in an evenly distributed mean sequencing depth for *NF1* and *SPRED1* of 400X. On average, 98.5% of all targeted bases were covered by at least 20X and 96% by at least 100X (Supplementary Figure S1). A good uniformity between samples and between amplicons was obtained (Supplementary Figures S2 and S3). The only uncovered region in *NF1* corresponded to the six first nucleotides of exon 25 (Supplementary Figure S4). This exon was then sequenced using Sanger method. As expected, these low-covered genomic regions are characterized by low complexity and a high GC content. Sequence targets with these two characteristics are usually difficult to amplify, resulting in reduced coverage for these sites. However, as shown above, this was the case for only a very small proportion of all bases intended to be analysed in this study, located in only one exon (exon 25). The sequence quality metrics of these data warrant a confident detection of variants in all samples. For a comprehensive summary of the obtained sequencing results of a typical 48 sample run, see Supplementary Table S1 and Supplementary Figure S1. We evaluated the NGS analysis costs at approximately €90 for consumables per sample, not including DNA extraction and dosage, bioinformatics analysis, interpretation, Sanger validation, diagnostic reporting, and depreciation of infrastructure costs (Supplementary Figure S5).

### Validation of method

A comprehensive *NF1* mutation screening (NGS and exon 25 Sanger sequencing) was performed for 30 probands of the validation cohort. The selection of the samples for this study was carried out to include as many different types of *NF1* mutations as possible, to simulate a real-world diagnostics scenario, including point mutations, insertions and/or deletions, and exon CNAs, so that we could test the performance of our approach for all these types of genetic alterations (Supplementary Table S2). The 30 tested mutated samples included one whole *NF1* gene deletion, one single exon deletion, five nonsense point mutations, eight missense point mutations, ten short insertion/deletion frameshift mutations, and five intronic mutations. To assess the sensitivity of our assay to detect pathogenic variants, we blindly inspected all mapped sequence reads from the 30 samples with previously defined mutations in *NF1*. The molecular diagnostic strategy that we present here has blindly identified all previously known pathogenic variants in the 30 samples studied. This represents a mutation detection rate of 100%, with zero false-positive calls: all of the called variants were validated by Sanger sequencing, as described in a previous publication using the same amplification-based approach on a PGM instrument (Life Technologies).<sup>27</sup> In our validation sample ( $N=30$ ) with a 100% experimental sensitivity, the probability of a



false-negative event can be estimated at  $3/N$  ( $= 10\%$  with  $N = 30$ ) corresponding to an overall sensitivity  $\geq 90\%$ , calculated using the 'rule of 3' estimate of power according to sample size.<sup>28</sup>

### Diagnostic routine application: *NF1* and *SPRED1* point mutations, and large complete or partial deletions identified in the prospective cohort

A prospective mutation screening of the *NF1* and *SPRED1* genes was performed in 279 probands of the validation cohort. An *NF1* alteration was found in 246/279 (88%) patients. Intragenic *NF1* point variants were found in 226/246 samples (92%). As expected, variant distribution across the *NF1* gene did not exhibit any hotspot domain, and among the 180 distinct point variants, only 24 were recurrent (Supplementary Table S3). Small insertions and/or deletions accounted for 29% of the 246 *NF1* variants identified in our study ( $N = 70/246$ ), most of them with frameshift consequences ( $N = 60$ ). Splice alterations accounted for 23% ( $N = 58/246$ ). Nonsense variants corresponded to 29% ( $N = 70/246$ ), whereas missense variants accounted for  $\sim 11\%$  ( $N = 28/246$ ). Deletion or duplication of several *NF1* exons accounted for  $\sim 2\%$  of all *NF1* alterations ( $N = 6/246$ ): five multi-exons deletions and one multi-exon duplication were found. Whole *NF1* gene deletions were found in  $\sim 6\%$  ( $N = 14/246$ ) of all *NF1* alterations. The 180 distinct intragenic variants and the six multi-exon deletions or duplication (Supplementary Table S3) were reported according to the standard human genome variation society (HGVS) nomenclature (<http://www.hgvs.org/mutnomen/>). These variants have been deposited in the LOVD ([https://grenada.lumc.nl/LOVD2/mendelian\\_genes/home.php?select\\_db=NF1](https://grenada.lumc.nl/LOVD2/mendelian_genes/home.php?select_db=NF1)).

