Tohoku J. Exp. Med., 2017, **242**, 173-181

Targeted Next-Generation Sequencing Newly Identifies Mutations in *Exostosin-1* and *Exostosin-2* Genes of Patients with Multiple Osteochondromas

Xiaoyan Guo,[1] Mingrui Lin,[2] Tengfei Shi,[1] Wei Yan[3] and Wenxu Chen[1]

[1]Department of Laboratory Medicine, The Second Hospital of Fuzhou Affiliated to Xiamen University, Fuzhou, Fujian, China
[2]Intensive Care Unit, The Affiliated People's Hospital of Fujian Traditional Medical University, Fuzhou, Fujian, China
[3]Department of Bone Tumors, The Second Hospital of Fuzhou Affiliated to Xiamen University, Fuzhou, Fujian, China

Multiple osteochondromas (MO) is one of the most common benign bone tumors in humans with an autosomal dominant hereditary mode. MO is a genetic heterogeneity disease with variable number and size of osteochondromas, as well as changeable number and location of diseased bones. Mutations in *Exostosin-1*/*Exostosin-2* (*EXT1*/*EXT2*) genes are the main molecular basis of MO. *EXT1* and *EXT2* genes encode exostosin 1 and exostosin 2, respectively, both of which are transmembrane glycosyltransferases that elongate the chains of heparin sulfate (HS) at HS proteoglycans (HSPGs). HSPGs are considered to be involved in regulating the proliferation and differentiation of chondrocytes. Owing to large size of *EXT1*/*EXT2* genes and lack of mutation hotspots, molecular diagnosis of MO is challenging. Here, we applied targeted next-generation sequencing (t-NGS) in mutation screening of *EXT1*/*EXT2* genes for 10 MO patients. The results were compared and validated with Sanger sequencing. Overall, nine mutations identified by t-NGS were confirmed with Sanger sequencing, excluding two variants of false positive, suggesting the reliability of mutation screening by t-NGS. The nine mutations identified by t-NGS include two missense mutations (*EXT1*: c.1088G>A and c.2120C>T), one splicing mutation (*EXT2*: c.744-1G>T), and six nonsense mutations (*EXT1*: c.351C>G, c.1121G>A, and c.1843_1846dup; EXT2: c.67C>T, c.561delG, and c.575T>A). In summary, our paper provides the primary data of the application of t-NGS in MO molecular diagnosis, including six newly identified mutations (*EXT1*: c.1843_1846dup, c.1088G>A, c.351C>G, and c.2120C>T and *EXT2*: c.744-1G>T and c.575T>A), which further enrich the mutation database of MO from the Chinese population.

Keywords: *EXT1*/*EXT2* genes; molecular diagnosis; multiple osteochondromas; mutation screening; targeted next-generation sequencing
Tohoku J. Exp. Med., 2017 July, **242** (3), 173-181.

Introduction

Multiple osteochondromas (MO) is a commonly benign skeletal disease characterized by multiple cartilage-capped bony protuberances in the growth plate of long and tubular bones. MO is an autosomal dominant disease with an estimation of 1/50,000 in Caucasian population (Schmale et al. 1994). MO is a genetic heterogeneity disease with variable number and size of osteochondromas, as well as changeable number and location of diseased bones. Furthermore, complications are usually encountered in the patients, such as deformities of involved joints, functional limitations of joints and limbs, compression of nerves and blood vessels due to protruding masses, and short stature (Jennes et al. 2012; Bozzola et al. 2015; Aouini et al. 2015). However, peripheral chondrosarcoma is deemed to be the most serious complication with an incidence of 0.5-2% (Wicklund et al. 1995).

MO mainly results from mutations in two tumor-suppressor genes: *Exostosin-1* (*EXT1*) (MIM*608177) (Ahn et al. 1995) and *Exostosin-2* (*EXT2*) (MIM*608210) (Stickens et al. 1996; Wuyts et al. 1996). Nevertheless, no obvious mutations are detected from *EXT1* or *EXT2* gene in some families or sporadic patients with a rate of 10-15%, suggesting that some special types of mutations (inversions, translocations and somatic mosaicism) provoke the disease

Received December 6, 2016; revised and accepted June 16, 2017. Published online July 7, 2017; doi: 10.1620/tjem.242.173.
Correspondence:Wenxu Chen, Department of Laboratory Medicine, The Second Hospital of Fuzhou Affiliated to Xiamen University, No.47, Shangteng Road, Fuzhou, Fujian 350007, China.
e-mail: chenwx1969@126.com

(Vink et al. 2005; Szuhai et al. 2011). Most mutations (80%) in *EXT1*/*EXT2* genes are nonsense, splice-site, and small frameshift mutations and often introduce a premature stop codon that silences the functional genes (Jennes et al. 2009). Others are missense mutations (Cheung et al. 2001) and single-/multi-exon deletions/insertions (Jennes et al. 2008, 2011). Approximately 90% of MO patients have an affected parent, while about 10% of individuals develop MO owing to a de novo mutation (Jennes et al. 2009). So far, 436 variants from *EXT1* gene and 223 variants from *EXT2* gene have been described in Multiple Osteochondroma Mutation Database (http://medgen.ua.ac.be/LOVDv.2.0/home.php).

EXT1 and *EXT2* genes encode exostosin 1 and exostosin 2, respectively, both of which act as transmembrane glycosyltransferases in the endoplasmic reticulum and Golgi apparatus. Exostosin 1 and exostosin 2 form a hetero-oligomeric complex that plays a role in the elongation of heparan sulfate (HS); HS is associated with the transduction of signal, which regulates the differentiation, ossification, and apoptosis of the chondrocyte (Senay et al. 2000; Esko and Lindahl 2001). Molecular diagnosis of MO is relatively complex and challenging due to the big size of the two genes and the absence of mutation hotspots. Generally speaking, Sanger sequencing combined with multiplex ligation-dependent probe amplification (MLPA) (White et al. 2004) and array comparative genomic hybridization (array-CGH) (Szuhai et al. 2011) is able to cover almost all types of mutations in *EXT1*/*EXT2* genes, including point mutations, small fragment insertions/deletions, single-/multi-exon deletions/insertions and some inversions, translocations or somatic mosaicism. However, time-consuming and multi-step procedures, as well as high cost are the main problems from these methods. Whole genome sequencing or exome sequencing is an alternative technique for screening mutations in MO with a broad spectrum (Zhang et al. 2013; Liu et al. 2015), but they are also costly and need long checking-time. Second-generation sequencing, targeted next-generation sequencing (t-NGS), has already been applied in molecular diagnosis for neurofibromatosis type 1 (NF1) and Duchenne and Becker muscular dystrophies (DMD) with excellent accuracy and high efficiency (Pasmant et al. 2015; Okubo et al. 2016). T-NGS specially aims to targeted genes and has successfully detected almost all subtle mutations (point mutation, small deletions/insertions, mosaic mutation) and large deletions/insertions for a large cohort in a few days. Compared to the whole genome sequencing or exome sequencing, t-NGS is a cost-effective method for some genetic diseases with known pathogenic genes. In this paper, we introduced and validated a t-NGS approach through screening mutations for 5 EXT-mutated MO patients and prospectively analyzing 5 other MO patients, thereby developed a new diagnostic program for MO.

