Seshat output (Short)

cDNA_Variant	Mutation nomenclature according to HGVS standards using the coding sequence as reference (position 1 refers to the A of the start ATG): reference sequence NM_000546.5
HG19_Variant	Mutation nomenclature according to HGVS standards using the genomic sequence as reference Reference sequence: NC_000017.10 for genome build NCBI37/hg19
NG_017013.2_variant	Mutation nomenclature according to HGVS standards using the RefSeq Gene NG_017013. sequence as reference http://www.ncbi.nlm.nih.gov/nuccore/NG_017013.2 This sequence is also the reference used by the Locus Reference Genomic (http://ftp.ebi.ac.uk/pub/databases/lrgex/LRG_321.xml
SNP_ID*	The SNP database now includes several pathogenic variants of the TP53 gene http://www.ncbi.nlm.nih.gov/snp * Note of caution: Since 2011 (build 134), dbSNP started accepting submissions of germ line and somatic variations associated with various types of diseases and changed its name to "database of Short Genetic Variation" keeping the dbSNP acronym. Several frequent *TP53* variants (rs121912651, c.742C>T, p.Arg248Trp or rs11540652, c.743G>A, p.Arg248Gln) are included in dbSNP, but other hot spot variants are missing, whereas rare somatic variants may be included. This heterogeneity caused by biased dbSNP submissions is misleading, as it does not reflect the true occurrence and frequencies of *TP53* variants. Therefore, without further distinction, we can no longer assume that variants in dbSNP are associated with the lack of effect on disease and tumour characteristics Common SNPs such as rs1042522 (p.P72R), rs1800371 (p.P47S), rs1800372 (p.R213R) or rs1800370 (p.P36P) are not included in the database..
Transcript t1 MN_000546.5	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG (position 1 refers to the A of the start ATG): reference sequence NM_000546.5*
Transcript t2 NM_001126112.2	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG (position 1 refers to the A of the start ATG): reference sequence NM_001126112.2*
Transcript t3 NM_001126114.2	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG (position 1 refers to the A of the start ATG): reference sequence NM_001126114.2*
Transcript t4 NM_001126113.2	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG (position 1 refers to the A of the start ATG): reference sequence NM_001126113.2*
Transcript t5 NM_001126115.1	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG (position 1 refers to the A of the start ATG): reference sequence NM_001126115.1*
Transcript t6 NM_001126116.1	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG (position 1 refers to the A of the start ATG): reference sequence NM_001126116.1*
Transcript t7 NM_001126117.1	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG (position 1 refers to the A of the start ATG): reference sequence NM_001126117.1*
Transcript t8 NM_001126118.1	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG (position 1 refers to the A of the start ATG): reference sequence NM_001126118.1*

Protein p1 TP53_alpha	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG: reference sequence LRG_321p1*
Protein p3 TP53_beta	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG: reference sequence LRG_321p3*
Protein p4 TP53_gamma	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG: reference sequence LRG_321p4*
Protein p8 Delta40_TP53_alpha	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG: reference sequence LRG_321p8*
Protein p9 Delta 40_TP53_beta	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG: reference sequence LRG_321p9*
Protein p10 Delta 40_TP53_gamma	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG: reference sequence LRG_321p10*
Protein p5 Delta 133_TP53_alpha	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG: reference sequence LRG_321p5*
Protein p6 Delta 133_TP53_beta	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG: reference sequence LRG_321p6*
Protein p7 Delta 133_TP53_gamma	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG: reference sequence NP_001119589.1*
Protein p11 Delta160_TP53_alpha	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG: reference sequence NP_001263626.1*
Protein p12 Delta160_TP53_beta	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG: reference sequence NP_001263627.1*
Protein p13 Delta160_TP53_gamma	*Mutation nomenclature and coordinates are described according to HGVS, NCBI and LRG: reference sequence NP_001263628.1*
Records_Number	Number of occurrences of the mutant in the database.
Variant_Classification	Translational effect of the mutation (Missense, Nonsense, Synonymous, Nonstop, In_frame_Del, Inframe_Ins, Frameshift_Del or Frameshift_Ins). These 8 items are identical to the entries used in MAF file Four novel items are used in the TP53 mutation database: In_frame_Del_Complex, Inframe_Ins_Complex, Frameshift_Del_Complex or Frameshift_Ins_Complex) for mutations that span one or more than one exon-intron site.
Variant_Type	Variant type as defined in MAF file **SNV:** Single Nucleotide Variant **DNP:** Change in two consecutive bases (dinucleotide variant) **TNP:** Change in three consecutive bases (tri-nucleotide variant) **ONP:** Change in four or more consecutive bases (oligo-nucleotide variant) **INS:** Insertion **DEL:** Deletion
Comment_1_Frequency	Specific information related to the frequency of the mutation in the database. Four categories have been defined: i: This mutation is very frequent ii) This mutation is frequent iii: This mutation is not frequent iv: This mutation is rare see Leroy et al. TP53 Mutations in Human Cancer: Database Reassessment and Prospects for the Next Decade. Human Mutation (2014) 35, 672-688
Comment_2_Activity	Specific information related to the residual activity of this TP53 mutant in the

