Supplementary methods genomics

TGS

Targeted NGS were performed using Ion AmpliSeq™ workflow, combining multiple primer pairs in a multiplex PCR to generate libraries followed byPGM/Ion Torrent sequencing (ThermoFisher Scientific, Courtaboeuf, france). The variant calling is performed with Torrent Suite™ software, variantCaller (v4.x and higher; ThermoFisher Scientific) using GRCh37 (h19) reference.. Variants annotation were then annotated using dbsnp (v138) (http://www.ncbi.nlm.nih.gov/SNP), COSMIC (v69), and dbNSFP (V2.1), using SnpSift (v4.0E) [20]. Variant effects on the genome were inferred using SnpEff (v4.0E). Variants with an allele frequency of 5% or more, tumour read count ≥25 reads. Variants were then subsequently filtered on the basis of their genomic effect: only missense, non-sense, frameshifts, start/stop codon gain or loss, in-frame InDels and splice site variants and not known as common polymorphisms (<0.1% in 1000G and ESV database) were retained as variants of interest and are reviewed by molecular geneticist. Reported somatic genetic variants were compared to litterature dta by the by molecular geneticist and were classified as pathogenic variant (hotspot mutations, already described mutations with oncogenic effect in litterature or deleterious mutations in tumor suppression gene) ; unknown pathogenicity variant (rare variant in classical targetable gene, variant in fonctional domain with strong deleterious effect predict by various algorytm (sift/polyphen/splicing sites etc...) and probably non-pathogenic variant (variant without any information related to oncogenic or deleterious effect of mutation. The targetable genes were defined on the basis of the TARGETv2 list of actionable molecular abnormalities (TARGET_db_v2_05042014.xlsx) published by Van Allen *et al.* (Van Allen EM. Nat Med 2014; 20: 682-688) completed by litterature data. The sensitivity of the TGS analysis was validated with multiple run, repeated along the trials, with multiplex reference standard control (Horizon Diagnostics, Cambridge, UK), harboring several low allelic frequency mutations. More over the Gustave roussy genomic platform participate to a French national quality program for molecular analysis (Gen&Tiss ; *www.genetiss.org/*) covering proficiency testing of main mutation in standard targetable genes (BRAF/EGFR/KRAS/NRAS/PIK3CA..).
CP1 and CHP2 available on www.ampliseq.com
MOSC3 panel :

Gene	RefSeq	target exons
ABL1	NM_007313.2	4 to 7
AKT1	NM_005163.2	3;6
AKT2	NM_001626.3	3
AKT3	NM_005465.3	3
ALK	NM_004304.3	20 to 26
APC	NM_000038.5	16 (partial)
ATM	NM_00005.3	8;9;12;17;26;34 to 36;39;50;54;59;61;63
BRAF	NM_004333.4	11;15
BRCA1	NM_007294.3	2 to 23
BRCA2	NM_000059.3	2 to 27
CDH1	NM_004360.2	3;8;9
CDKN2A	NM_00077.4	2
CSF1R	NM_005211.3	7;22
CTNNB1	NM_NM	3
DDR2	NM_001014796	4 to 19
EGFR	NM_005228.3	3;7;12;15;18 to 21
ERBB2	NM_004448.2	8;19to21
ERBB3	NM_001982.3	1 to 28
ERBB4	NM_005235.2	3 to9;15;23
EZH2	NM_004456.4	16
FBXW7	NM_033632.2	2 to 12
FGFR1	NM_023110.2	4;7;12;14;15
FGFR2	NM_000141.4	7;9;12;14
FGFR3	NM_000142.4	7;9;14;16;18
FGFR4	NM_002011.3	2 to 18

Gene	RefSeq	target exons
FLT1	NM_002019.4	1 to 30
FLT3	NM_004119.2	11;14;16;20
GNA11	NM_002067.2	5
GNAQ	NM_002072.3	5;8
GNAS	NM_000516.4	8;9
HNF1A	NM_000545.5	3;4
HRAS	NM_005343.2	2 to 4
IDH1	NM_005896.2	4
IDH2	NM_002168.2	4
INPP4B	NM_003866.2	5 to 27
JAK2	NM_004972.3	14
JAK3	NM_000215.3	4;13;16
KDR	NM_002253.2	1 to 30
KEAP1	NM_203500.1	2 to 6
KIT	NM_000222.2	2;9to11;13;15;17;18
KRAS	NM_033360.2	2 to 4
MAP2K1	NM_002755.3	2;3
MAP2K4	NM_003010.2	1 to 11
MAP3K1	NM_005921.1	1 to 20
MET	NM_001127500.	2 (partial);11;14;16 to 19
MLH1	NM_000249.3	12
MLL3	NM_170606.2	8;9;43
MPL	NM_005373.2	1
MTOR	NM_004958.3	1 to 58
NF1	NM_001042492.	1 to 58