Materials and Methods

MO patients and samples

All MO cases were inpatients from the Second Hospital of Fuzhou Affiliated to Xiamen University, they were definitely diagnosed based on clinical manifestations, imaging and histopathological examinations. After the signing of informed consent, peripheral blood samples were obtained from 10 MO families, and 10 probands from these families were picked out to test and verify the clinical utility of t-NGS. Among them, 5 probands had been identified with mutations by Sanger sequencing previously, the remaining 5 probands were directly screened for mutations by t-NGS. Normal controls were healthy people with normal physical examinations in the hospital. This study was approved by the Ethics Committee of the Second Hospital of Fuzhou Affiliated to Xiamen University.

T-NGS of EXT1 and EXT2 genes

Genomic DNA was extracted from peripheral blood lymphocytes of MO probands and normal controls according to the procedures of SE Blood DNA Kit (Omega Bio-Tek, CA, USA), DNA purification was processed by Ampure Beads, and the purified DNA was tested by t-NGS with minimum amount of 30 ng for each sample. Samples were divided into two groups which were tested on an Ion Personal Genome Machine (PGM) sequencer (Thermo Fisher Scientific, MA, USA) one after another in different times: group 1 included five MO subjects with known variants which had been identified by Sanger sequencing; group 2 contained five newly gathered MO subjects without any information of gene mutation (Table 1). The design of multiplex primer pools targeting *EXT1*/*EXT2* genes was based on Ion AmpliSeq Designer software (Thermo Fisher Scientific) (Table 2). The targeted regions covered 100% of *EXT1*/*EXT2* genes, which mainly involved 100% of the coding regions, part of intronic regions from exon–intron boundary (10 base pairs), the 5′- and 3′-untranslated regions (UTRs). Theoretically the entire coverage could reach 99.2%. The *EXT1* gene (reference sequence NM_000127.2) was theoretically covered at 100%, the *EXT2* gene (reference sequence NM_207122.1) was theoretically covered at 98.4% (64 base pairs in the 5′ UTR of *EXT2* gene were not covered). Targeted fragments were sequenced using an Ion AmpliSeq Library Kit 2.0 (Thermo Fisher Scientific) and performed on the Ion PGM sequencer utilizing 318 Chip (Thermo Fisher Scientific) based on the manufacturer's protocol.

T-NGS bioinformatics analysis: variant identification

Ion Torrent Suite software 4.0.1 (Life Technologies) was used to extract sequence data and align with sequences of *EXT1* and *EXT2* genes in human. Variants such as single nucleotide variants (SNVs), small insertions/deletions were picked out through the Variant Caller plugin. The major variant screening parameters were set as follows: (1) For SNVs and small insertions/deletions, minimum allele frequency (MAF) was set for ≥ 10%. (2) For SNVs, minimum sequencing depth was set for ≥ 6 ×, and ≥ 15 × for small insertions/deletions. (3) For small insertions/deletions, minimum sequencing depth on either strand was set for ≥ 5 ×. The 10% MAF cutoff filter was used to avoid any possible false-negative event. Potential pathogenic variants selected from this filter were further analyzed based on two algorithms: variant score and variant frequency. Variant score was calculated through the NextGENe software (Softgenetics), while variant frequency referred to the 137 SNP database.

Table 1. Clinical phenotype and genotype of MO probands in the study.

Subject (group)	Sex	Onset (years)	Involved sites	Functional limitation*	Family history	Gene	Exon/ Intron	Type of mutation	DNA change	Protein change
1(1)	male	4	femurs, tibia, fibula	+	yes	EXT1	Exon9	Frameshift/ Nonsense	c.1843_1846 dup	p.Tyr616Ter
2(1)	male	8	limbs,ribs, shoulder	—	no	EXT1	Exon3	Missense	c.1088G>A	p.Gly363Glu
3(1)	male	10	femur, tibia, fibula	++	no	EXT2	Intron4	Splice site	c.744-1G>T	-
4(1)	male	1	shoulders, chest, limbs	+	no	EXT2	Exon 3	Nonsense	c.575T>A	p.Leu192Ter
5(1)	male	10	forearm,knees	++	yes	EXT2	Exon 3	Frameshift/ Nonsense	c.561delG	p.Leu269Ter
6(2)	male	2	wrist, shoulder, knees,ankle	+	no	EXT1	Exon 1	Nonsense	c.351C>G	p.Tyr117Ter
7(2)	female	13	knees, tibial	—	no	EXT1	Exon 3	Nonsense	c.1121G>A	p.Trp374Ter
8(2)	male	6	forearms, wrists,knees	+	yes	EXT1	Exon11	Missense	c.2120C>T	p.Thr707Met
9(2)	male	19	femurs,tibias, fibulas, shoulder	—	no	EXT2	Exon2	Nonsense	c.67C>T	p.Arg23Ter
10(2)	male	16	shoulder,knee	—	no	-	-	-	-	-

*+, joint deformity; ++, joint deformity, nerve compression or injury, or shortened limb; −, normal in joints and nerves.