A *SPRED1* alteration was found in 10/279 ( $\sim 4\%$ ) patients (Supplementary Table S4). Among the eight distinct point variants, only one was recurrent: the substitution  $c.46C>T$ , p.Arg16\* was found in three unrelated patients. All the eight distinct *SPRED1* variants were supposed to lead to *SPRED1* protein loss-of-function as four of them were nonsense mutations and four were short frameshift insertions/deletions. Clinical features of the 10 patients carrying *SPRED1* variants are summarized in Supplementary Table S4. All *SPRED1* variants have been deposited in a database created using the LOVD software and are accessible at [https://grenada.lumc.nl/LOVD2/mendelian\\_genes/home.php?select\\_db=SPRED1](https://grenada.lumc.nl/LOVD2/mendelian_genes/home.php?select_db=SPRED1).

All *NF1* and *SPRED1* point variants and short insertions and/or deletions identified by NGS were confirmed using Sanger sequencing of the corresponding exons. All single and multi-exon deletions were confirmed using the MLPA analysis. Finally, an *NF1* or *SPRED1* alteration was found in 256/279 (92%) samples by our 48 multiplexed sample NGS strategy. No *NF1* or *SPRED1* mutation was found in 23 probands.

All the negative cases ( $N = 23/279$ ; 8%) were then analysed using both DNA and cDNA Sanger sequencing and MLPA approaches as previously described<sup>11</sup> but no *NF1* or *SPRED1* mutation was found. Among the 23 probands with no *NF1* or *SPRED1* mutation, 14 negative patients (14/23) showed a mild sporadic form of *NF1*, including café-au-lait spots or freckling. Three negative patients (3/23) showed a typical familial form of *NF1*. The six remaining negative patients (6/23) were affected by a sporadic segmental *NF1*. This peculiar mosaic-localized form of *NF1* may explain the absence of detectable somatic *NF1* mutation in blood.<sup>29–31</sup>

### Detection of mutations in homopolymeric stretches

In the seven cytosine homopolymeric stretch, wild-type control samples recurrently showed artifact deletions of one cytosine with a mean 6.1% MAF and artifact duplication with a mean 4.7% MAF

(Supplementary Figure S6). The 10 true-positive duplications (control samples) were observed with a mean 38.3% MAF. The two true-positive deletions (control samples) were observed with a mean 50.2% MAF. Filtering against MAFs allowed differentiating between true-positive and artifact deletions in the homopolymeric stretch. Variants in this homopolymeric stretch were therefore correctly called with our bioinformatics pipeline.

### Mosaic detection

The six *NF1*-negative samples from patients showing a segmental form of *NF1* were then analysed by a new NGS run with a greater depth of coverage in order to increase detection sensitivity. Multiplexing of six patients instead of 48 allowed a 2600X depth. One *NF1* point variant (exon 21:  $c.2446C>T$  leading to p.Arg816\*) was found in one patient among the six segmental patients. Variant allele frequency was  $\sim 7\%$  indicating a mosaic mutation.

### Tumour somatic DNA screening: two *NF1* hits detection in MPNSTs

Retrospective analysis of 15 *NF1*-associated MPNST samples also allowed the identification of two *NF1* alterations in 13/15 (80%) MPNSTs (Supplementary Table S5).

## DISCUSSION

Molecular diagnosis of *NF1* is challenging owing to the large size of the gene, and the lack of mutational hotspots with a complex mutational spectrum.<sup>11</sup> To date, *NF1* alterations have been identified in routine diagnostics with a number of techniques. In our previous experience, *NF1* mutation screening included intragenic *NF1* microsatellite analysis (for large deletion screening), real-time PCR-based gene dosage (large deletion confirmation), and *NF1* Sanger sequencing at both cDNA (point mutation screening), and DNA levels (point mutation confirmation).<sup>11</sup> When a molecular abnormality was not identified at the cDNA level, exon deletion/duplication screening was performed using MLPA. In case of a negative result, mutation screening was performed by sequencing the 57 constitutive coding exons and the alternative spliced exon 31 at the DNA level. These procedures were expensive, time-consuming, and labour-intensive. Here, we report our experience of a targeted NGS approach for *NF1* and *SPRED1* mutations and CNA screening, in a large cohort of *NF1* patients. We confirm this method as a reference for routine molecular analysis.