Gene	RefSeq	target exons
NFE2L2	NM_006164.3	2
NOTCH1	NM_017617.3	24;27;34
NOTCH2	NM_024408.	34
NOTCH4	NM_004557.3	1 to 30
NPM1	NM_002520.6	12
NRAS	NM_002524.3	2 to 4
PDGFRA	NM_006206.4	12;14;15;18
PIK3CA	NM_006218.2	2;5;7;8;10;14;19;21
PIK3R1	NM_181523.1	10;12;14;15
PPP2R1A	NM_014225.5	5;6
PTEN	NM_000314.4	1 to 9
PTPN11	NM_002834.3	3;13
RB1	NM_000321.2	4;6;10;11;14;17;18;20 to 22
RET	NM_020975.4	10;11;13;15;16
ROS1	NM_002944.	38
SMAD4	NM_005359.5	3to6;8to12
SMARCB1	NM_003073.2	2;4;5;9
SMO	NM_005631.4	3;5;6;9;11
SRC	NM_005417.3	14
STK11	NM_000455.4	1 to 9
TP53	NM_000546.4	1 to 11
TP53	NM_000546.4	2;4to8;10
TSC1	NM_000368.4	1 to 23
TSC2	NM_000548.3	2 to 42
VHL	NM_000551.3	1 to 3

CGH array

DNA was restriction digested and controlled by Agilent Bioanalyzer on DNA 7500 chips (Agilent Technologies, Santa Clara, CA, USA) and labelled with Cy3-dUTP or Cy5-dUTP using Agilent Genomic DNA Labelling Kit PLUS. Hybridization was carried out on Agilent 4x180kb arrays for 24 hours at 65°C in a rotating oven (Robbins Scientific, Mountain View, CA) at 20 rpm, according to the manufacturer's instructions. A commercial DNA was used for control (Promega). Scanning was performed with an Agilent G2505C DNA Microarray scanner using default parameters. Quantification of Cy5 and Cy3 signals from scans was performed with Feature Extraction v10.5.1.1 (Agilent Technologies) using default parameters. Resulting raw signals and log2 (ratio) profiles were normalized and centered according to their dye composition (Cy5/Cy3) and local GC content. These profiles were segmented with the Circular Binary Segmentation algorithm through its implementation in the DNA copy package for R v(v2.6 to v3.1) using default parameters [23]. DNA copy number imbalances were detected considering a minimum of 3 consecutive probes and a minimal absolute amplitude threshold that was specific for each profile. Manual reprocessing using the *undo.sd* parameter to a maximum value of 0.75 was performed in case for a few noisier profiles. Profiles were centered using the most centered out of the three most populated peaks of the smoothed log2(Test/Ref) distribution. Aberration levels were called by setting a log2(Test/Ref) threshold automatically adapted to the internal noise for each profile, considered as one-fourth of the median value of the absolute differences between consecutive log2(Test/Ref) measures along the genome. All genomics coordinates were established using the human genome as defined by the UCSC build hg19 (GRCh37). The copy number alterations detected with CGHa were transformed into log2 ratio. Gain were defined by x0.3 to x0.7 log2 ratio, amplifications >x0.7 log2 ratio, loss x0 to -0.5 log2 ratio and deletions < 0.5 log2 ratio. The alterations with a length less than 10 Mb were considered of interest.

