

SUPPLEMENTAL METHODS

Sample processing

Peripheral blood (PB) leukocytes were isolated by erythrolysis with ammonium chloride. Mononuclear (MNC) and granulocyte fractions were separated using density gradient centrifugation (Histopaque 1077 and 1119, Sigma-Aldrich). MNC were vitally frozen in fetal bovine serum containing 10% dimethyl sulfoxide (Sigma-Aldrich) and stored at -70°C. Samples taken into PAXgene Blood RNA tubes (PreAnalytiX) were processed according to manufacturer's recommendations.

TP53 Sanger sequencing

Sanger sequencing of the *TP53* gene was performed according to IARC *TP53* Database Protocol (<http://p53.iarc.fr/ProtocolsAndTools.aspx>)¹.

Colony-forming cell assay

The colony forming cell assay was performed from vitally frozen PB mononuclear cells using methylcellulose medium (MethoCult H4434 Classic, StemCell Technologies, Vancouver, Canada) according to manufacturer's protocol with modifications². Single colonies were picked, *TP53* was Sanger-sequenced and JAK2 and CALR were analysed by fragmentation analysis as described previously³.

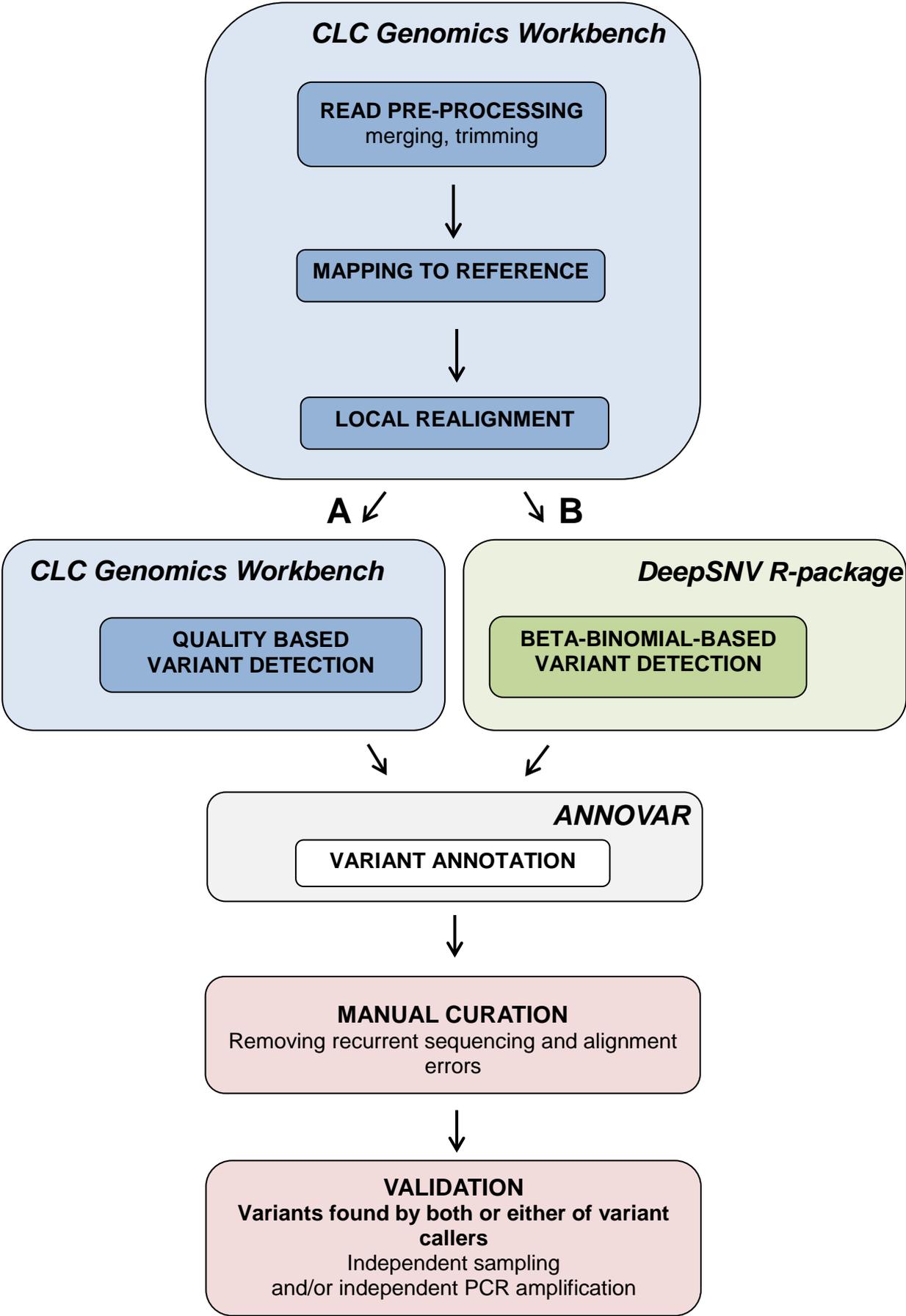
SNP arrays analysis

CEL files were analyzed using the Chromosome Analysis Suite software (Affymetrix), v3.1.0.15 and annotated using NetAffx 33.1 annotation dataset. All chromosomes were then manually checked in order to identify pathogenic aberrations with frequency down ~10% (as assessed based on a cohort of 100 hematological patients with available FISH results) and to exclude genome structural variations detected in the normal population (collected in

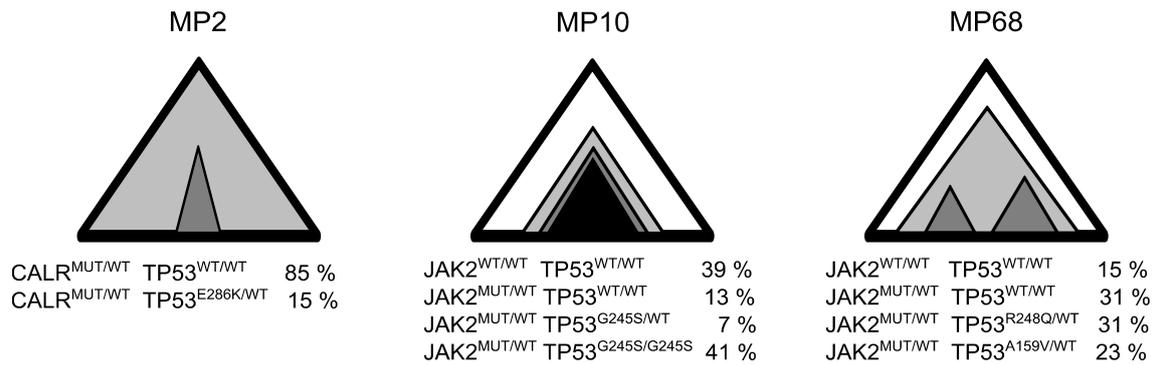
the Database of Genome Variants and Affymetrix CytoScan-specific dataset). Special attention was paid to cancer-associated regions.

1. Bouaoun L, Sonkin D, Ardin M, et al: *TP53* Variations in Human Cancers: New Lessons from the IARC *TP53* Database and Genomics Data. *Hum Mutat* 37:865-76, 2016
2. Olcaydu D, Harutyunyan A, Jäger R, et al: A common JAK2 haplotype confers susceptibility to myeloproliferative neoplasms. *Nat Genet* 41:450-4, 2009
3. Klampfl T, Gisslinger H, Harutyunyan AS, et al: Somatic mutations of calreticulin in myeloproliferative neoplasms. *N Engl J Med* 369:2379-90, 2013

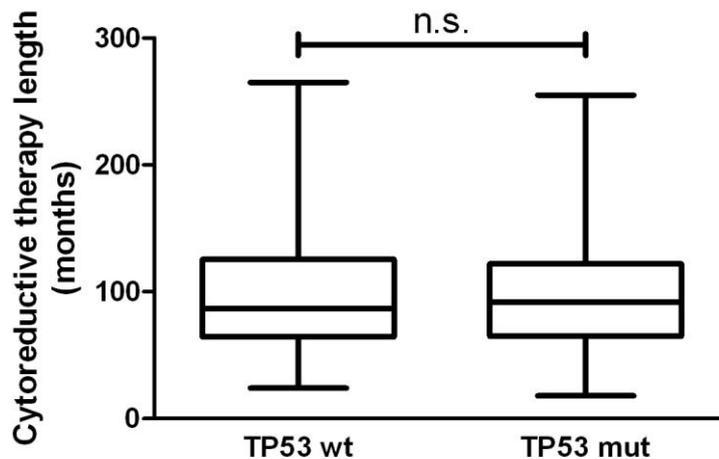
Supplemental Figure S1



Supplemental Figure S1: Scheme of bioinformatics pipeline used to call variants within NGS data. Each sample (represented by paired-end sequencing reads with length of 2x150 bp) was analyzed using CLC Genomics Workbench version 7.5 software (<http://www.clcbio.com>; Qiagen). Overlapping paired-end reads were merged and low quality ends were trimmed. Read alignment was performed on human reference genome (Genome Reference Consortium Human Build 37 patch release 9). Following local realignment, two independent algorithms for variant detection were applied: (A) Quality based variant detection using CLC Genomics Workbench version 7.5. Only high quality *TP53* variants were considered (quality score of $Q \geq 35$ guaranteeing more than 99.9% base call accuracy). Details on CLC settings are available upon request. The ANNOVAR program was used to annotate variants with gene and exonic function, exon number, position in cDNA and amino acid change (RefSeq) using reference sequence NM_000546. Exon and splice site variants with variant allelic frequency $\geq 0.2\%$ and variant read count ≥ 10 were selected. Recurrent sequencing and alignment errors and common exon polymorphisms were removed. The remaining variants were considered as mutations. (B) Read mappings were exported as BAM files and further analyzed using freeware R (deepSNV R-package). The shearwater algorithm from deepSNV package was used to compute Bayes classifier based on betabinomial model for variant calling with prior knowledge. Default settings of function “bbb” were applied, and priors obtained from COSMIC v.67 database. As a compound control sample we used samples from patients in whom no mutation was identified using CLC-based algorithm plus 20 control samples from young donors. Variants with Bayes factor posterior probability $p \leq 0.01$ were taken into account. Variants were annotated and processed as per CLC-based algorithm. Within compound control sample no mutations were found. Samples containing potential mutations identified by both or either approaches were validated from independent sampling and/or independent PCR amplification (Table S2). For over-time monitoring and validation of previously identified mutation, cut-off 0.1% was applied (minimal coverage per base ≥ 10000).

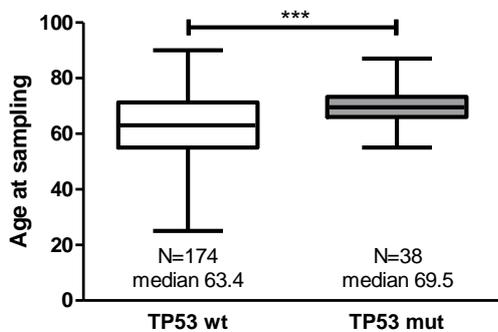


Supplemental Figure S2: Clonal structure assessed by colony-forming cell assay. In patient MP10, analysis was performed from the sample taken 14 months after mutation identification (~70% p.G245S). Number of analyzed colonies: MP2 - 34, MP10 - 44; MP68 - 13.