In the first step (method validation), we aimed to determine whether the sensitivity and specificity of our targeted NGS strategy was equal to those of our previous molecular diagnosis strategy. In total, 30 individuals with an *NF1* alteration were retrospectively sequenced using a multiplex PCR (Ampliseq) approach on a PGM sequencer (Life Technologies). The 316 chip capacity allowed mixing 48 bar-coded samples, providing a good average sequencing depth. On average, 98.5% of all targeted bases were covered by at least 20X and 96% by at least 100X, using 230 amplicons. All previously known pathogenic variants in the 30 samples studied were identified (including point mutations, short insertions and/or deletions, and exon CNAs; Supplementary Table S2), highlighting the accuracy of this strategy as a molecular diagnostic tool for *NF1*, as previously described.<sup>27</sup> As 48 samples were multiplexed on one chip, genotyping throughput was increased over 10 times, as compared with Sanger sequence genotyping. In our hands, the *NF1* and *SPRED1* diagnostic NGS analysis cost approximately €90 for consumables per sample (Supplementary Figure S5). The NGS approach thus appears to be far

cheaper than Sanger sequencing of the 57 constitutive coding *NF1* exons and the seven *SPRED1* exons.

In the second step, 279 DNAs from individuals referred with an *NF1* clinical diagnosis were prospectively sequenced in our routine molecular diagnosis laboratory, using the targeted NGS sequencing of the *NF1*- and *SPRED1*-coding sequences and intron–exon boundaries (25 bp). An *NF1* alteration was identified in 246 of 279 index cases (88%; Supplementary Table S3). This high mutation detection sensitivity can be achieved in well-phenotyped *NF1* patients. It is comparable to the one obtained with previous screening methodologies including Sanger sequencing.<sup>5,6,8,10,11</sup> Of course, the single DNA sequencing can miss some *NF1* alterations including deep intronic mutations that represent a weak proportion of the whole mutation types.<sup>11</sup> The distribution of the different types of mutations found by the NGS approach was the same than previously reported.<sup>11,32</sup> Our targeted NGS approach also provided a quantitative information (based on sequencing depth) that allowed identification of deletion of one or several exons in ~2% ( $N=6/246$ ) of all *NF1* alterations, in exact accordance with literature data.<sup>8</sup> Using the same analysis, whole *NF1* gene deletions were identified in ~6% ( $N=14/246$ ) of *NF1* patients, in accordance with previous data.<sup>14</sup> MLPA confirmed the identification of these rearrangements.

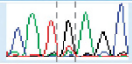
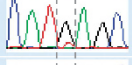
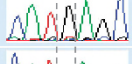



*SPRED1* screening was included in the NGS-targeted design. Therefore, *SPRED1* mutation testing was systematically performed in the *NF1* screening. A *SPRED1* loss-of-function mutation was identified in 10 of 279 index cases (~4%). All *SPRED1*-mutated patients were addressed for *NF1* clinical suspicion. Interestingly, *SPRED1* loss-of-function mutated patients represent one-third (10/33) of *NF1* mutation-negative patients in this cohort, highlighting the relevance of simultaneous *NF1* and *SPRED1* screening. Their phenotype included café-au-lait spots, freckling, learning disabilities, and development delay (Supplementary Table S4), as previously reported.<sup>21,22,33,34</sup> Because of the important clinical overlap with *NF1*, it is impossible to diagnose the syndrome exclusively based on clinical feature alone. In patients presenting solely with café-au-lait spots and/or frecklings, a molecular confirmation has to be performed.

Molecular diagnosis in *NF1* is helpful to confirm the diagnostic, notably in patients with paucisymptomatic forms: in the case of

pediatric forms (café-au-lait macules are often the only clinical findings in young children), segmental forms, and in Legius syndrome. Molecular diagnosis has implication for the family and prenatal testing can be offered. Individuals with *NF1* generally require routine clinical management to detect and manage the manifestations. On the basis of the current understanding of Legius syndrome, affected individuals are not at increased risk for the typical *NF1*-associated tumours,<sup>21,22,33,34</sup> and generally do not require the same intensity of clinical management. A diagnosis of Legius syndrome compared with *NF1* may provide both peace of mind due to decreased risk of a tumour phenotype and cost savings for the individual due to less aggressive medical management.<sup>35</sup> Distinguishing between these two entities is thus important for prognosis, counseling, prenatal diagnosis, and clinical management.