Variant validation through Sanger sequencing and pathogenicity analysis for the variants

EXT1/*EXT2* variants filtered from the t-NGS approach were validated separately by Sanger sequencing of the corresponding mutated fragments. Sequence alignments and analysis were carried out respectively through DANMAN8.0 (Lynnon Biosoft, USA) and Chromas2 (Technelysium Pty Ltd, Tewantin, qLD, Australia). Two currently published genomic DNA sequences (NC_000008.11 and NC_000011.10) were chosen as reference sequences. Exons of *EXT1* gene complied with the NCBI nomenclature (numbered 1-11), wheras exons of *EXT2* gene were numbered 1-15. The primers, procedures and annealing temperatures of PCR before Sanger sequencing referred to two literatures (Wuyts et al. 2005; Tian et al. 2014).

The exostosin sequences of 18 species were downloaded from NCBI (http://www.ncbi.nlm.nih.gov/) to analyze the conservation of mutation sites. Three bioinformatics softwares—PolyPhen (Ramensky et al. 2002), SIFT (Ng and Henikoff 2003) and PROVEAN (Choi and Chan 2015) were selected to predict the functional effect of amino acid changes. The CRYP-SKIP server (http://cryp-skip.img.cas.cz/), MaxEntScan (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq_acc.html) and Human Splicing Finder (HSF) (http://www.umd.be/HSF/) were used to analyze the splice mutation. The corresponding fragments which mutated in MO patients from 100 normal controls (200 alleles) were sequenced for polymorphism analysis.

Results

Clinical, imaging, histopathological examinations and pedigree investigation

From all subjects invited in this study we obtained detailed physical, imaging examinations and pedigree investigation. The diagnosis was all confirmed by histopathological examination finally. The affected parts mainly involved long bones, adjacent joints, nerve and their function (Table 1).

Identification of EXT1/EXT2 variants by t-NGS

Five MO patients in group 1 (Subjects 1-5) with probable causative variants through Sanger sequencing were detected by t-NGS on a 318 chip successfully, and normal control (NC) was set for assessing the efficiency of capturing the targeted area. A total of 78.3 MB data was obtained, and the coverage of the targeted area was close to or exceeded 95% except subject 5 (Table 3). Finally, five variants in subjects 1-5 were all discovered by t-NGS, in which two variants were identified in *EXT1* gene and three in *EXT2* gene (Fig. 1, Table 1). Then, five newly gathered MO patients with unknown variants (Subjects 6-10 in group 2) were detected by t-NGS; 56 MB data were achieved and the coverage of the targeted area was 93.55% (20×) (Table 3). However, only four suspicious pathogenic variants were separately detected in subjects 6-9 by t-NGS, and the variants were validated by Sanger sequencing (Fig. 1, Table 1). No pathogenic variant was detected in subject 10 by t-NGS, or by Sanger sequencing.

In addition, three variants (*EXT1*: c.1065C>T and c.1761G>A; and *EXT2*: c.28 C>A) were detected by t-NGS and confirmed by Sanger sequencing. However, they were all synonymous mutations in which c.1065C>T was found in subject 3, c.1761G>A in subjects 4-6 and 9, and c.28 C>A in subjects 3, 4, 6 and 10.

Exclusion of false positive and pathogenicity analysis of three non-synonymous SNVs

Subjects 5, 6, and 10 were all found with a point mutation of one thymine (c.59T>G, p.Leu20Trp) and a deletion of one thymine (c.64delT, Tyr22Ilefs*114) by t-NGS, both of which were discovered from a homopolymeric stretch in exon 1 of *EXT1* gene; however, the two variants were verified to be false positive by Sanger sequencing, ultimately.

Table 2. Multiplex primers designed for *EXT1* and *EXT2* genes.

