Supplemental information: Methods

Next-generation sequencing (NGS) method. Molecular profiling of both the tissue biopsy specimen and peripheral blood sample was performed to identify any pathogenic somatic and germline mutations with potential clinical significance. The samples were transported to MedGenome Labs Ltd. (Bengaluru, India) for performing a laboratory-developed College of American Pathologists (CAP)-accredited NGS assay, data analysis, interpretation and clinical report generation. MedGenome's TumorFocus panel and comprehensive hereditary cancer gene panel were processed for somatic and germline mutation calling, respectively. The TumorFocus panels are a multigene tumor agnostic panel enabling detection of single nucleotide variations (SNVs), small insertions and deletions (InDels), copy number variations (CNVs) and gene fusions in 76 genes of diagnostic, prognostic, and therapeutic significance, curated as per the National Comprehensive Cancer Network (NCCN) guidelines. The hereditary cancer gene panel covers complete coding segments (CDS) of 143 clinically significant genes, promoter regions of relevant genes and other critical non-coding/coding pathogenic variants <100 bp, as documented in clinical databases for the detection of SNVs, InDels and CNVs, curated as per the American College of Medical Genetics and Genomics (ACMG) and NCCN guidelines.

Histopathologic assessment was done on tissue sections (3-4 μm) prepared from the formalin-fixed paraffin-embedded (FFPE) tissue block and the tumor content was determined after Hematoxylin and Eosin (H&E) staining. The block with >30% tumor content was selected for genomic DNA and RNA extraction (QIAmp DNA FFPE Tissue Kit; cat. no. 56404; and RNeasy FFPE Kit; cat. no. 73504; both from Qiagen GmbH). The QIAsymphony DNA mini-Kit (cat. no. 931236; Qiagen GmbH) was used for DNA extraction from the PB sample. The extracted nucleic acid was subjected to a quality check for concentration using Qubit (Thermo Fisher Scientific, Inc.) and Tape station (Agilent Technologies, Inc.). Approximately 100-200 ng of the DNA/RNA was used for targeted library preparation using a custom hybrid capture kit (Twist Library Prep Kit MF with AMP Mix; cat. no. 100876; Twist Bioscience). RNA was converted to cDNA using NEBNext Ultra II RNA Library Prep Kit (cat. nos. E7771L and E6111L, for 'NEBNext Ultra II RNA First Strand Synthesis Module' and 'NEBNext Ultra II Non-Directional RNA Second Strand Synthesis Module', respectively; New England Biolabs, Inc.). The sequencing library quality was checked using Qubit (Thermo Fisher Scientific, Inc.) and Tape station (Agilent Technologies, Inc.). The quality-passed library was loaded at a concentration of 120 pM, and sequenced as 2x150 bp paired-end reads on an Illumina NovaSeq 6000 System (Illumina, Inc.) according to the manufacturer's protocol using the NovaSeq 6000 S4 Reagent Kit v1.5 (300 cycles; cat. no. 20028312) to an average sequencing depth of around 250x for somatic and >80-100X for germline mutation analysis. The sequenced data were processed using a well-validated bioinformatics pipeline. Variants were annotated using an in-house pipeline with the VEP program against the Ensembl release 99 human gene model. NGS analysis quality metrics checks were performed as per the CAP guidelines. Only non-synonymous and splice site variants in coding regions were used for clinical interpretation and gene fusions with both established and rare fusion partners were identified using a combination of validated and published fusion detection algorithms. Clinically relevant mutations were cross-referenced with published variant data in literature and databases like ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/), OncoKB (https://www.oncokb.org/), HGMD (https://www.hgmd.cf.ac.uk/ac/index.php), BRCA Exchange (https://brcaexchange.org/) and LOVD (https://www.lovd.nl/). Common variants were filtered based on minor allele frequencies from 1000Genome Phase 3, ExAC, gnomAD and an internal MedGenome database, which consists of >2 million high quality germline variants in the coding region across 20,000+ genes from >100,000 South Asian individuals. This database helps identify common population variants and is useful for variant prioritization in rare diseases and cancer database. The biological effect of variants was assessed using prediction algorithms such as PolyPhen, SIFT, Mutation Taster2 and LRT. The annotated variants were uploaded into a proprietary CAP-validated variant interpretation and reporting clinical diagnostic software 'Oncominer' developed in-house and visualized on Integrative Genomics Viewer. MedGenome has independently developed 'Oncominer', a proprietary, end-to-end, CAP-validated software for somatic variant analysis, interpretation and clinical reporting. This platform is used routinely for MedGenome's data analysis and reporting of >200,000 oncology samples to date, and its use has been documented in the published medical literature (1). For somatic mutation, variants with a minimum variant allele frequency (VAF) of 3-5% and fusions with >10 junction reads were prioritized and classified according to AMP/ASCO/CAP [PMID:27993330] and NCCN guidelines. In case of germline data analysis, clinically significant variants with a minimum VAF of 15-18% were prioritized and classified according to the ACMG [PMID: 27854360] and NCCN guidelines.