By a proof-of-concept mutation dilution experiment, we showed that NGS may be the method of choice for *NF1* mosaic detection: NGS showed a high sensitivity with a quantitative aspect (Figure 1). Bioinformatic filtering parameters were adapted for somatic mutation detection: MAF was set at  $\geq 5\%$ . A greater depth of coverage was obtained by multiplexing six negative samples (instead of 48) from segmental *NF1* patients in one NGS run. This approach allowed the identification of one mosaic *NF1* point mutation with a ~7% minor allele frequency. As a more important sequencing depth is needed to increase sensitivity, such mosaic screening is more expensive and time-consuming than constitutional mutation detection. In routine molecular diagnostic, this application may be a second-line *NF1* mutation screening in selected sporadic *NF1* patients (with segmental form of the disease, for example) with no mutation found by NGS screening at a 'regular' sequencing depth. The aggregated data obtained from all patients with mosaic *NF1* clearly indicate that the percentage of cells carrying the *NF1* first hit mutation is often too low to be reliably detected in blood by a Sanger approach.<sup>30</sup> The frequency of mosaicism in sporadically affected *NF1* patients is currently still largely unknown.<sup>36</sup>

We also showed that our NGS *NF1* mutation-screening strategy can be applied to tumour analysis. Two *NF1* hits were found in 13/15 of tested MPNSTs (Supplementary Table S5). This observation is in accordance with previous data showing a ~90% identification rate of the two *NF1* hits in MPNSTs.<sup>37</sup> One-third (8/28) of the *NF1* hits were

| Theoretical c.5017G>T mutation allele frequency | NGS results                        |                               | Sanger results                     |   |
|---|------------------------------------|-------------------------------|------------------------------------|---|
|   | Observed mutation allele frequency | Number of mutated reads/depth | Observed mutation allele frequency | Electropherogram<br>C A T T A G C   |
| 25%   | 27.3%                              | 544/1991                      | Clearly detected in both strands   |  |
| 12.5%   | 13.3%                              | 231/1739                      | Detected in forward strand         |  |
| 6.25%   | 7.1%                               | 141/1988                      | Not detected                       |  |
| 3.1%  | 3.2%                               | 63/1990                       | Not detected                       |  |
| 1.6%  | 2.2%                               | 58/2642                       | Not detected                       |  |
| 0.8%  | 1.1%                               | 24/2202                       | Not detected                       |  |

**Figure 1** Mutation dilution sequencing using NGS and Sanger methods. NGS allowed the mutation detection in all dilutions and also presented a quantitative aspect, providing the number of mutated reads. A good correspondence was found between the theoretical and the NGS observed mutation allele frequencies. Sanger sequence analysis with Seqscape v2.5 allowed mutation visualization only at 25 and 12.5%.

large rearrangements. The high prevalence of such rearrangements in MPNST contrast with the 5–10% found in constitutional DNAs from NF1 patients and is in accordance with previously identified MPNST somatic *NF1* mutation spectrum.<sup>37</sup> As molecular analysis of the *NF1* gene is challenging, *NF1* tumour mutation-spectrum may have been less studied than other genes with gain-of-function mutations. However, *NF1* was recently identified as one of the most significantly somatic mutated genes in the Cancer Genome Atlas (TCGA) that provided molecular tumour maps using exome sequencing in more than 3000 tumours from 12 cancer types.<sup>38,39</sup> As *NF1* encodes a negative regulator of RAS-MAPK signalling, it has recently been suggested that cooperation between *BRAF* and *NF1* could lead to melanoma development and *BRAF* inhibitor resistance.<sup>40–42</sup> NGS will allow high throughput and cost-effective screen of *NF1* somatic alterations to confirm these data.

In conclusion, NGS constitutes an immense improvement for *NF1* molecular alteration screening, as a genotyping method comparable to Sanger sequencing-based typing in quality, while being more cost-effective, and higher throughput.

### CONFLICT OF INTEREST

The authors declare no conflict of interest.

### ACKNOWLEDGEMENTS

This work was supported in part by grants from *Ligue Française Contre les Neurofibromatoses*, and *Association Neurofibromatoses et Recklinghausen*. We thank the Platform of Biological Ressources, Assistance Publique Hôpitaux de Paris (AP-HP), Hôpital Henri Mondor, Créteil, France, for providing MPNST samples.