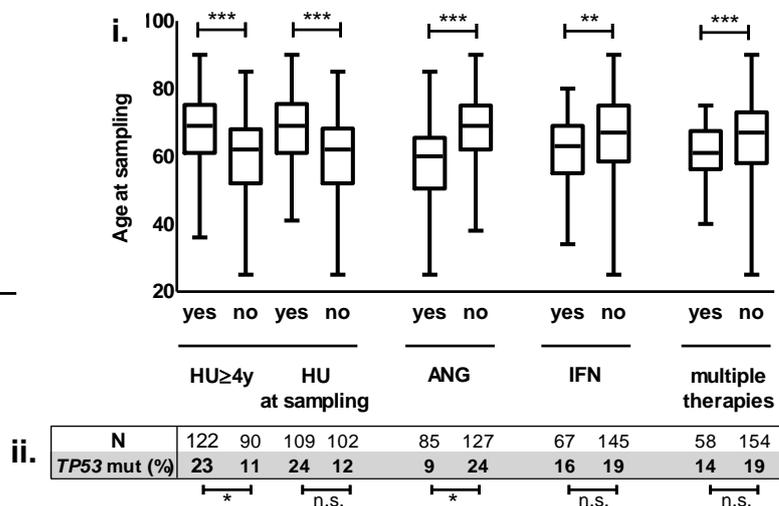


Supplemental Figure S3: Cytoreductive therapy length (months) in *TP53*-wt nad *TP53*-mutated cases. Box-and-whiskers plot, line indicates median. Difference was tested using Mann-Whitney test.

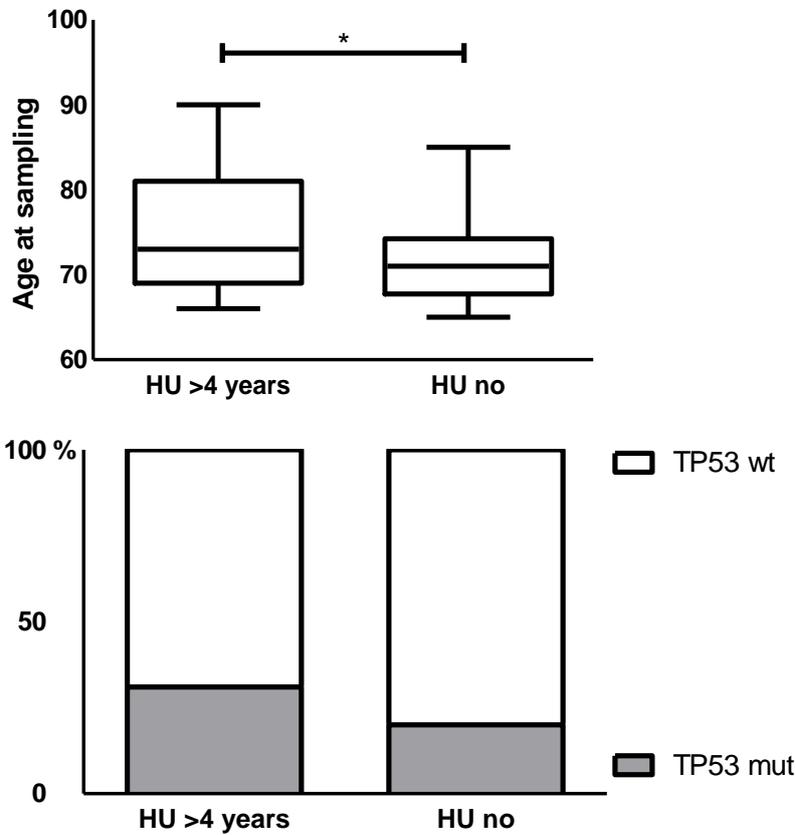
A



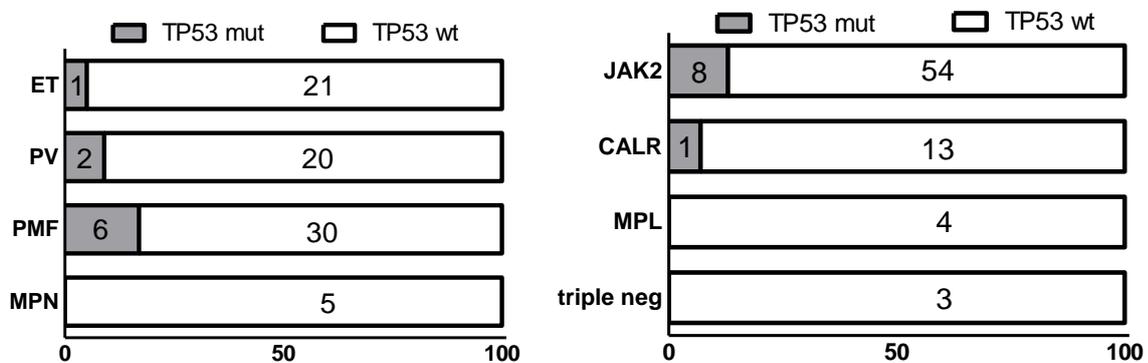
B



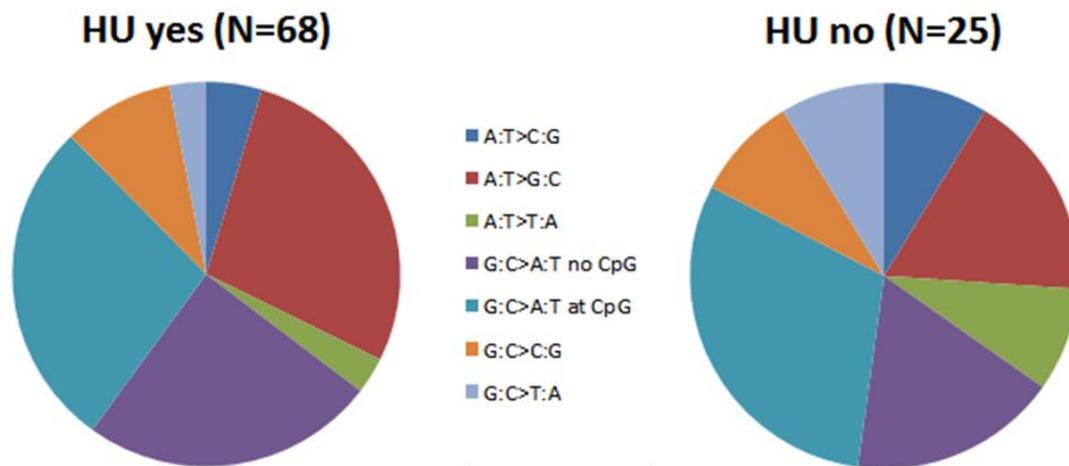
Supplemental Figure S4: Analysis limited to patients treated with HU ≥ 4 years and treated with non-HU drugs. A: Age at sampling in *TP53*-mut and *TP53*-wt patients ($P=0.0009$; Kruskal-Wallis test). B: Age at sampling (i) and *TP53* mutation frequency (ii) according to therapy parameters (HU-yes/HU-no $P=2.09 \times 10^{-6}$ and $P=0.03$; HU at sampling yes/no $P=2.23 \times 10^{-5}$ and $P=0.058$; ANG-yes/no $P=1.1 \times 10^{-9}$ and $P=0.01$; IFN-yes/no $P=0.0092$ and $P=0.85$; multiple therapies during disease course yes/no $P=0.0001$ and $P=0.42$ for age (Kruskal-Wallis test) and *TP53* mutation frequency (Fisher exact test), respectively). Lines within boxes indicate median, box limits - 25th and 75th percentiles, whiskers - minimum and maximum.



Supplemental Figure S5: Age at sampling and proportion of *TP53* mutations over 65 years of age (Age: $P=0.0262$; Mann-Whitney test; Mutation frequency: n.s.; Fischer exact test)



Supplemental Figure S6: Distribution of *TP53* mutations in untreated patients

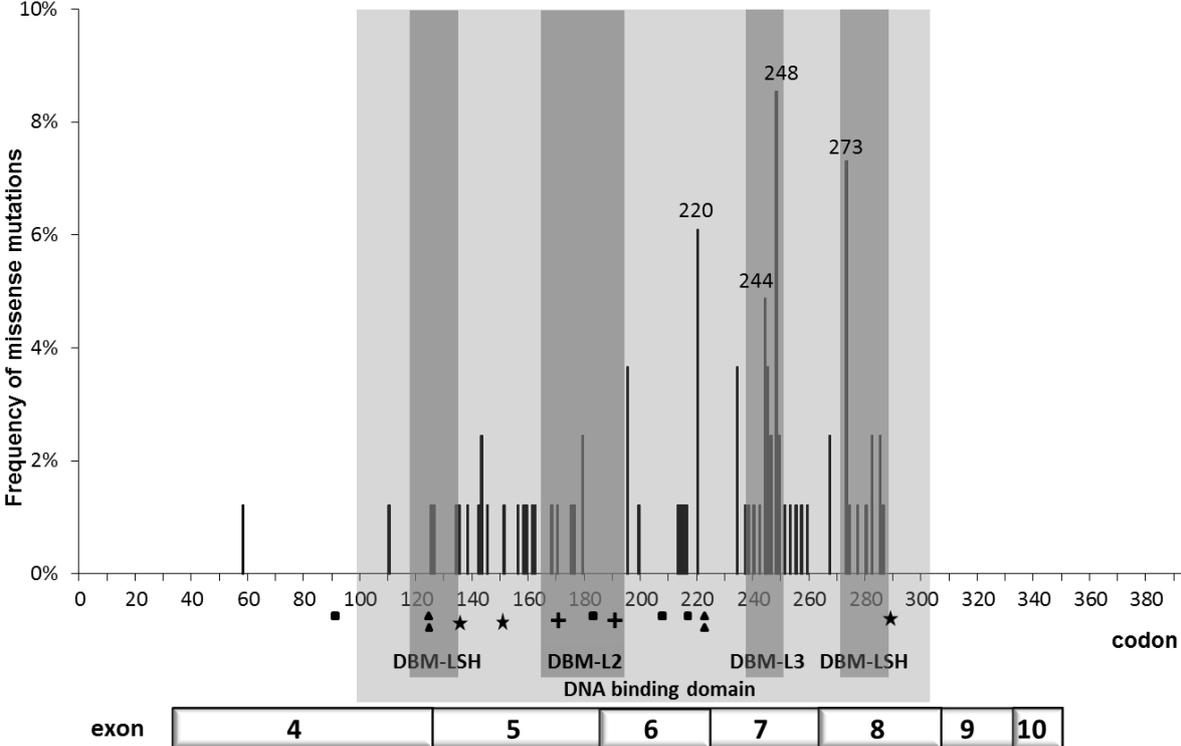


Supplemental Figure S7: Mutational profile of all *TP53* mutations found was similar in patients treated with HU (HU yes) and other therapies or untreated (HU no) (n.s., chi-square test).

Supplemental Figure S8 (next page): Location of the mutations with respect to the functional domains of the p53 protein. **A:** All identified mutations (Table S5). Aminoacids in direct contact with DNA are highlighted in dark gray. Exons 2,3 and 11 were not included in the study since the mutations are very rare in these regions. These exons were analyzed in multiple samples from 48 patients (29 *TP53*-mutated and 19 *TP53*-wt) and none mutation was identified. **B:** Mutations of which allelic fractions increased during follow-up (Table 3). *DBM* - DNA binding motif; *L* - loop, *LSH* - loop/sheet/helix. **C:** Type of mutation (all identified mutations).