Amplicon_ID	Ion_AmpliSeq_Fwd_Primer	Ion_AmpliSeq_Rev_Primer
AMPL7153088408	CCAGGAGAGGTGATAATGTTAAACCC	GCTTTGCGTAAATTCATGCACATG
AMPL7153088419	GACCTTTTGGATTCATCTTCTTTGAAAGT	GTTTGTACTGATTTCACCTCTGATTGGA
AMPL7153088440	GGCCTTTAGTTCTGTATGACATCTTCA	TTGAGCATCATCCTTTCCAAATATCATCA
AMPL7153088468	AGATATCTAGGGCCAAGACCCAA	TCTTTCTGTCTCTGAGAAGAGGCTT
AMPL7153088475	CTGCCAAGGCACGGCTAAAAGA	TGAGGATGGGAGAATTGTCCTGA
AMPL7153088483	GCAGCAGCCAATATAAAAAGCTATGT	GTTGACATTTGTTAACACTGTTGATTGC
AMPL7153088491	GCAAGGGAAGAGGGCTCTTCTA	CCACTTGTCATCATGTGATAATGGC
AMPL7153413892	GGATCAGCAAAACTAGTTTGTAATCTCTTG	GGTATACATTACAATGACCTGAGGAGATTT
AMPL7153413893	CCTTTATAGCTAGTTGACCCAAGTTGT	CACTGACCAATAAAACTGTGTTTGGTC
AMPL7153413898	CAAAAGAATGCAGTGTGGTGTCA	TGCTGTCCTTATATCTCCAACAAAGC
AMPL7153413901	GGATGTTGTTTCTGCTTGTGAAATGA	ATTCAGCTCCTGTCCCTCTGTA
AMPL7153413910	AGTAATTCCTGTTCCTCTCCACAGT	GCCCACCACACTAAACCTCATC
AMPL7153413916	GTGAAATCGAAGCAGGACTCCAT	GTTTTATTTCGGAGGCTTGCAGTTTAG
AMPL7153564403	GGTGTAACGAGGCAGGATGAAT	CCAAATTCACTGCAGTCATCCATG
AMPL7153572905	ACCAGCTGCAATTTTCCAATCAC	CTGAGCCTTTGCGAGAGGTAAT
AMPL7153609335	CGCGTGGTGTCTCGTTTG	AGGTCGGATCGGTCCGAA
AMPL7153670076	CGGATTCCTCAAAGTCGCTCAA	TCTCATCATTATCCTTGTTTCTCTCCTTCT
AMPL7153670489	GCTGCTTTTCTGACCCGTGTTA	CATCCAAAATTGATCCAGCTGAGAGA
AMPL7153670507	ACAACTTTGGGAATAAAGGAATTAGCCT	CTCATGTGGCTAGCACTGGAAT
AMPL7153670512	GATACCTGTTTGGATAACTCAGCACT	GTGTATCATTCTCTCAGTTTTGTCACCT
AMPL7153841029	CCAGTCCGCTCCTTCCTTTC	TAGTTCCCGGTCACCCAGTC
AMPL7153841036	TGGCCATGACAATGATGTCTGT	AAAAAGGCCAATGAGAACTCAACTC
AMPL7154029887	ATGAGCCGGTGACCAAGCT	CGGGCAGTCATTGTCACAG
AMPL7154030005	CGTAACCTCACAAATCCCTGCAT	GTAAGTGTAACATTCAGAACCGGGTAA
AMPL7155111390	GGATATCAAGAACTGGGTAGATGCT	GATATTTGCCCAAGATCACAAAGCAA
AMPL7155115749	CCCTGACACAGTTCTACCTTTGG	GTTTTCAGCAGTCCTCACAACTTTTAA
AMPL7155115750	CCAGATTCTCTCTGGCCCAAAA	TTTTAGAAAAGACAAGCAGTCATAGGAAGT
AMPL7155115752	CTTGGTCACTTGACCAAAAGCA	CCAGGAAGTTCATGGCAATATCTTCA
AMPL7155115761	GGTTGAATAGTCTTTTCAAGTGTCATTTGC	ACAATGGAGAAGAGGGTGATATAGTAGATT
AMPL7155115762	TGCCCTCATCCCAAGAATGAAG	CTAACAACCGGCACATCACGGAT
AMPL7155115763	CGAGTCCTCAAATGACTGGAATGTAG	GAGATGGCCATGAGCAGTTCATTA
AMPL7155115764	CGTCTCTGTCAGCAACACCATC	CTGAGGGCCACTCAAGTATCTC
AMPL7155116106	TGTCACCTACTTGGCTCGAGAA	GATTCTCGCTGTGACAGAGACA
AMPL7155116107	CCAGCCCAGACACTTACTTCTC	GGTACCTGACAGGGATAGGATCAG
AMPL7155116108	CCCGTTATGGACGTGATATAAGGC	TCCGACCCAACTTTGATGTTTCTATT
AMPL7155116109	CTCCTGTCCTGGGATGATCCTTA	GATACTTTAGACAGAGACCAGTTGTCAC
AMPL7155116110	CTCTGCACTTTGGATCTCAAATTGTG	AACTCCAGCATCTACAAAGGCA
AMPL7156300394	ATTTTGCTGTTATCTCTCAACCTCTTGA	CACGTTTCATAAGCTGCCAATGTT
AMPL7156300395	CAAAGATGACTTTCCTGAGAAGCTGA	TCATAAGAAATACAAGAACAGAGGTTGCA
AMPL7156300396	AGAACAAGAACCTAGAATGAATATCCAAGC	GTCTGGGTACCCTTTACATGATTCC
AMPL7156300397	GAAGAGAAGCGTGTTAGCCCAT	CATATCAAAAAGCCAAACCAAAGAGGAA
AMPL7156300398	GCTGAGTCAGGATCCTGTCAGT	ATGATTTGGGACCTTCTTACGCTT
AMPL7156300403	CGCATGTGGGCGATTTCTTTAA	ATCCGAGAGGCAAGGCAATCACT
AMPL7156300404	AAAAAGCTCCCGATACCCAATCAA	CGCTGCAGAAGGGAATAAAGAGA
AMPL7156300406	AGAAGCCAGTGAGGTGAGTTTG	AGACCAGGTCTCTATTTTGAGGAAGAA
AMPL7156684414	AGAAGGCTGAGAGAAGTGTATAAAGGA	GGTTATTTTGATCAAGTGCATCTCTTTGTT
AMPL7156685223	TAGTTGTTGTCTAGTAACTGACTCTTGTCT	CACATGAACAAAATGATCTTGAACCCA
AMPL7157032248	ACGCCAGGAGGAATGTCTGATA	CAACTTTTACGGGTACCAGCTAGA
AMPL7157032249	AACAGAGGGTCTGTACTAGCCAT	TCAACTGGGCTTAAAGAGGTATAATTCAC
AMPL7160389514	CCTACCGCTGTGTTCTTCTCTC	CTCTTGACCCAGGCAGGACACAT
AMPL7160389576	CTGCAGCCACGAGAAGCTTCAAC	CTTGCTTTCCAGCGCTTCATTAG
AMPL7160803748	CTGAGAGCAGGATGAAATAGCGTT	GTGCAACCCAAAGATGAAAGACCGA