- 1 Carey JC, Baty BJ, Johnson JP, Morrison T, Skolnick M, Kivlin J: The genetic aspects of neurofibromatosis. *Ann NY Acad Sci* 1986; **486**: 45–56.
- 2 National Institutes of Health Consensus Development Conference Statement. Neurofibromatosis. *Arch Neurol Chicago* 1988; **45**: 575–578.
- 3 Tucker T, Wolkenstein P, Revuz J, Zeller J, Friedman JM: Association between benign and malignant peripheral nerve sheath tumors in NF1. *Neurology* 2005; **65**: 205–211.
- 4 Clementi M, Barbujani G, Turolla L, Tenconi R: Neurofibromatosis-1: a maximum likelihood estimation of mutation rate. *Hum Genet* 1990; **84**: 116–118.
- 5 Fahsold R, Hoffmeyer S, Mischung C *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**: 790–818.
- 6 Messiaen L, Callens T, Mortier G *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**: 541–555.
- 7 Ars E, Kruyer H, Morell M *et al*: Recurrent mutations in the *NF1* gene are common among neurofibromatosis type 1 patients. *J Med Genet* 2003; **40**: e82.
- 8 Wimmer K, Yao S, Claes K *et al*: Spectrum of single- and multiexon *NF1* copy number changes in a cohort of 1100 unselected *NF1* patients. *Genes Chromosomes Cancer* 2006; **45**: 265–276.
- 9 Pros E, Gómez C, Martín T, Fábregas P, Serra E, Lázaro C: Nature and mRNA effect of 282 different *NF1* point mutations: focus on splicing alterations. *Hum Mutat* 2008; **29**: E173–E193.
- 10 Valero MC, Martín Y, Hernández-Imaz E *et al*: A highly sensitive genetic protocol to detect *NF1* mutations. *J Mol Diagn* 2011; **13**: 113–122.
- 11 Sabbagh A, Pasmant E, Imbard A *et al*: *NF1* molecular characterization and neurofibromatosis type 1 genotype-phenotype correlation: the French experience. *Hum Mutat* 2013; **34**: 1510–1518.
- 12 Mautner VF, Kluwe L, Friedrich RE *et al*: Clinical characterisation of 29 neurofibromatosis type-1 patients with molecularly ascertained 1.4 Mb type-1 *NF1* deletions. *J Med Genet* 2010; **47**: 623–630.