Supplemental Figure S8

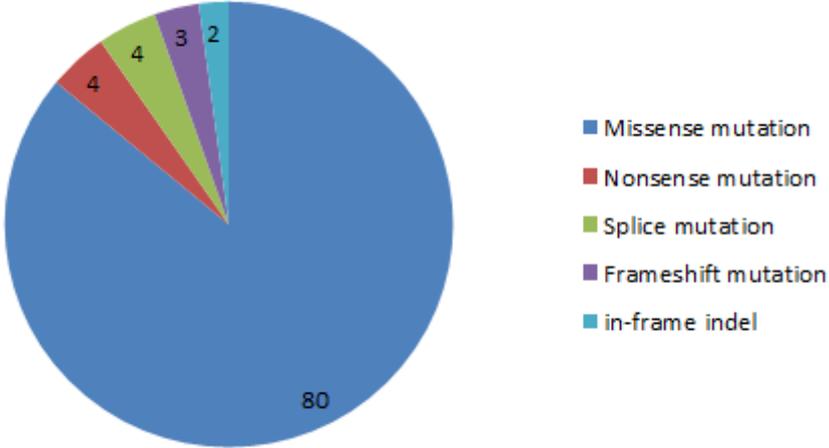
A



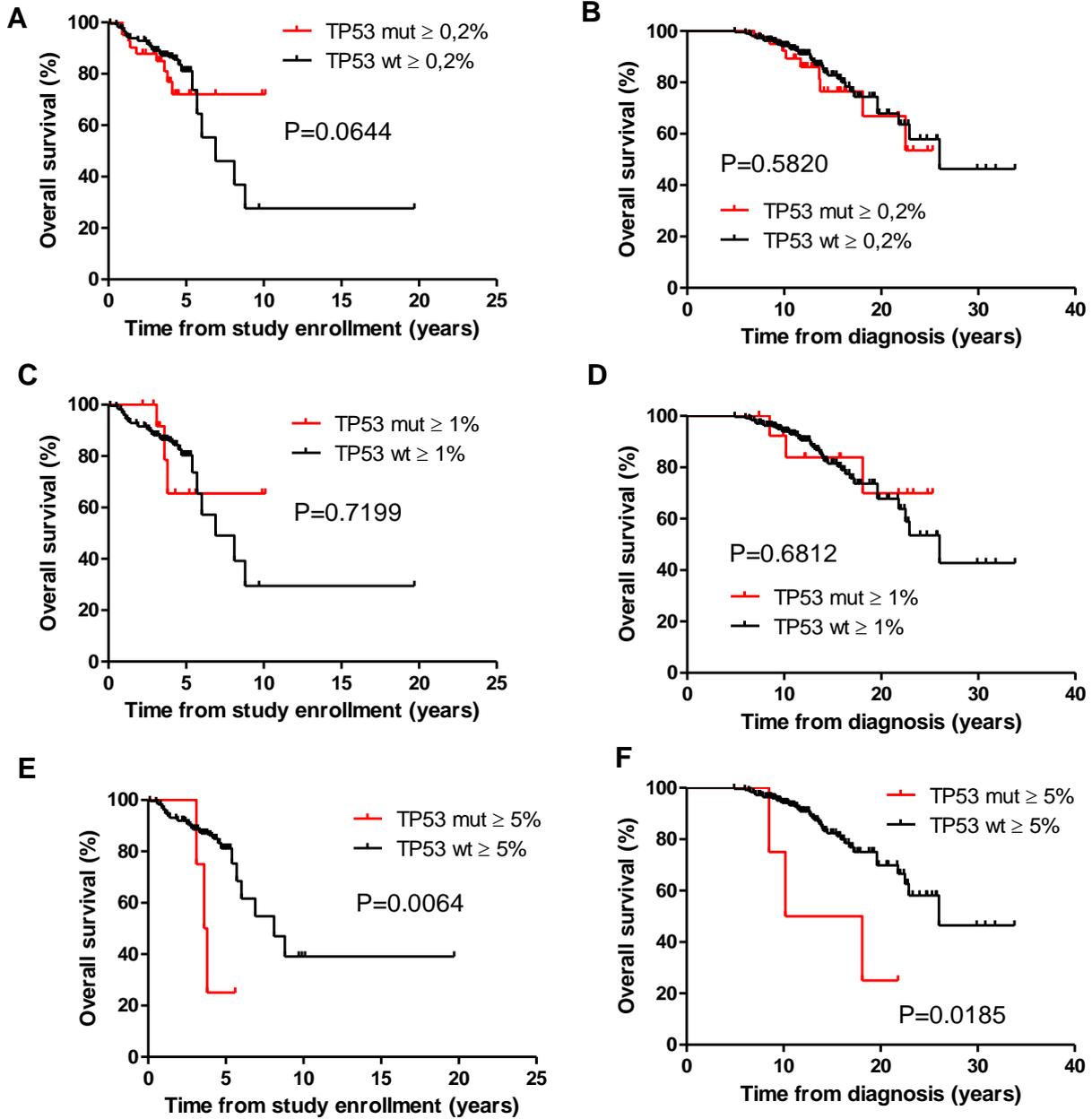
B



C



Supplemental Figure S9



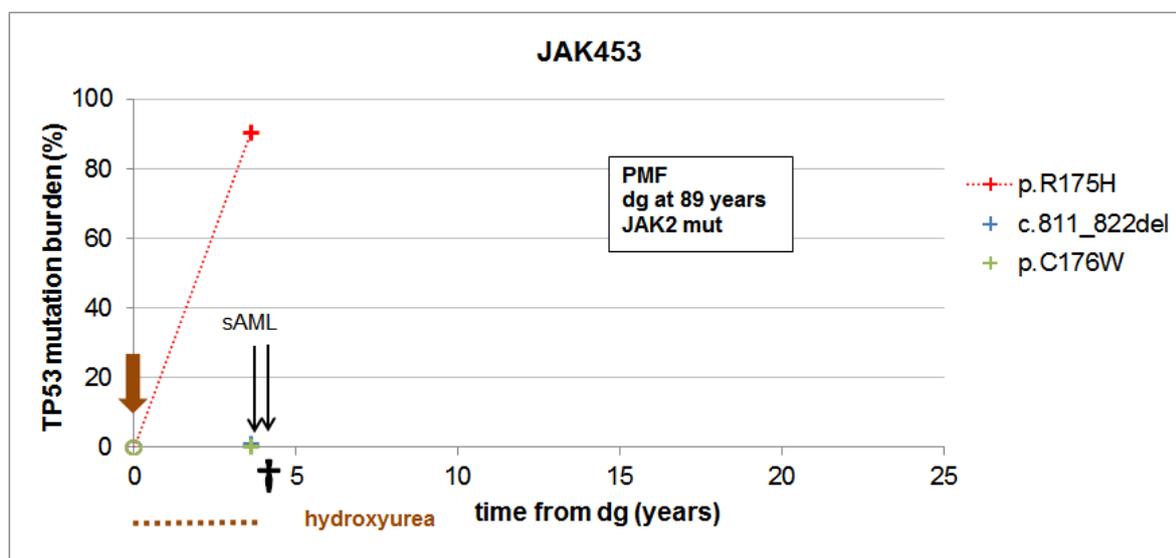
G

Patient	% VAF (sum [#])	OS* (y)	OS (y)	Final % VAF (sum [#])	sAML	Cause of death
JAK646	16 (17)	3.4	3.4	unkn.	no	cardiac failure
MP68	8 (17)	3.6	8.5	3 (5)	yes	sAML TP53wt
MP10	11 (12)	3.8	10	86 (88)	no	accident
186A	8 (8)	3.1	18	unkn.	yes	sAML
MP2	11 (11)	3.4	21	2 (2)	no	alive

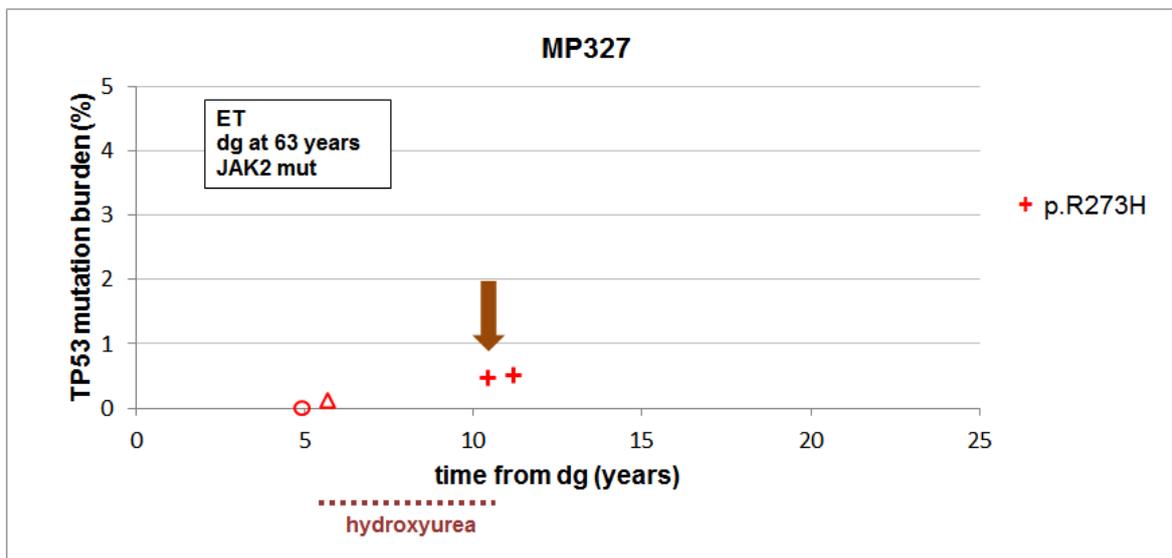
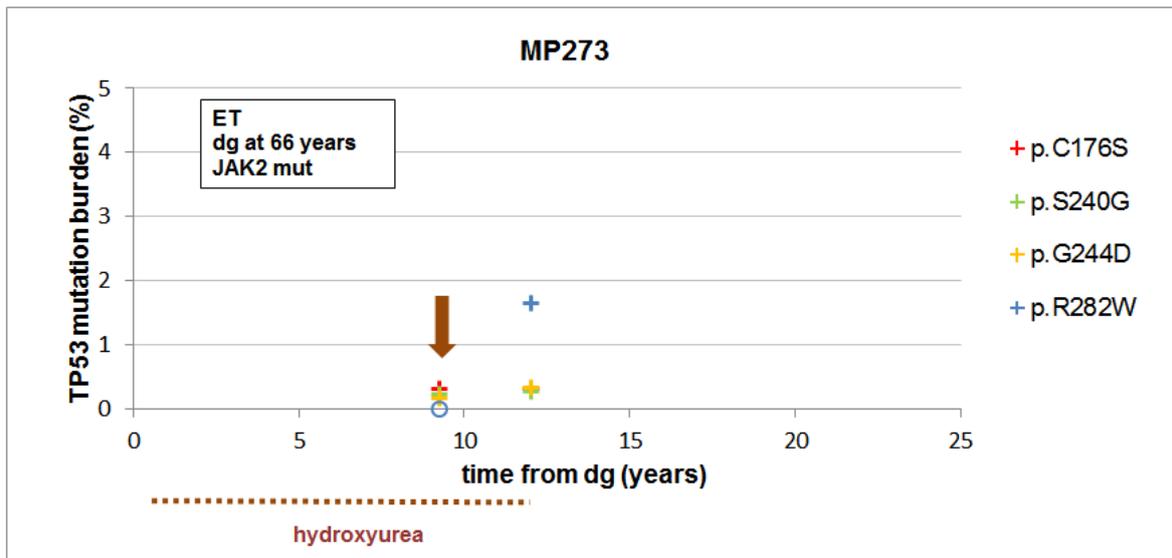
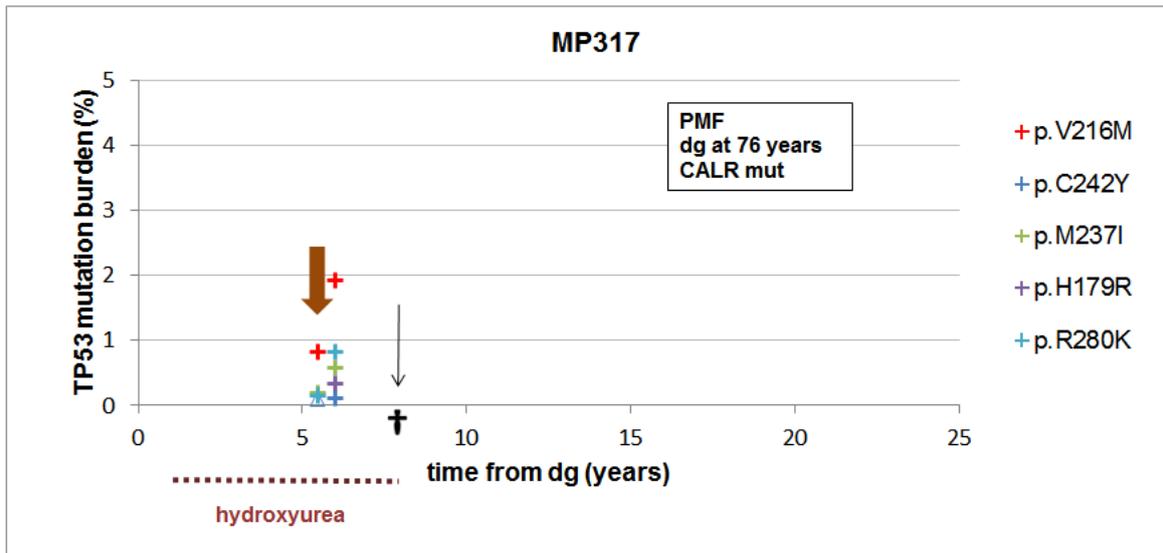
Supplemental Figure S9: A-D: Overall survival in treated patients according to the *TP53* mutational status from time of study enrollment (A, C) and from diagnosis (B, D). **E-F:** Overall survival in all examined patients according to *TP53* mutation with VAF>5% from study enrollment (E) and diagnosis (F). **G:** Information on patients with *TP53* mutation above 5% VAF (at study enrollment) shown in graphs E and F. #sum of all *TP53* mutated subclones; *overall survival from study enrollment