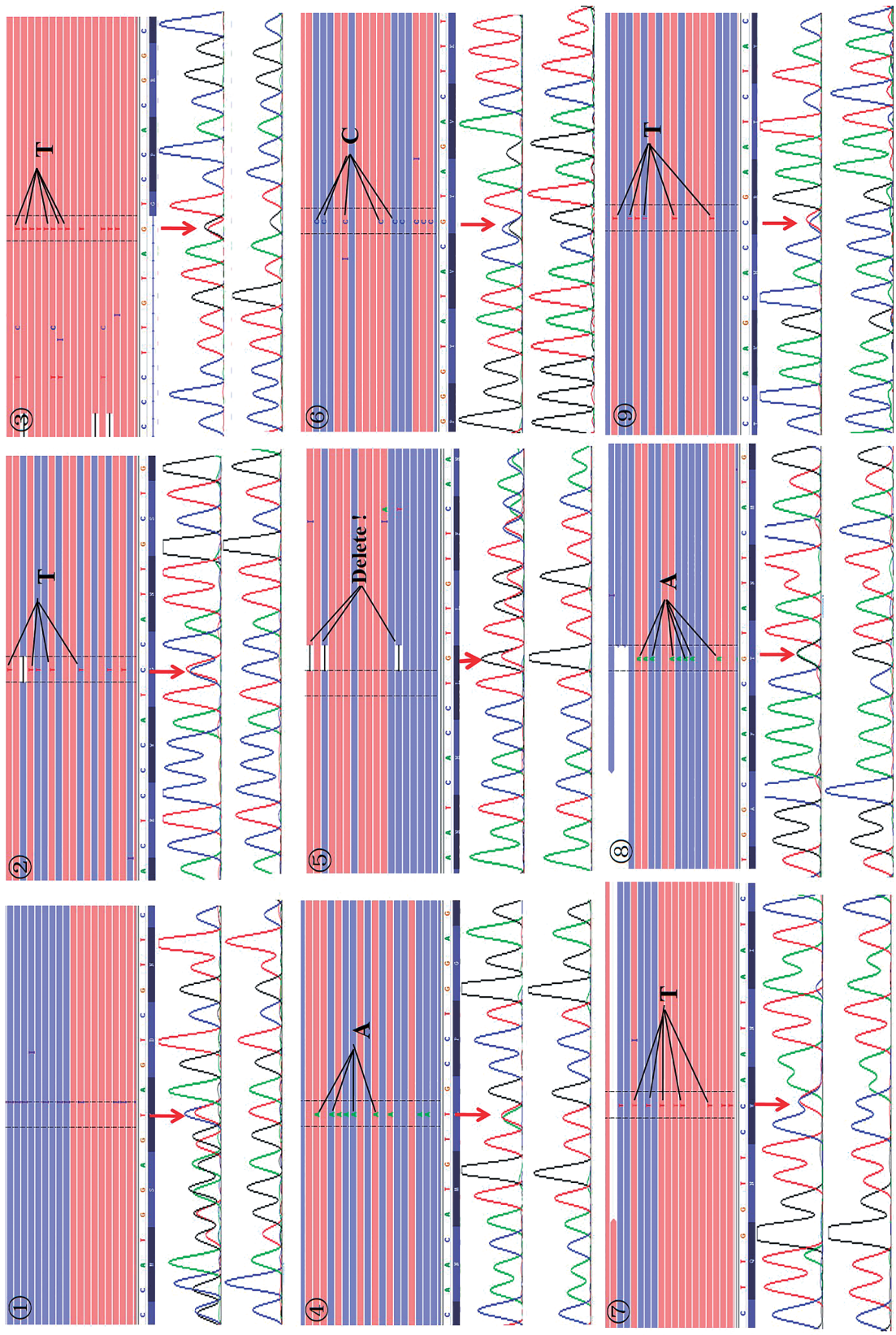

Fig. 1. Variants detected in *EXT1*/*EXT2* genes from 9 MO subjects.
①-⑤: Five variants identified by Sanger sequencing previously were also detected by t-NGS (*EXT1*: ① c.1843_1846dup, ② c.1088G>A; *EXT2*: ③ c.744-1G>T, ④ c.575T>A, ⑤ c.561delG). ⑥-⑨: Four variants detected by t-NGS were validated by Sanger sequencing (*EXT1*: ⑥ c.351C>G, ⑦ c.1121G>A, ⑧ c.2120C>T; *EXT2*: ⑨ c.67C>T). The top of the figure was a partial view from t-NGS. Blue bar indicated reverse sequencing, pink bar indicated forward sequencing, each bar was a read. The lower part was the corresponding results of Sanger sequencing, the red arrows indicated the positions of the variants (MT: mutant type, WT: wild type).

Table 3. Summary reports for 10 MO patients from t-NGS.

Subject	Mapped Reads	On Target	Mean Depth	Uniformity	1×Coverage	20×Coverage	100×Coverage
NC	25579	96.20%	418.6	97.84%	100%	100%	97.8%
1	23119	96.14%	391.9	97.25%	100%	97.85%	94.59%
2	37227	96.86%	639.3	97.83%	100%	99.99%	97.84%
3	28130	95.77%	474.9	96.37%	100%	99.89%	95.96%
4	117274	96.39%	1992	97.78%	100%	98.73%	97.84%
5	4912	87.68%	60.08	94.99%	100%	88.96%	16.3%
6	17293	88.21%	204.8	97.77%	100%	97.80%	82.80%
7	3254	91.79%	41.97	96.64%	100%	81.54%	2.31%
8	40308	96.69%	656.1	97.84%	100%	97.84%	97.84%
9	56511	96.18%	904.6	97.84%	100%	98.13%	97.84%
10	8088	87.93%	94.65	93.01%	100%	92.43%	39.15%

Table 4. Prediction of the pathogenicity for two missense mutations by bioinformatics.

nsSNP	PolyPhen		SIFT		PROVEAN	
	Prediction	Score	Prediction	Score	Prediction	Score
p.Gly363Glu	Probably damaging	0.998	Damaging	0.04	Deleterious	-4.37
p.Thr707Met	Possibly damaging	0.953	Tolerated	0.066	Neutral	-1.67

EXT1 and *EXT2* genes are composed of 11 exons and 15 exons, respectively (Fig. 2A). C.1088G>A/p.Gly363Glu, c.2120C>T/p.Thr707Met and c.744-1G>T were three newly discovered non-synonymous SNVs. PolyPhen, SIFT and PROVEAN predicted that c.1088G>A was probably a pathogenic mutation while c.2120C>T was possibly damaging (Table 4). Multiple sequence alignment of exostosin 1 showed that both amino acids in the sites of 363 and 707 were highly conserved in 18 different species (see Fig. 2C). Moreover, both exon 3 and exon 11 of *EXT1* gene in 100 normal controls were sequenced without related mutations. The CRYP-SKIP server predicted that the probability of cryptic splice site activation (*P*CR-E) was 0.7, while exon 5 skipping was 0.3 (1-*P*CR-E) due to c.744-1G>T in *EXT2* gene (Fig. 2B). HSF indicated that the altered acceptor site (atTC) was probably affecting splicing and its consensus value (0-100) dropped to 55.77 (agTC: 84.72). HSF also provided a potential acceptor site (agGT) at 31 base pairs downstream the wild-type acceptor site (agTC) with a consensus value of 90.89, MaxEntScan illustrated that the mutated-type acceptor site had a lower MAXENT (1.96) while the MAXENT of wild-type acceptor site was 10.56.

Discussion

Molecular diagnosis of MO is challenging. In previous studies, Sanger sequencing or denaturing high performance liquid chromatography (DHPLC) combined with Sanger sequencing of EXT genes was usually preferred. If there were no subtle mutations from Sanger sequencing, MLPA was performed to detect exon deletion/duplication, or array-CGH was used to search for some rare types of mutations, such as unbalanced translocations or mosaic deletions/duplications. As t-NGS is able to detect almost all types of mutations in MO except some unbalanced translocations, t-NGS is helpful to simplify the diagnostic process of MO. In our study, t-NGS approach was successfully applied in mutation screening for 10 MO patients; it is probably an alternative technique in routine molecular diagnosis for MO in the future.