- 13 De Raedt T, Brems H, Wolkenstein P *et al*: Elevated risk for MPNST in *NF1* microdeletion patients. *Am J Hum Genet* 2003; **72**: 1288–1292.
- 14 Pasmant E, Sabbagh A, Spurlock G *et al*: *NF1* microdeletions in neurofibromatosis type 1: from genotype to phenotype. *Hum Mutat* 2010; **31**: E1506–E1518.
- 15 Castle B, Baser ME, Huson SM, Cooper DN, Upadhyaya M: Evaluation of genotype-phenotype correlations in neurofibromatosis type 1. *J Med Genet* 2003; **40**: e109.
- 16 Easton DF, Ponder MA, Huson SM, Ponder BA: An analysis of variation in expression of neurofibromatosis (NF) type 1 (*NF1*): evidence for modifying genes. *Am J Hum Genet* 1993; **53**: 305–313.
- 17 Sabbagh A, Pasmant E, Laurendeau I *et al*: Unravelling the genetic basis of variable clinical expression in neurofibromatosis 1. *Hum Mol Genet* 2009; **18**: 2768–2778.
- 18 Pasmant E, Vidaud M, Vidaud D, Wolkenstein P: Neurofibromatosis type 1: from genotype to phenotype. *J Med Genet* 2012; **49**: 483–489.
- 19 Upadhyaya M, Huson SM, Davies M *et al*: An absence of cutaneous neurofibromas associated with a 3-bp inframe deletion in exon 17 of the *NF1* gene (c.2970–2972 delAAT): evidence of a clinically significant *NF1* genotype-phenotype correlation. *Am J Hum Genet* 2007; **80**: 140–151.
- 20 Brems H, Beert E, de Ravel T, Legius E: Mechanisms in the pathogenesis of malignant tumours in neurofibromatosis type 1. *Lancet Oncol* 2009; **10**: 508–515.
- 21 Brems H, Chmara M, Sahbatou M *et al*: Germline loss-of-function mutations in *SPRED1* cause a neurofibromatosis 1-like phenotype. *Nat Genet* 2007; **39**: 1120–1126.
- 22 Pasmant E, Sabbagh A, Hanna N *et al*: *SPRED1* germline mutations caused a neurofibromatosis type 1 overlapping phenotype. *J Med Genet* 2009; **46**: 425–430.
- 23 Brems H, Pasmant E, Van Minkelen R *et al*: Review and update of *SPRED1* mutations causing Legius syndrome. *Hum Mutat* 2012; **33**: 1538–1546.
- 24 Yu H, Zhao X, Su B *et al*: Expression of *NF1* pseudogenes. *Hum Mutat* 2005; **26**: 487–488.
- 25 Ferner RE, Huson SM, Thomas N *et al*: Guidelines for the diagnosis and management of individuals with neurofibromatosis 1. *J Med Genet* 2007; **44**: 81–88.
- 26 Loman NJ, Misra RV, Dallman TJ *et al*: Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol* 2012; **30**: 434–439.
- 27 Balla B, Arvai K, Horváth P *et al*: Fast and robust next-generation sequencing technique using ion torrent personal genome machine for the screening of neurofibromatosis type 1 (*NF1*) gene. *J Mol Neurosci* 2014; **53**: 204–210.
- 28 Mattocks CJ, Morris MA, Matthijs G *et al*: A standardized framework for the validation and verification of clinical molecular genetic tests. *Eur J Hum Genet* 2010; **18**: 1276–1288.
- 29 Biesecker LG, Spinner NB: A genomic view of mosaicism and human disease. *Nat Rev Genet* 2013; **14**: 307–320.
- 30 Maertens O, De Schepper S, Vandensompele J *et al*: Molecular dissection of isolated disease features in mosaic neurofibromatosis type 1. *Am J Hum Genet* 2007; **81**: 243–251.
- 31 Ruggieri M, Huson SM: The clinical and diagnostic implications of mosaicism in the neurofibromatoses. *Neurology* 2001; **56**: 1433–1443.
- 32 Messiaen LM, Wimmer K: Pitfalls of automated comparative sequence analysis as a single platform for routine clinical testing for *NF1*. *J Med Genet* 2005; **42**: e25.
- 33 Spurlock G, Bennett E, Chuzhanova N *et al*: *SPRED1* mutations (Legius syndrome): another clinically useful genotype for dissecting the neurofibromatosis type 1 phenotype. *J Med Genet* 2009; **46**: 431–437.
- 34 Messiaen L, Yao S, Brems H *et al*: Clinical and mutational spectrum of neurofibromatosis type 1-like syndrome. *JAMA* 2009; **302**: 2111–2118.
- 35 Muram TM, Stevenson DA, Watts-Justice S *et al*: A cost savings approach to *SPRED1* mutational analysis in individuals at risk for neurofibromatosis type 1. *Am J Med Genet A* 2013; **161**: 467–472.
- 36 Messiaen L, Vogt J, Bengesser K *et al*: Mosaic type-1 *NF1* microdeletions as a cause of both generalized and segmental neurofibromatosis type-1 (*NF1*). *Hum Mutat* 2011; **32**: 213–219.
- 37 Upadhyaya M, Kluwe L, Spurlock G *et al*: Germline and somatic *NF1* gene mutation spectrum in *NF1*-associated malignant peripheral nerve sheath tumors (MPNSTs). *Hum Mutat* 2008; **29**: 74–82.
- 38 Kandath C, McLellan MD, Vandin F *et al*: Mutational landscape and significance across 12 major cancer types. *Nature* 2013; **502**: 333–339.
- 39 Ciriello G, Miller ML, Aksoy BA, Senbabaoglu Y, Schultz N, Sander C: Emerging landscape of oncogenic signatures across human cancers. *Nat Genet* 2013; **45**: 1127–1133.
- 40 Gibney GT, Smalley KS: An unholy alliance: cooperation between *BRAF* and *NF1* in melanoma development and *BRAF* inhibitor resistance. *Cancer Discov* 2013; **3**: 260–263.
- 41 Whittaker SR, Theurillat JP, Van Allen E *et al*: A genome-scale RNA interference screen implicates *NF1* loss in resistance to *RAF* inhibition. *Cancer Discov* 2013; **3**: 350–362.
- 42 Maertens O, Johnson B, Hollstein P *et al*: Elucidating distinct roles for *NF1* in melanomagenesis. *Cancer Discov* 2013; **3**: 338–349.

Supplementary Information accompanies this paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)