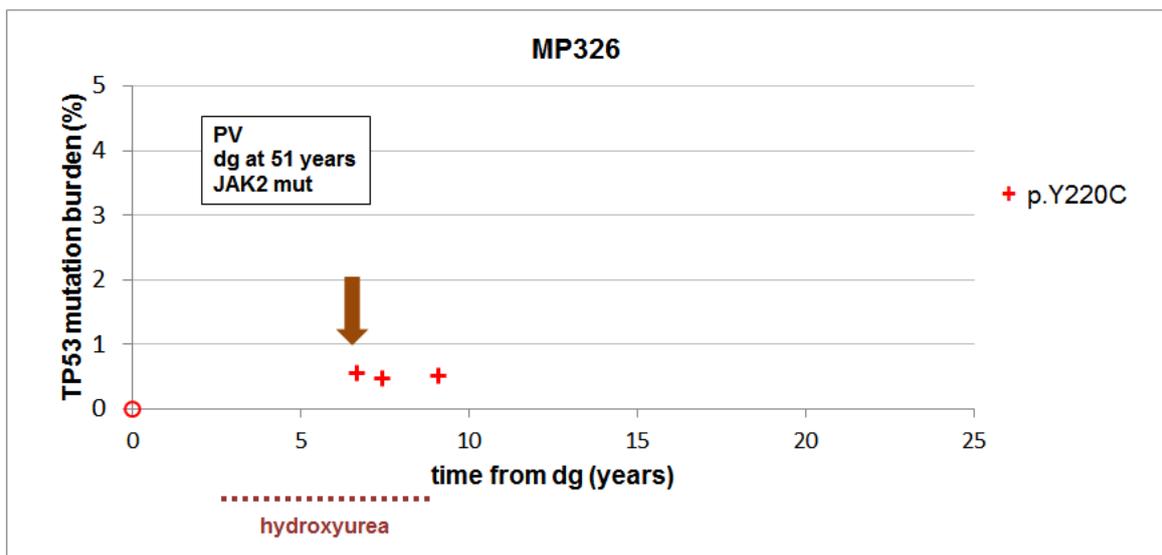
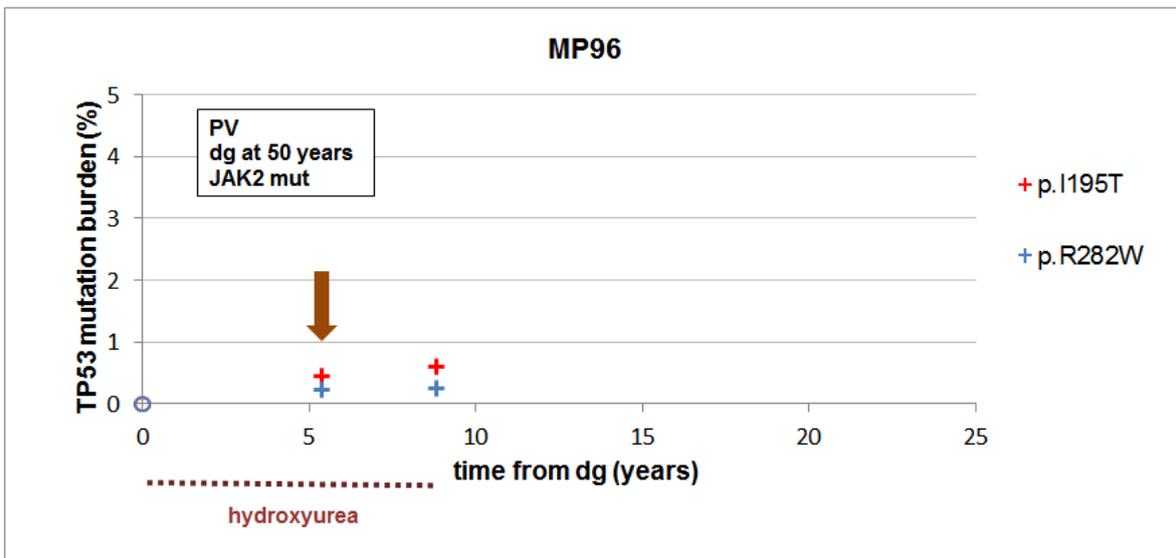
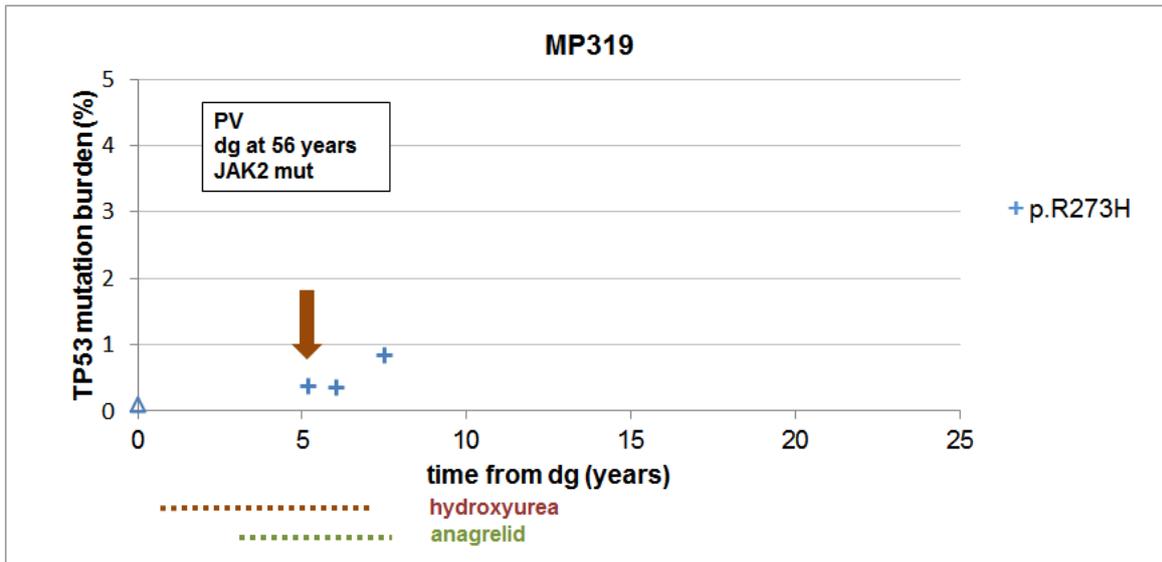
Supplemental Figure S10: Graphic overview of *TP53* mutation development in individual patients. **A:** Patient developing *TP53*-mutated sAML in whom no *TP53* mutation was identified at diagnosis. **B:** Patients treated with HU before the study enrollment. **C:** Patients treated with IFN and/or ANG or untreated for MPN before the study enrollment.

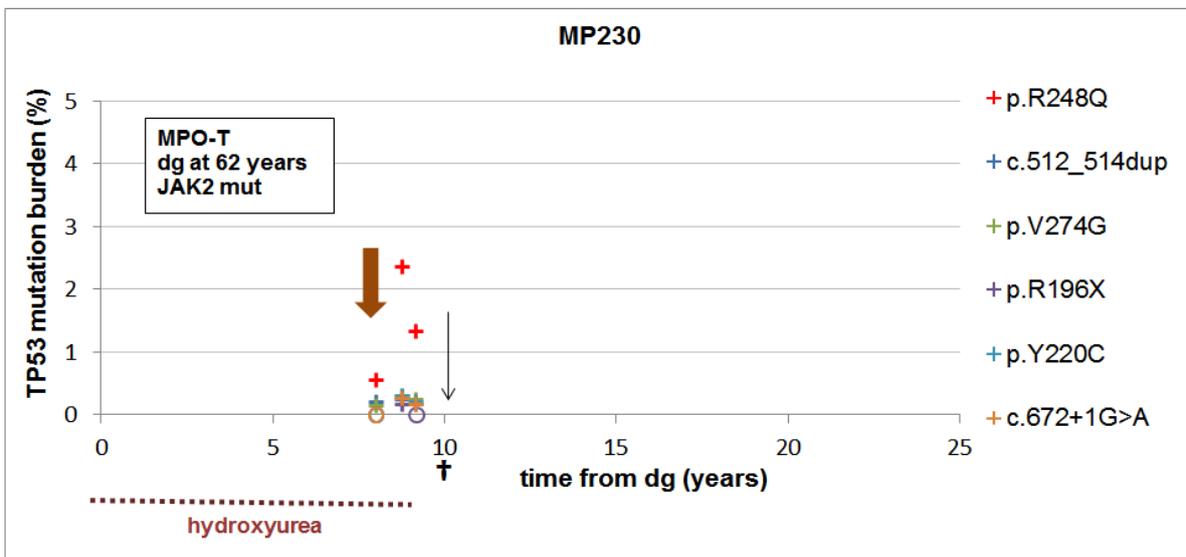
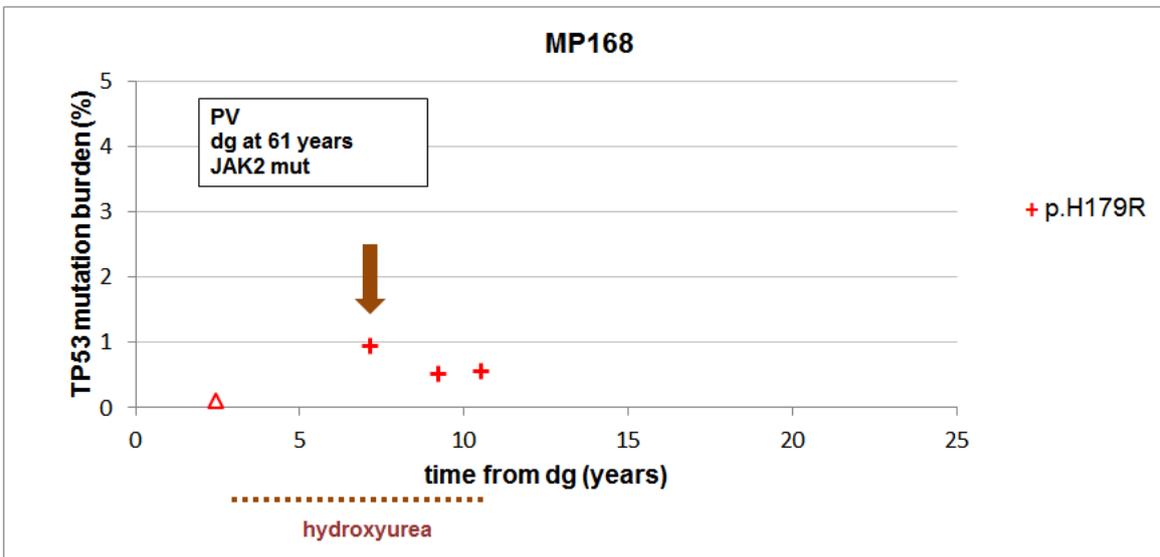
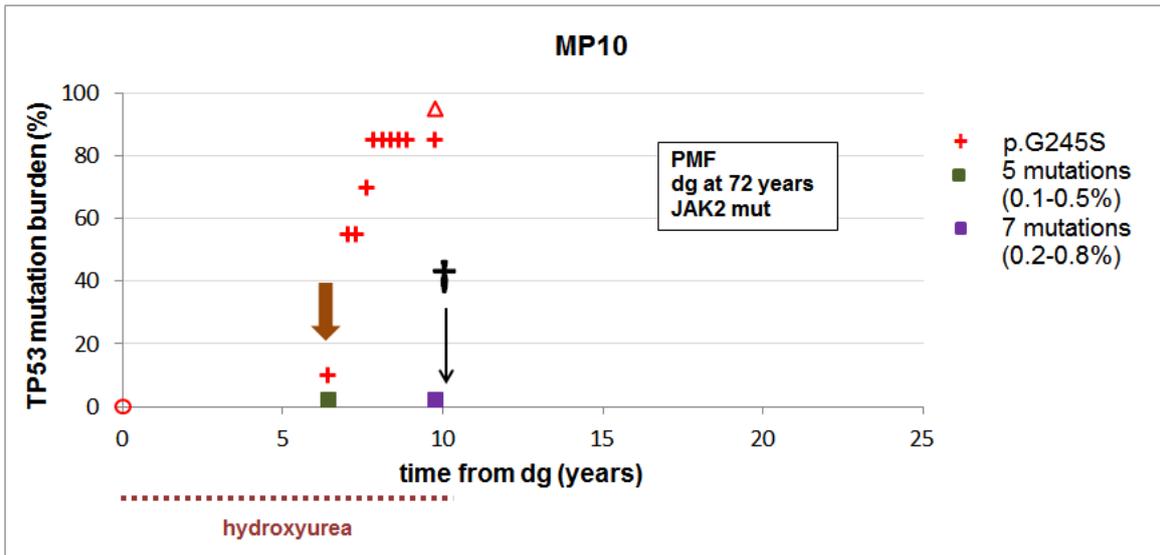
A.