After definite diagnosis, 10 MO patients were chosen for our study, nine of them were males and one female, their average onset age was close to 9 years old. Three patients had at least one affected parent, while seven were sporadic patients. The first group with known mutations was selected to evaluate the sensitivity and specificity of the t-NGS strategy compared to Sanger sequencing. Finally, the five known mutations were also detected by t-NGS, which illustrated the accuracy and sensitivity of t-NGS. Among these mutations, c.1088G>A and c.1843_1846dup were identified in *EXT1* gene; c.1088G>A is a missense mutation which results in glutamic acid instead of glycine at the 363 amino acid of exostosin 1, and c.1843_1846dup is a frameshift mutation which is expected to create a premature stop codon at the 616 amino acid of exostosin 1. C.561delG, c.575T>A and c.744-1G>T were found in *EXT2* gene; c.561delG is a frameshift mutation which could create a premature stop codon at the 269 amino acid of exostosin 2, c.575T>A could also introduce a pre-

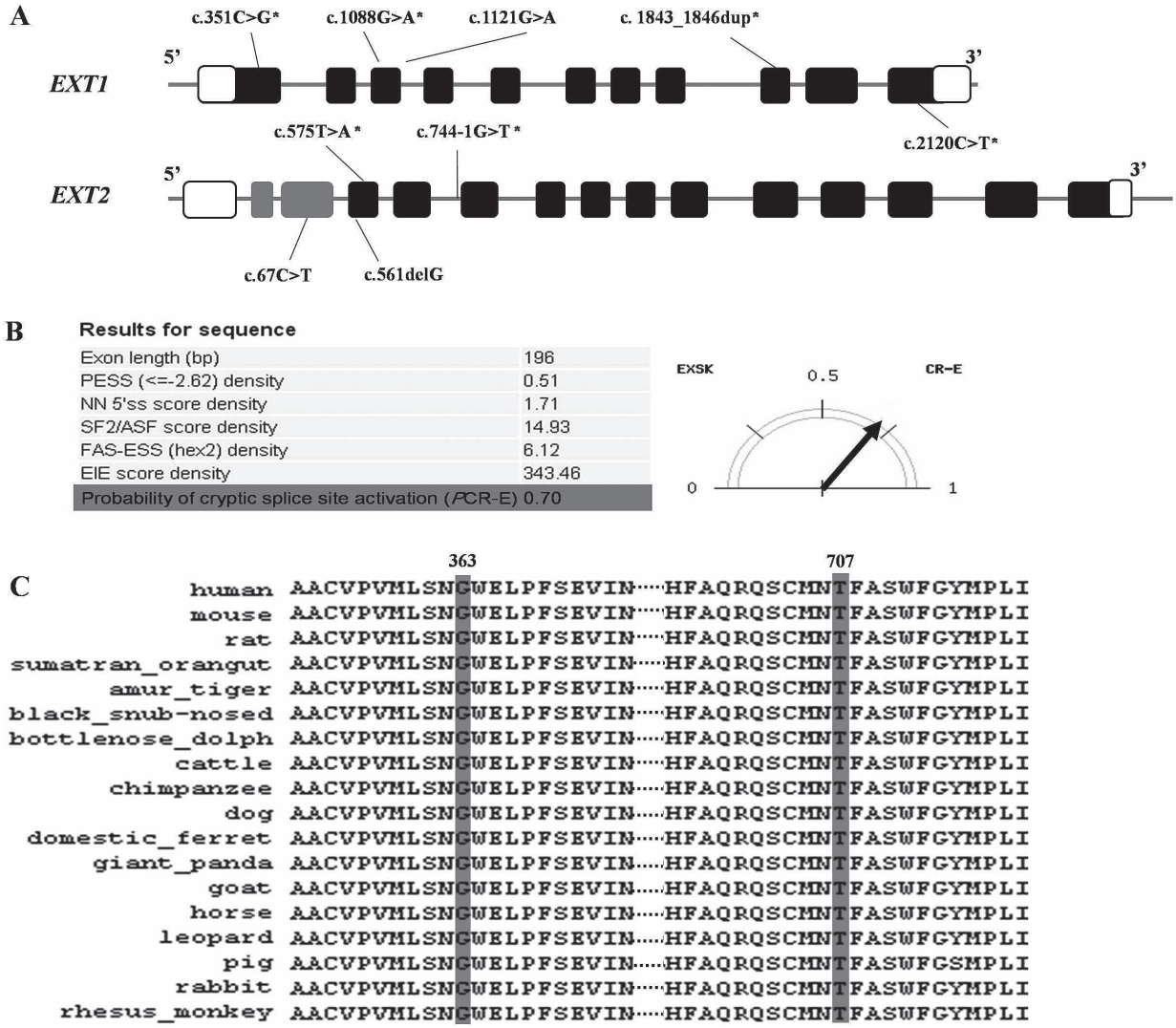

Fig. 2. Mutation spectrum of the *EXT1*/*EXT2* genes and pathogenicity analysis of three nonsynonymous mutations.
A: *Newly identified mutations in the MO probands in the present study, black boxes indicate the constitutive exons of *EXT1* and *EXT2* genes, gray boxes indicate the alternative exons of *EXT2* gene; B: Prediction for c.744-1G>T in *EXT2* gene from CRYP-SKIP; C: Part results from multiple sequence alignment of exostosin 1 among eighteen species.

mature stop codon at the 192 amino acid of exostosin 2, and c.744-1G>T is a splice mutation that could cause skipping of exon 5 of *EXT2* gene or abnormal splicing due to the mutation (Fig. 2A, Table 1).

The second group was prospectively sequenced on PGM sequencer, and four probable pathogenic mutations were picked out, in which c.351C>G, c.1121G>A and c.2120C>T were from *EXT1* gene and c.67C>T from *EXT2* gene. All of them were point mutations and confirmed by Sanger sequencing. C.351C>G and c.1121G>A were both nonsense mutations which led to early appearance of stop codons at the 117 amino acid and the 374 amino acid of exostosin 1, respectively. c.2120C>T is a missense mutation which could replace threonine by methionine at the 707 amino acid of exostosin 1, and c.67C>T in *EXT2* gene is a nonsense mutation which could truncate exostosin 2 at the 23 amino acid. One patient (Subject 10) with negative result from t-NGS was also confirmed by Sanger sequencing (Fig. 1, Table 1). Meanwhile, three synonymous mutations were identified by t-NGS in exonic regions of *EXT1*/*EXT2* genes from six MO patients, and the results were consistent with Sanger sequencing. Considering that all the variants identified by Sanger sequencing were also detected by t-NGS in our study, t-NGS may be a better alternative mutation screening method in molecular diagnosis of MO.