	database based on the overall transcriptional activity (TA) on 8 different promoters as measured by Kato et al. For each mutant, the median of the 8 promoter-specific activities (expressed as percent of the wild-type protein) has been calculated. For **missense variants,** five categories have been defined: • No activity: median <=20 • Partial activity: median >20 and <=75 • Fully active: median >75 and <=140 • Hyper active: median >140 • No data: this mutant has not been tested For **nonsense variants**, one category has been used The activity of truncated p53 is assumed to be nil For **frameshift** variants, two categories have been used: The consequence of this in-frame mutation is unknown (In-frame of 15 bp or less). The activity of truncated p53 is assumed to be nil (out-of-frame insertion and deletion, in-frame mutation >18 bp or mutation across an intron:exon junction). For **synonymous** variants, two categories have been used: This synonymous mutation is known to impair TP53 splicing. Synonymous mutation with unknown consequences. For mutations that target the canonical AG **splice-acceptor site** or GT **splice-donor site**: Splicing defect: impaired TP53 activity Activity for each individual promoter is also available (see the various rows in the database: WAF, MDM2, BAX, 14-3-3-s, AIP, GADD45, NOXA and P53R2).
Comment_3_Isoforms	Number of TP53 isoforms targeted by the mutation
Comment_4_Prediction	Several prediction algorithms have been used to predict TP53 loss of function (SIFT, Mutassessor, Provean, PolyPhen, see the corresponding rows for each individual analysis). A prediction index has been deduced from the various analyses Damaging Probably damaging Tolerated Note of caution: for TP53 mutation, the sensitivity of the various algorithms is never higher than 80% and the relation btween loss of function and pathogenicity is not straightforward.
Comment_5_Outliers	Indicates whether or not the mutation is associated with outlier publications. See Edlund K, Larsson O, Ameur A, Bunikis I, Gyllensten U, Leroy B, Sundstrom M, Micke P, Botling J, Soussi T. 2012. Data-driven unbiased curation of the TP53 tumor suppressor gene mutation database and validation by ultradeep sequencing of human tumors. Proc Natl Acad Sci USA 109:9551–9556. for more info on outliers studies. Rare mutants only found in outlier studies should be considered to be

	suspicious.
Comment_6_Splicing	Indicates whether or not the mutation could impair TP53 splicing. All TP53 gene substitutions have been analysed by using mutpred_splice M. Mort *et al.*, *Genome Biol* **15**, R19 (2014) courtesy of M. Mort A MutPred Splice general score probability cutoff of ≥0.70 was used to indicate a predicted SAV. For mutations close to an exon, a cutoff of ≥0.60 was used. Raw data for mutpred_splice are also available in this table.
Comment_7_Sequence	Indicates the presence of homopolymeric tracts at the position of the mutation.
Comment_9_SNP	Specific comments regarding the specificity of each SNP including novel SNP detected in new sequencing projects
Comment_10_population	Population data and frequency of the SNP in various databases.
Pathogenicity	We used this specific standard terminology for TP53 variants: 'pathogenic', 'likely pathogenic', 'uncertain significance' (VUS), 'likely benign'. A new TP53 specific algorithm was used to define TP53 variant pathogenicty (T Soussi et al. manuscript in preparation).
Final comment	Comment summary.
Mutalyzer_comment	Comment generated by mutalyzer when the mutation has not been correctly annoted.
CHROM	User data
POS	User data
ID	User data
REF	User data
ALT	User data