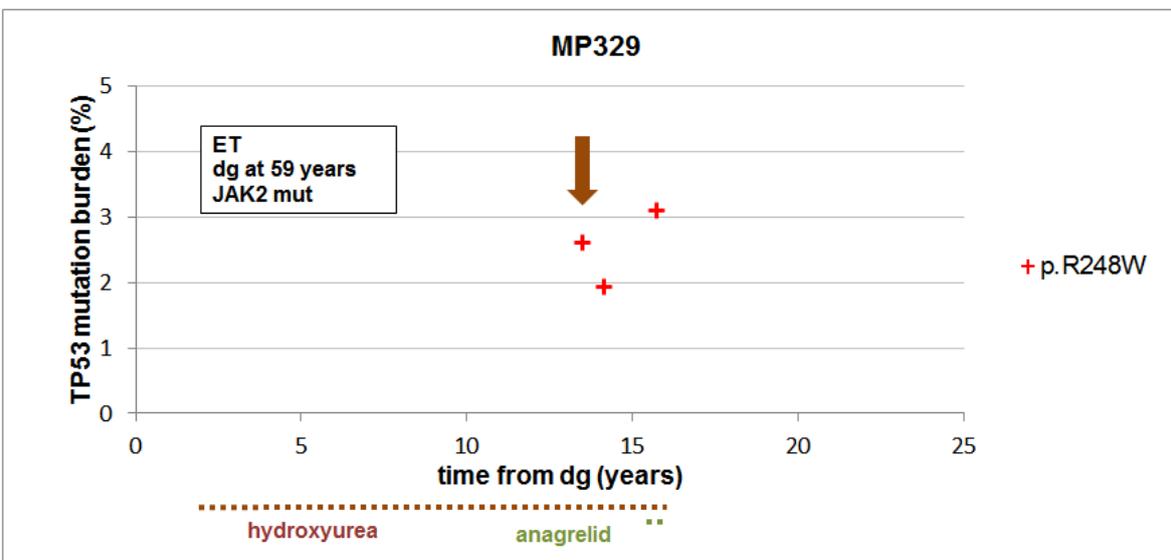
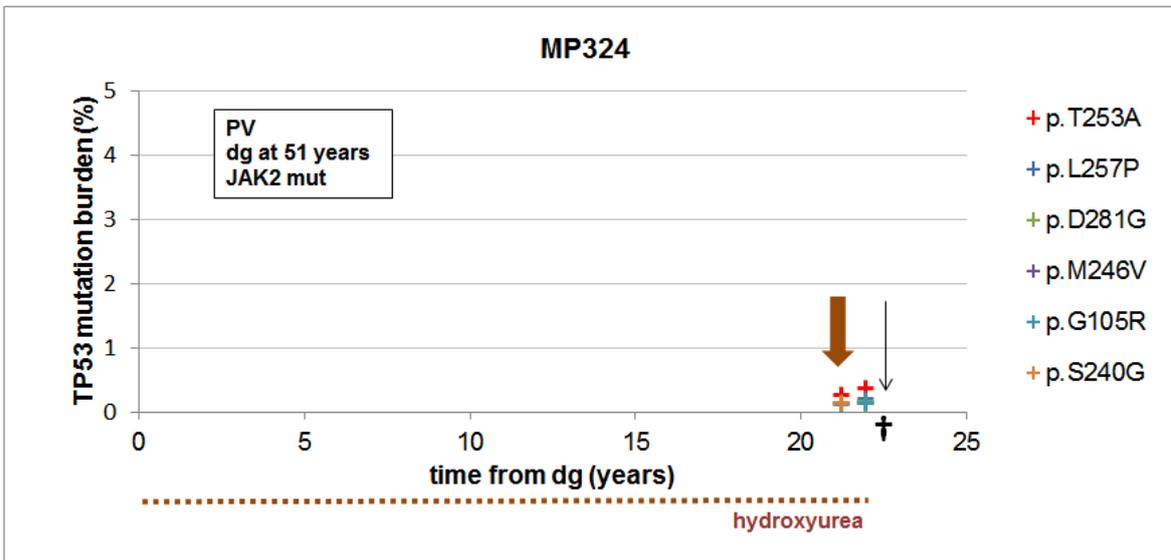
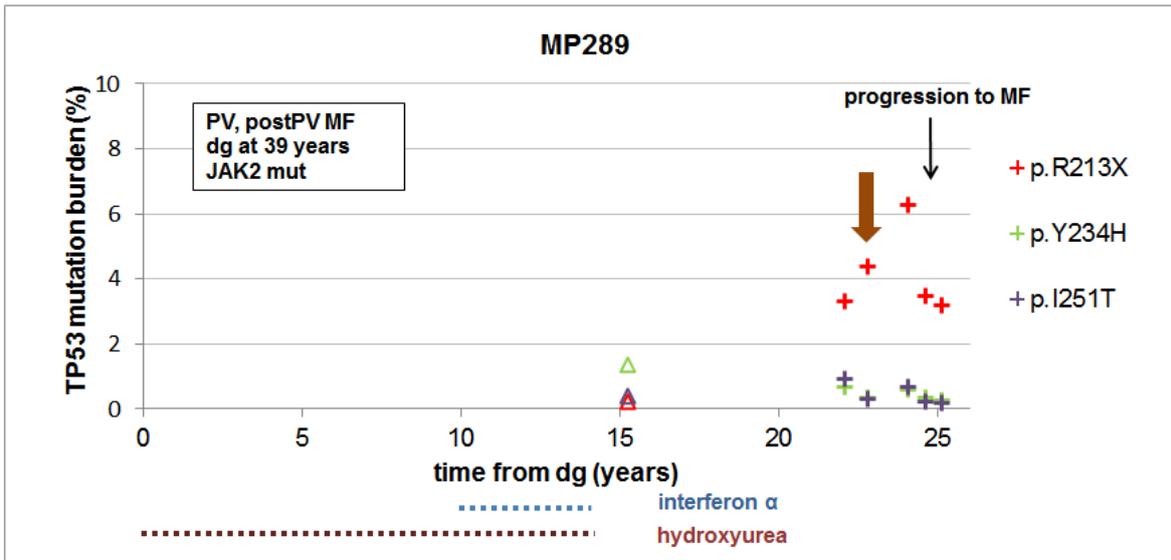


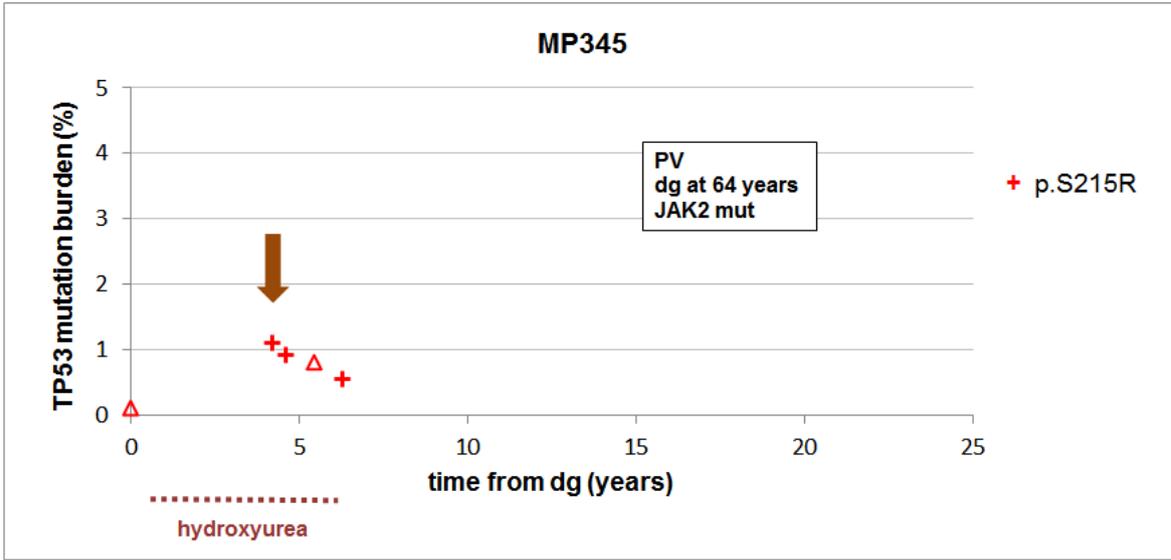
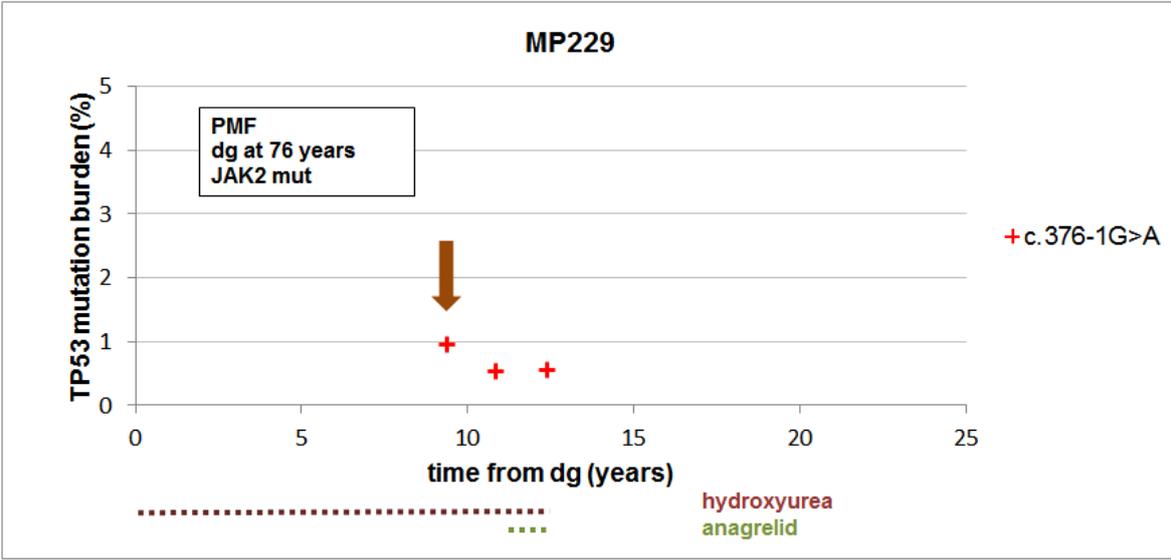
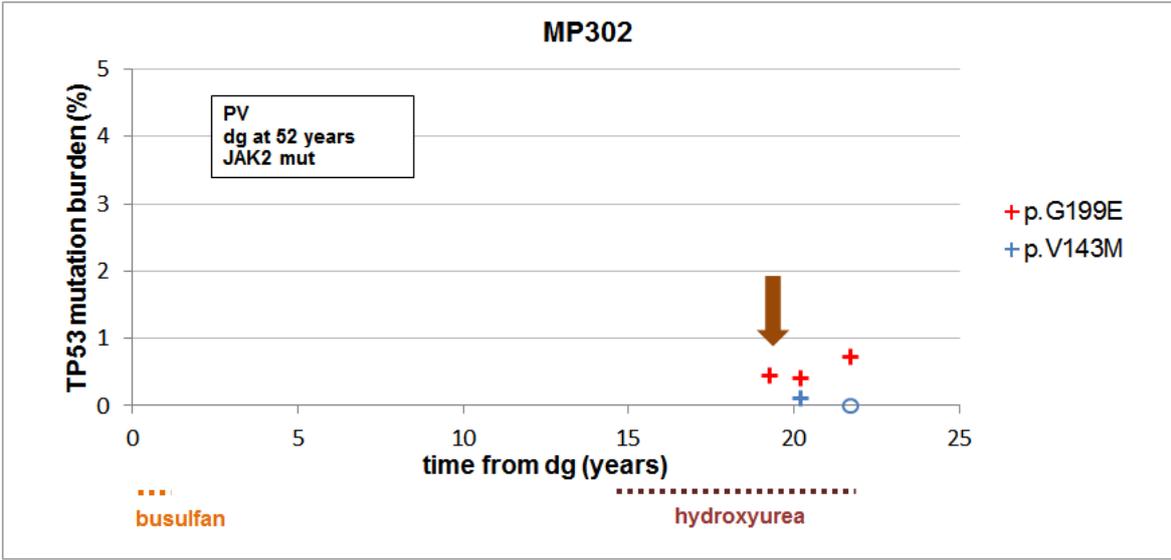
B.