However, two suspicious pathogenic mutations (c.59T>G, c.64delT) were detected in *EXT1* gene by t-NGS from three MO probands (Subjects 5, 6, 10), but they were determined to be false positive by Sanger sequencing. Thus, Ion Torrent PGM had limited accuracy especially in case of homopolymers of any length (Loman et al. 2012). Subject 10 was discovered with no pathogenic mutation by t-NGS, and the subsequent DNA Sanger sequencing confirmed the negative result. The unsuccessful detection may

exist in both methods, especially if there is exon deletion/duplication, unbalanced translocation, or mosaic mutation in subject 10. We were regretful that we had not verified the result through cDNA Sanger sequencing, MLPA, or array-CGH as recommended (Szuhai et al. 2011; Sabbagh et al. 2013). However, the negative patient shows a sporadic form of moderate MO phenotype (Table 1) (Francannet et al. 2001); there may be a low-abundance mosaic mutation in the patient, which may explain the undetectable somatic EXT mutation from peripheral blood. However, t-NGS is helpful to detect mosaic mutation (Pasmant et al. 2015).

Meanwhile, 6 mutations were newly identified in our study (*EXT1*: c.1843_1846dup, c.1088G>A, c.351C>G, and c.2120C>T; and *EXT2*: c.744-1G>T and c.575T>A), while the remaining 3 mutations had been reported previously (Multiple Osteochondroma Mutation Database) (Fig. 2A). C.1843_1846dup, c.351C>G and c.575T>A are all nonsense mutations which probably cause the transcripts with early stop codons to activate the mechanism of RNA surveillance and nonsense-mediated mRNA decay (Gerards et al. 2010). Thus, the functional exostosin 1 and exostosin 2 would be reduced. C.1088G>A and c.2120C>T are missense mutations, prediction of computer and bioinformatics, as well as polymorphism analysis of both mutation sites suggest that the two new mutations were probably damaging and responsible for the occurrence of MO (Fig. 2C, Table 4). C.744-1G>T in *EXT2* gene was a splice mutation, both HSF and MaxEntScan indicated that the splicing probability of 3′ splice site of exon 5 would decrease due to the mutation, and a potential splice site could be activated according to the prediction of HSF and the CRYP-SKIP server, moreover, if the potential acceptor site (agGT) at the 31 base pairs downstream the wild-type acceptor site is activated, exon 5 of *EXT2* gene will be aberrantly spliced and a premature termination codon is to appear at 31 base pairs downstream the potential acceptor site of the abnormally spliced exon 5, then exostosin 2 is truncated and functional protein is also reduced.

In short, we applied t-NGS in molecular diagnosis of MO for the first time, and we successfully identified 9 mutations in *EXT1*/*EXT2* genes from 10 MO subjects. Among the nine mutations, six mutations were newly discovered, which further enriched the mutation database of MO from the Chinese population.

Acknowledgments

Supported by the Science and Technology Project of Fuzhou (2015-S-141-14) and Youth Research Project of Health and Family Planning in Fujian Province (2016-2-41).

Conflict of Interest

The authors declare no conflict of interest.

References

Ahn, J., Ludecke, H.J., Lindow, S., Horton, W.A., Lee, B., Wagner, M.J., Horsthemke, B. & Wells, D.E. (1995) Cloning of the putative tumour suppressor gene for hereditary multiple exostoses (*EXT1*). *Nat. Genet*., **11**, 137-143.

Aouini, F., Garali, W., Saaidi, A., El Mahdi, A., Mechergui, S., Jabeur, C., Ben Romdhane, N. & Manaa, J. (2015) Nerve and deep vein compression by femoral artery pseudoaneurysm in a patient with multiple exostosis. *Ann. Cardiol. Angeiol*., **64**, 113-115.

Bozzola, M., Gertosio, C., Gnoli, M., Baronio, F., Pedrini, E., Meazza, C. & Sangiorgi, L. (2015) Hereditary multiple exostoses and solitary osteochondroma associated with growth hormone deficiency: to treat or not to treat? *Ital. J. Pediatr*., **4**, 41-53.

Cheung, P.K., McCormick, C., Crawford, B.E., Esko, J.D., Tufaro, F. & Duncan, G. (2001) Etiological point mutations in the hereditary multiple exostoses gene *EXT1*: a functional analysis of heparan sulfate polymerase activity. *Am. J. Hum.Genet*., **69**, 55-66.

Choi, Y. & Chan, A.P. (2015) PROVEAN web server: a tool to predict the functional effect of amino acid substitutions andindels. *Bioinformatics*, **31**, 2745-2747.

Esko, J.D. & Lindahl, U. (2001) Molecular diversity of heparan sulfate. *J. Clin. Invest*., **108**, 169-173.

Francannet, C., Cohen-Tanugi, A., Le Merrer, M., Munnich, A., Bonaventure, J. & Legeai-Mallet, L. (2001) Genotype-phenotype correlation in hereditary multiple exostoses. *J. Med. Genet*., **38**, 430-434.

Gerards, M., van den Bosch, B., Calis, C., Schoonderwoerd, K., van Engelen, K., Tijssen, M., de Coo, R., van der Kooi, A. & Smeets, H. (2010) Nonsense mutations in *CABC1/ADCK3* cause progressive cerebellar ataxia and atrophy. *Mitochondrion*, **10**, 510-515.

Jennes, I., de Jong, D., Mees, K., Hogendoorn, P.C., Szuhai, K. & Wuyts, W. (2011) Breakpoint characterization of large deletions in *EXT1* or *EXT2* in 10 multiple osteochondromas families. *BMC Med. Genet*., **12**, 85.

Jennes, I., Entius, M.M., Van Hul, E., Parra, A., Sangiorgi, L. & Wuyts, W. (2008) Mutation screening of *EXT1* and *EXT2* by denaturing high-performance liquid chromatography, direct sequencing analysis, fluorescence in situ hybridization, and a new multiplex ligation-dependent probe amplification probe set in patients with multiple osteochondromas. *J. Mol. Diagn*., **10**, 85-92.