About 30-50 ng of the genomic DNA was extracted from the PB sample and subjected to a Digital multiplex ligation-dependent probe amplification (dMLPA) assay according to the manufacturer's protocol (SALSA® digitalMLPA™ Probemix D001-Hereditary Cancer Panel 1; cat. no. D001; MRC Holland BV). Raw data were analyzed using Coffalyser dMLPA software (MRC Holland BV) in combination with the corresponding LOT-specific product sheet to identify copy number variations, i.e., large deletions and duplications.

H&E staining and immunohistochemistry (IHC). H&E staining was performed at MedGenome Labs Ltd. on a Leica Autostainer XL ST-5010 following an in-house protocol. IHC was done by a manual method using heat-induced epitope retrieval and using the secondary antibody PolyExcel HRP/DAB Detection System, cat no. IPS 007, Ready To Use (RTU), by PathnSitu Biotechnologies. Fixative used: 10% neutral buffered formalin at room temperature; duration, 36-48 h. Thickness of the sections: 3 microns. Stain used: H&E stain. Microscope type: Light microscope, Olympus CX23 (Olympus Corp.). Magnification: 10x10.25 and 40x10.65 objective.

Antibodies used:

Cytokeratin (CK)7: Cat. no. Z2026; 1:200 dilution, by GenomeMe

CK19: Cat. no. IHC019-7; RTU, by GenomeMe
CEA: Cat. no. IHC543-7; RTU, by GenomeMe
Caudal-type homeobox 2: Cat. no. IHC302-7; RTU, by GenomeMe
Mucin (MUC)2: Cat. no. MR1367; RTU, by PathnSitu Biotechnologies
MUC5AC: Cat. no. PM233; RTU, by PathnSitu Biotechnologies
Hepatocyte paraffin 1: Cat. no. PM115; RTU, by PathnSitu Biotechnologies
Arginase 1: Cat. no. PM230; RTU, by PathnSitu Biotechnologies
Glypican 3: Cat. no. MM1179; RTU, by PathnSitu Biotechnologies
MutL Homolog (MLH) 1: Cat. no. CM098; 1:200 dilution, by PathnSitu
Biotechnologies
MLH2: Cat. no. CR055; 1:200 dilution, by PathnSitu Biotechnologies
MLH6: Cat. no. CR056; 1:200 dilution, by PathnSitu Biotechnologies
Postmeiotic segregation increased 2: Cat. no. CR067; 1:200 dilution, by
PathnSitu Biotechnologies
Programmed cell death ligand 1: Cat. no. 790-4905; RTU, by Roche Ventana

Reference

1. Sreenath ND, Singh K, Rastogi S, Biswal N, Singh A, Barwad A, Mridha AR, Gamanagatti S, Shamim SA, Dharmashaktu Y, *et al*: Clinical impact of integrating RNA-based next-generation sequencing into the diagnostic evaluation of soft tissue sarcomas: Insights from a single-center multidisciplinary workflow. JCO Glob Oncol 12: e2500440, 2026.