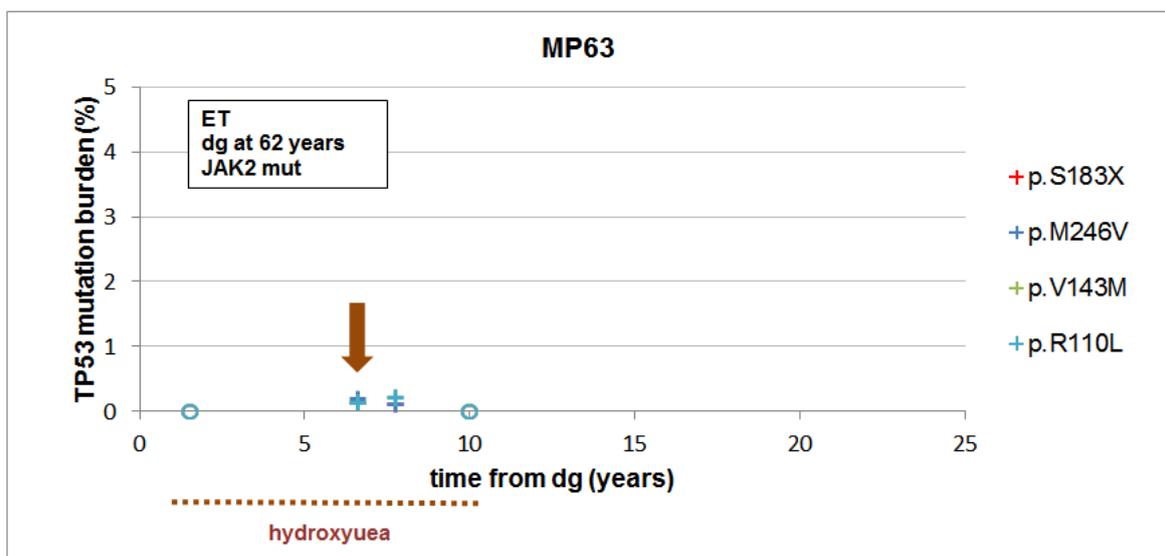
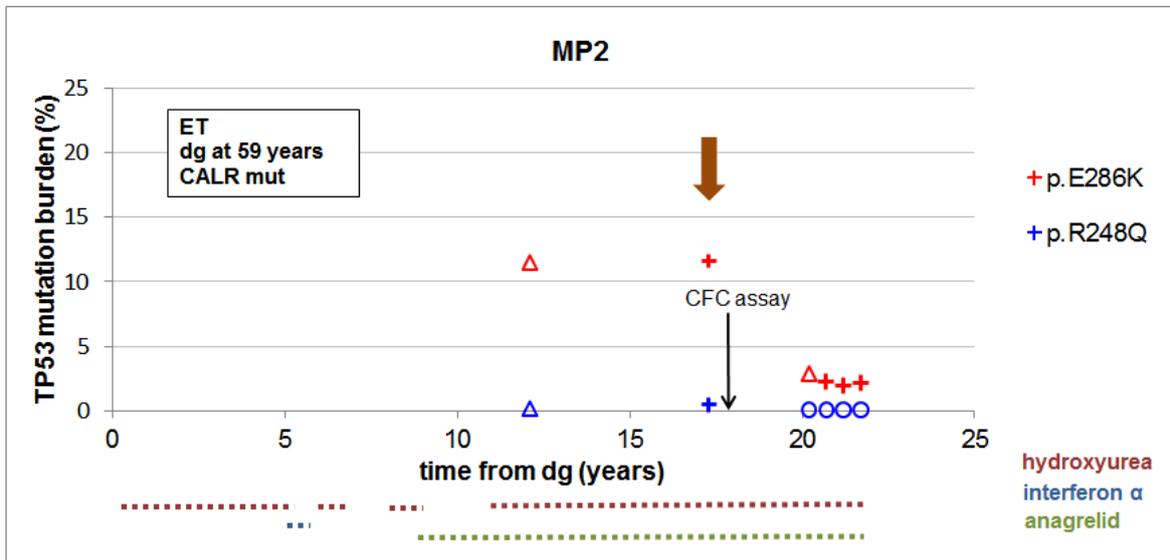
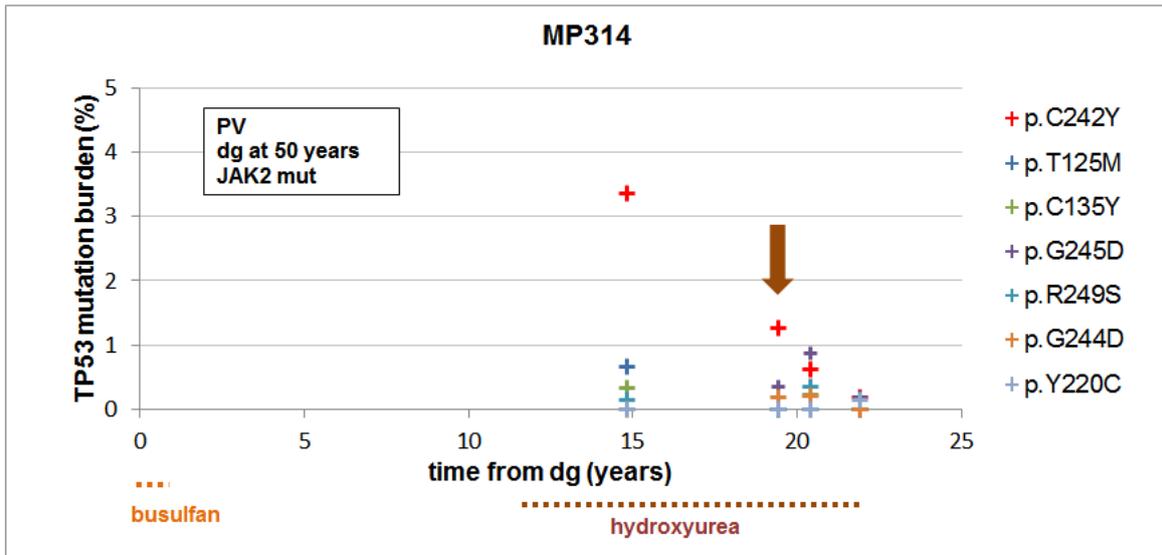


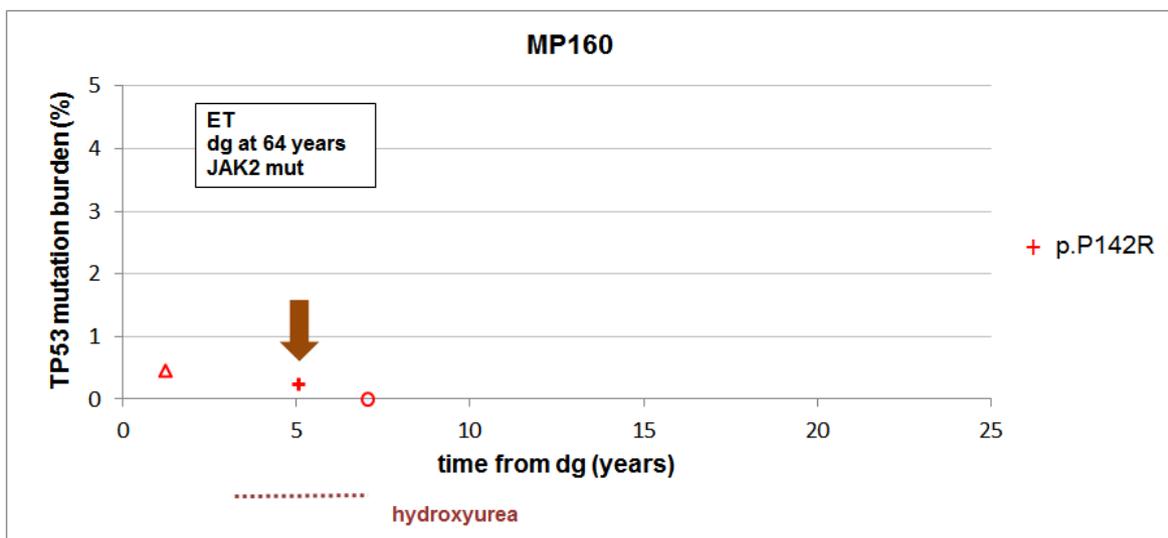
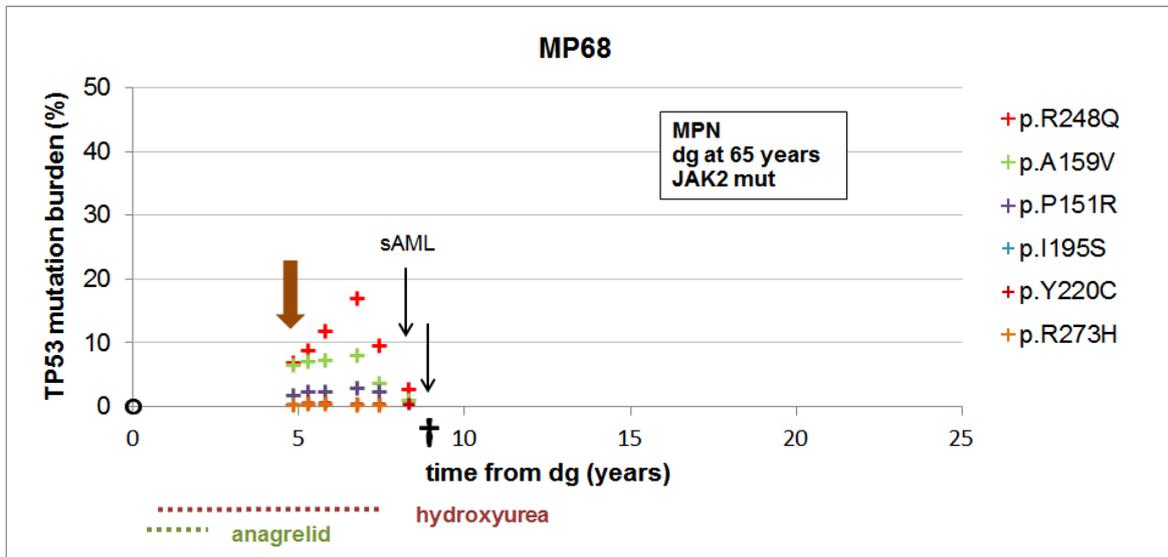
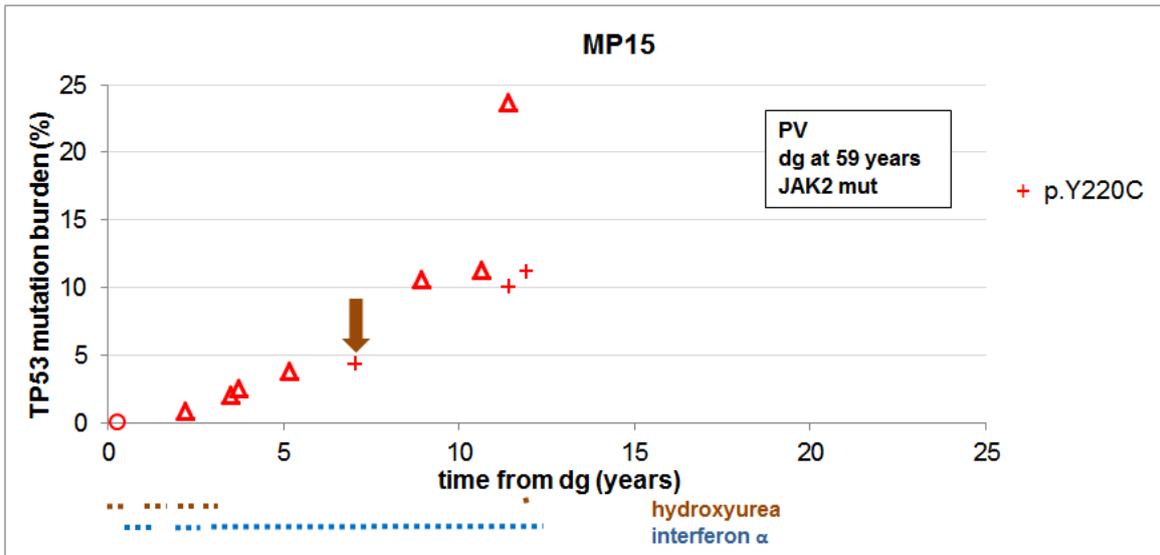