Jennes, I., Pedrini, E., Zuntini, M., Mordenti, M., Balkassmi, S., Asteggiano, C.G., Casey, B., Bakker, B., Sangiorgi, L. & Wuyts, W. (2009) Multiple osteochondromas: mutation update and description of the multiple osteochondromas mutation database (MOdb). *Hum. Mutat*., **30**, 1620-1627.

Jennes, I., Zuntini, M., Mees, K., Palagani, A., Pedrini, E., De Cock, G., Fransen, E., Vanden Berghe, W., Sangiorgi, L. & Wuyts, W. (2012) Identification and functional characterization of the human *EXT1* promoter region. *Gene*, **492**, 148-159.

Liu, H., Wu, S., Duan, L., Zhu, W., Zhang, S., Hu, X., Jia, W., Yang, G., Liu, C., Li, W., Yang, L., Guo, L., Lin, Y., Wang, Y., He, M., et al. (2015) Identification of a novel *EXT1* mutation in patients with hereditary multiple exostosis by exome sequencing. *Oncol. Rep*., **33**, 547-552.

Loman, N.J., Misra, R.V., Dallman, T.J., Constantinidou, C., Gharbia, S.E., Wain, J. & Pallen, M.J. (2012) Performance comparison of benchtop high-throughput sequencing platforms. *Nat. Biotechnol*., **30**, 434-439.

Ng, P.C. & Henikoff, S. (2003) SIFT: predicting amino acid changes that affect protein function. *Nucleic. Acids Res*., **31**, 3812-3814.

Okubo, M., Minami, N., Goto, K., Goto, Y., Noguchi, S., Mitsuhashi, S. & Nishino, I. (2016) Genetic diagnosis of Duchenne/Becker muscular dystrophy using next-generation sequencing: validation analysis of *DMD* mutations. *J. Hum.*

Genet., **61**, 483-489.

Pasmant, E., Parfait, B., Luscan, A., Goussard, P., Briand-Suleau, A., Laurendeau, I., Fouveaut, C., Leroy, C., Montadert, A., Wolkenstein, P., Vidaud, M. & Vidaud, D. (2015) 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*., **23**, 596-601.

Ramensky, V., Bork, P. & Sunyaev, S. (2002) Human non-synonymous SNPs: server and survey. *Nucleic. Acids Res*., **30**, 3894-3900.

Sabbagh, A., Pasmant, E., Imbard, A., Luscan, A., Soares, M., Blanche, H., Laurendeau, I., Ferkal, S., Vidaud, M., Pinson, S., Bellanne-Chantelot, C., Vidaud, D., Parfait, B. & Wolkenstein, P. (2013) NF1 molecular characterization and neurofibromatosis type I genotype-phenotype correlation: the French experience. *Hum. Mutat*., **34**, 1510-1518.

Schmale, G.A., Conrad, E.U. 3rd. & Raskind, W.H. (1994) The natural history of hereditary multiple exostoses. *J. Bone Joint Surg. Am*., **76**, 986-992.

Senay, C., Lind, T., Muguruma, K., Tone, Y., Kitagawa, H., Sugahara, K., Lidholt, K., Lindahl, U. & Kusche-Gullberg, M.(2000) The EXT1/EXT2 tumor suppressors: catalytic activities and role in heparan sulfate biosynthesis. *EMBO Rep*., **1**, 282-286.

Stickens, D., Clines, G., Burbee, D., Ramos, P., Thomas, S., Hogue, D., Hecht, J.T., Lovett, M. & Evans, G.A. (1996) The *EXT2* multiple exostoses gene defines a family of putative tumour suppressor genes. *Nat. Genet*., **14**, 25-32.

Szuhai, K., Jennes, I., de Jong, D., Bovee, J.V., Wiweger, M., Wuyts, W. & Hogendoorn, P.C. (2011) Tiling resolution array-CGH shows that somatic mosaic deletion of the *EXT* gene is causative in *EXT* gene mutation negative multiple osteochondromas patients. *Hum. Mutat*., **32**, E2036-E2049.

Tian, C., Yan, R., Wen, S., Li, X., Li, T., Cai, Z., Du, H. & Chen, H. (2014) A splice mutation and mRNA decay of EXT2 provoke hereditary multiple exostoses. *PLoS One*, **9**, e94848.

Vink, G.R., White, S.J., Gabelic, S., Hogendoorn, P.C., Breuning, M.H. & Bakker, E. (2005) Mutation screening of EXT1 and EXT2 by direct sequence analysis and MLPA in patients with multiple osteochondromas: splice site mutations and exonic deletions account for more than half of the mutations. *Eur. J. Hum. Genet*., **13**, 470-474.

White, S.J., Vink, G.R., Kriek, M., Wuyts, W., Schouten, J., Bakker, B., Breuning, M.H. & den Dunnen, J.T. (2004) Two-color multiplex ligation-dependent probe amplification: detecting genomic rearrangements in hereditary multiple exostoses. *Hum. Mutat*., **24**, 86-92.

Wicklund, C.L., Pauli, R.M., Johnston, D. &. Hecht, J.T. (1995) Natural history study of hereditary multiple exostoses. *Am. J. Med. Genet*., **55**, 43-46.

Wuyts, W., Radersma, R., Storm, K. & Vits, L. (2005) An optimized DHPLC protocol for molecular testing of the *EXT1* and *EXT2* genes in hereditary multiple osteochondromas. *Clin. Genet*., **68**, 542-547.

Wuyts, W., Van Hul, W., Wauters, J., Nemtsova, M., Reyniers, E., Van Hul, E.V., De Boulle, K., de Vries, B.B., Hendrickx, J., Herrygers, I., Bossuyt, P., Balemans, W., Fransen, E., Vits, L., Coucke, P., et al. (1996) Positional cloning of a gene involved in hereditary multiple exostoses. *Hum. Mol. Genet*., **5**, 1547-1557.

Zhang, F., Liang, J., Guo, X., Zhang, Y., Wen, Y., Li, Q., Zhang, Z., Ma, W., Dai, L., Liu, X., Yang, L. & Wang, J. (2013) Exome sequencing and functional analysis identifies a novel mutation in *EXT1* gene that causes multiple osteochondromas. *PLoS One*, **8**, e72316.