**Supplementary Information**

Recommended guidelines for validation, quality control and reporting of TP53 variants in clinical practice

Bernard Leroy, Mandy L. Ballinger, Fanny Baran-Marszak, Gareth L. Bond, Antony Braithwaite,, Nicole Concin, Lawrence A. Donehower, Wafik S. El-Deiry, Pierre Fenaux, Gianluca Gaidano, Anita Langerød, Eva Hellstrom-Lindberg, Richard Iggo, Jacqueline Lehmann-Che, Phuong L. Mai, David Malkin, Ute M. Moll, Jeffrey N. Myers, Kim E. Nichols, Sarka Pospisilova, Patricia Ashton-Prolla, Davide Rossi Sharon A. Savage, Louise C. Strong, Patricia N. Tonin, Robert Zeillinger, Thorsten Zenz,Joseph F. Fraumeni Jr., Peter E.M. Taschner,Pierre Hainaut and Thierry Soussi



**Figure S1**

***TP53* gene and p53 mRNAs**. The *TP53* gene (upper part of the figure) is transcribed into eight different mRNAs. Transcripts t1 to t4 originate from promoter P1 localized upstream from the gene. Transcripts t5 to t8 originate from promoter P2 localized in intron 4. Translated exons are shown in blue. The two novel exons β and γ are shown in red and blue, respectively. Untranslated regions are shown in black. For transcripts t3, t4, t6, and t7, which include exons β or γ, exons 10 and 11 are noncoding. Transcript t8 encodes only p8 (DeltaTP53α) and exons 1 to 3 are noncoding (black boxes). Proteins translated from the various transcripts are described on the right. Frequent *TP53* germline SNP are shown on the *TP53* gene with both coding (green) and non-coding (blue) snp.

This figure includes only the eight major *TP53* transcripts as reported in LRG\_321. Other transcripts have been identified, but it is currently unknown whether they are expressed in normal cells or in all cell types. (http://www.ensembl.org/Homo\_sapiens/Gene/Summary?db=core;g=ENSG00000141510;r=17:7661779-7687550)



**Figure S2. Correspondence between the LRG, the NCBI RefSeq TP53 protein and their corresponding transcripts**. All *TP53* transcripts (except one) have alternative translation start sites and can encode two TP53 isoforms. Each TP53 isoform has a single LRG\_321 p (protein) number, but might be encoded by more than one transcript with its own t (transcript) number. In contrast, full-length TP53 protein (p1) and the Delta40p53α isoform (p8) have multiple RefSeq protein accession numbers. As RefSeq transcripts can be linked to only one RefSeq protein accession number in the NCBI’s data model, a different RefSeq transcript accession number is therefore assigned to the same mRNA sequence for each TP53 isoform it encodes.



**Figure S3. Correspondence between the LRG, the NCBI RefSeq *TP53* transcripts and their corresponding putative proteins.**