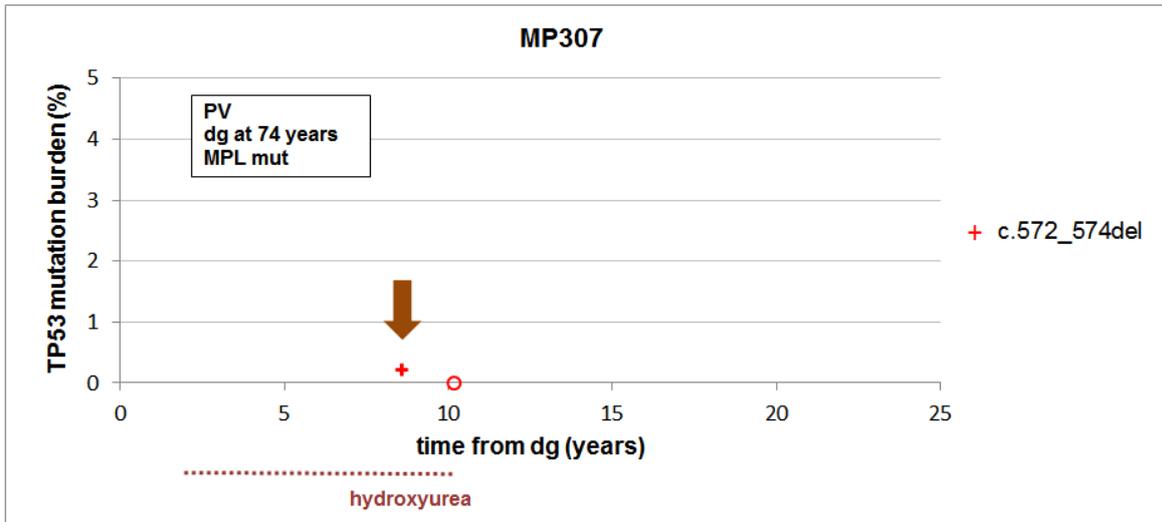




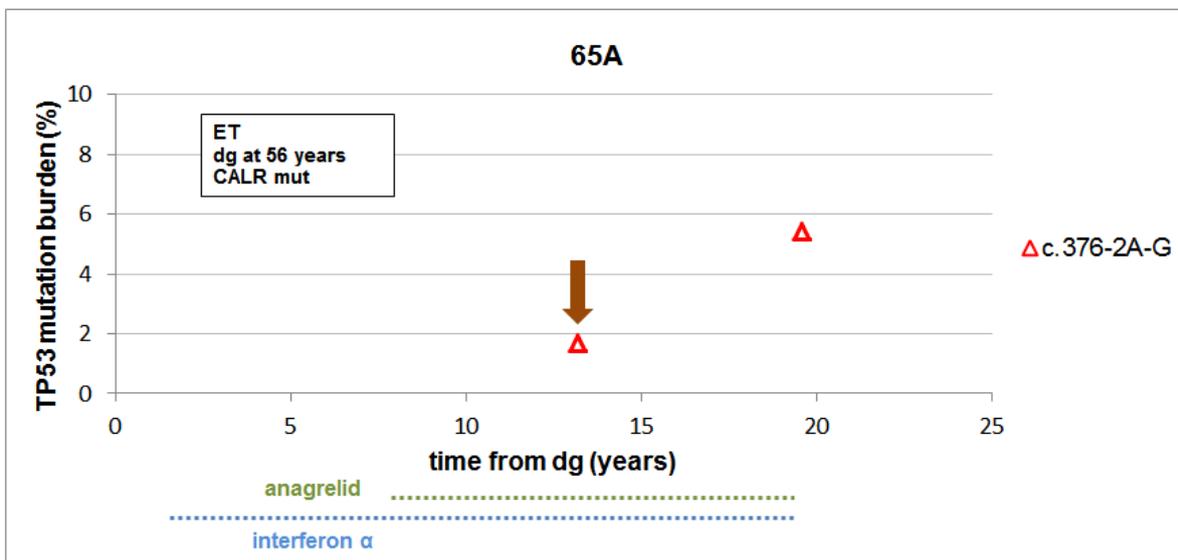
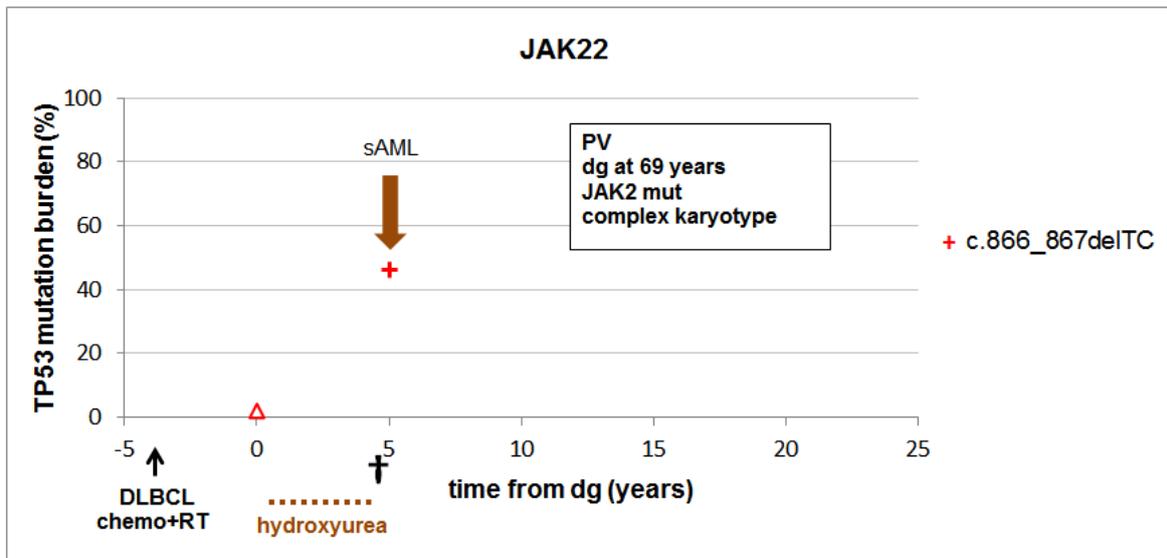