WES

Base calling was performed using the Real-Time Analysis software sequence pipeline (RTA2) with default parameters. Sequence reads were mapped to the human genome build (hg19) using Elandv2e (Illumina, CASAVA-1.8.2) allowing multiseed and gapped alignments. The duplicated reads were removed. CASAVA-1.8.2 was used to call single-nucleotide variants (SNVs) and short insertions/deletions (max. size was 300nt), taking into account all reads per position. Indels with Q(Indel) < 20, or regions with low mappability (QVCutoff < 90) were filtered out. An IntegraGen algorithm, which compares normal and tumor genotypes from exome sequencing data, determines the somatic nature of the variation. A somatic score was calculated for each variant ranging from 1 to 30, a score of 30 translating the highest confidence index. This score takes into account the frequencies and counts of mutated allele in both samples to minimize false positive variations. Finally, variants displaying mutated reads in the constitutional sample above 5 percent are considered as germline or false positive. The somatic variant caller handles indels similarly, analyzing the number of alignments covering a given position that include a particular indel (the variant count) versus the overall coverage at that position. Variants annotation was based on dbSNP (dbsnp132), the 1000 Genomes Project (phase1_release_v3_20121010), the Exome Variant Server (ESP5400_snps), and an in-house database (201 exomes whole exomes for SNVs and 130 exomes whole exomes for indels). Other information like quality score, homozygote/heterozygote status, count of variant allele reads, mutation type (somatic or germline) and somatic score, the presence of the variant in the COSMIC database (v67_241013) were reported. Copy number aberrations (CNA) were identified using exomeCNV and the Bioconductor DNACopy package (DNAcopy 1.32.0) by comparing tumor from each sample to the normal DNA exome data.

RNAseq Analysis

Library preparation, capture, sequencing, and Bioinformatics analysis were performed by IntegraGen, Evry, France. Libraries were prepared using the TruSeq Stranded mRNA kit (Illumina) following manufacturer's instruction. Briefly, the TruSeq stranded mRNA sample prep kit converts the poly-A containing mRNA in total RNA (1000ng engaged in the process) into a cDNA library using poly-T oligo-attached magnetic bead selection. Following mRNA purification, the RNA was chemically fragmented prior to reverse transcription and cDNA generation. The fragmentation step resulted in an RNAseq library that included inserts ranging in size from approximately 400mers. The cDNA fragments then went through an end repair process, the addition of a single 'A' base to the 3' end and then ligation of the adapters. Finally, the products were purified and enriched with PCR to create the final double stranded cDNA library, which was then purified and quantified by QPCR. Each transcriptome library was sequenced on an Illumina NextSeq 500 as paired-end 75b reads.

RNAseq base calling was performed using the Real-Time Analysis software sequence pipeline (RTA2) with default parameters. Sequence reads were mapped to the human genome build (hg19) using TopHat2.1.0, including Bowtie2. Single Nucleotides Variants (SNV) were detected with Samtools/BcfTools (Broad Institute).

IHC –MET

MET staining was previouspy described in Lacroix et.al ((Lacroix L et al. PLoS One. 2014 Jan 17;9(1PubMed PMID: 24465403)). Briefly, we performed MET and p-MET immunostaining using serial tissue sections. Briefly, antigen retrieval was performed with Cell Conditioning 1 (CC1) buffer at 95°C for 8 minutes, and then at 100°C for 28 minutes. After the endogen biotins blocking step, slides were incubated with the primary anti-antibody for one hour at 24°C for the rabbit anti-human c-MET (final dilution 1/50, clone SP44, reference M3444, Spring Bioscience, USA or clone CVD13, reference 18-2257, Invitrogen, USA) and at 37°C for anti-p-MET (Tyr1234/1235) (final dilution 1/50, clone D26, reference 3077, Cell Signaling Technologies, USA). A post-fixation step with glutaraldehyde (0.05% in NaCl 0.9% w/v) for 4 minutes at 24°C was done. For MET detection, the secondary antibody biotin-SP-conjugated Affinipure goat anti-rabbit (reference 111-065-003, batch 84328, Jackson ImmunoResearch Laboratories, Inc, USA) was incubated at 24°C for 32 minutes at 0.5 microg/mL. For p-MET, the secondary antibody biotin free peroxidase multimer anti-rabbit UltraMap™ was incubated at 24°C for 16 minutes. Immunostaining was done with 3,3-diaminobenzidine tetrahydrochloride (DAB) from DABMap™ chromogenic detection kit according to manufacturer's recommendations. A counter-staining step was done with hematoxylin and blueing reagent was applied. Stained slides were dehydrated and coverslipped with cytoseal XYL (8312-4, Richard-Allan Scientific, USA).