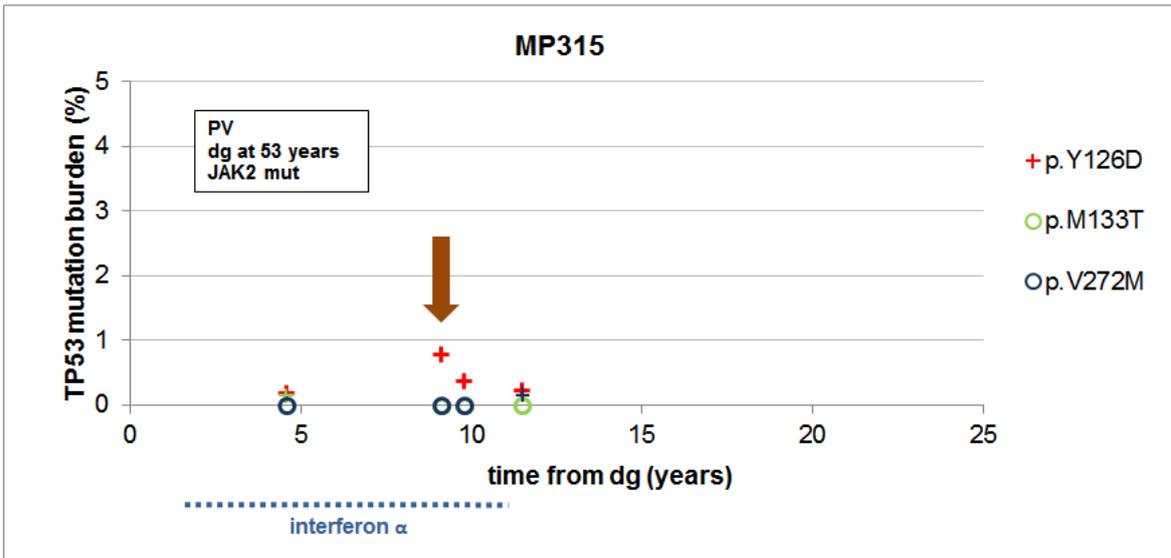
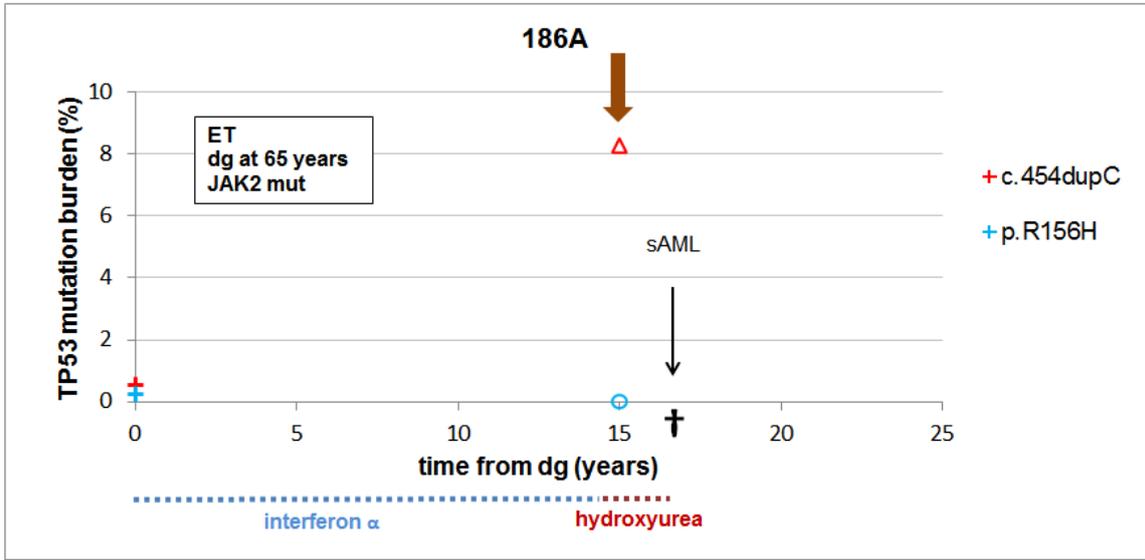
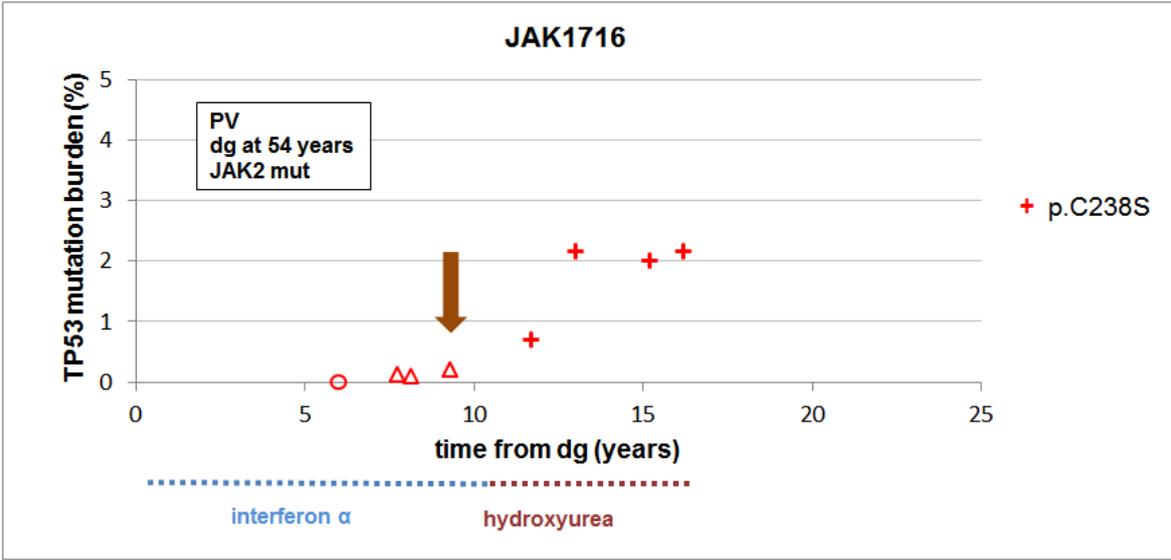


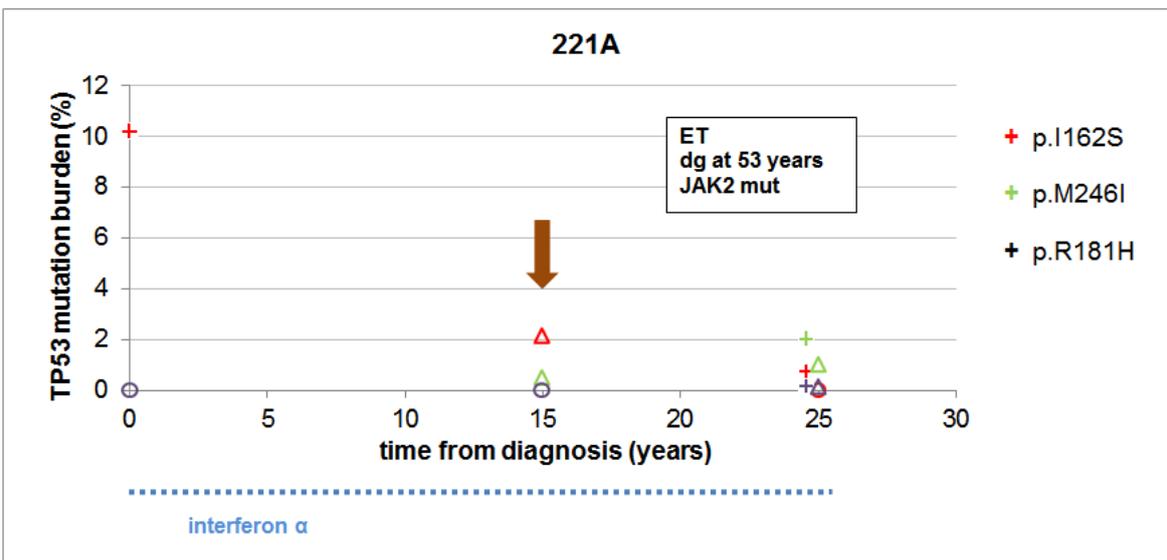
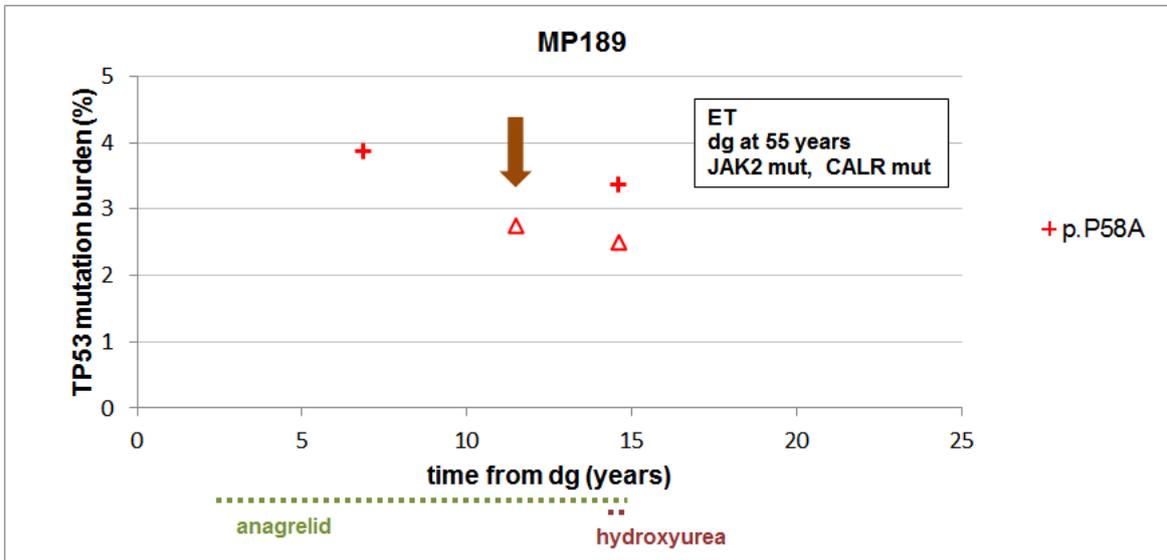
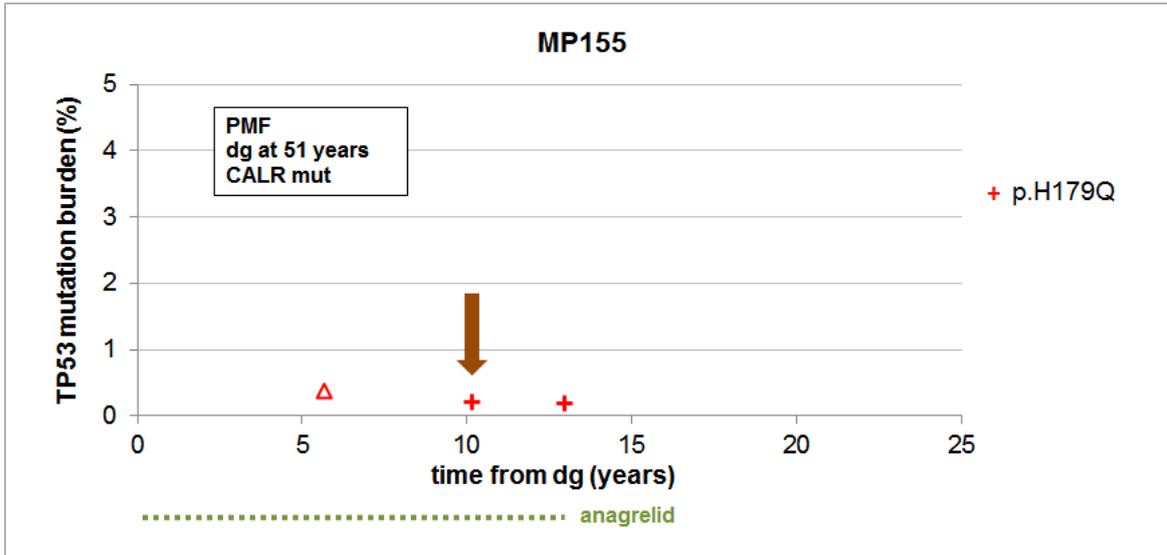


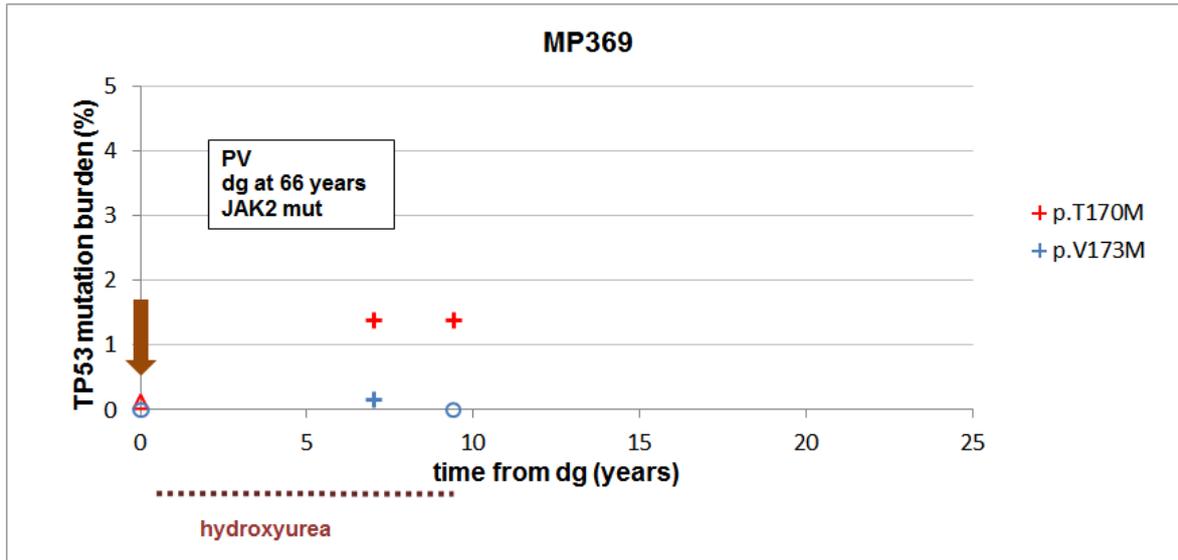


C.









Mutation identification



Mutation in granulocytes



Mutation in leukocytes



Mutation not found



Interferon α treatment



Anagrelid treatment



Hydroxyurea treatment



Busulfan treatment



Death