Supplementary Online Material

Materials and Methods

Patient materials: Subjects were selected from 869 unrelated women of French-Canadian ancestry (with Quebec-born grandparents) recruited from the general community. All women answered a medical questionnaire, and gave an informed consent. The study was approved by the Maisonneuve-Rosemont Hospital's Ethics Committee in 1998, and re-approved yearly in accordance with institutional bylaws. Granulocytes were collected from Ficoll density gradient centrifugation of peripheral blood. In addition, CD3+ cells were isolated by FACS plus buccal epithelial cells were collected [1].

Genetic Analysis: For exome sequencing, 3ug of high molecular weight DNA was sheared to an average size of 180bp+/-80bp. Fragments from 200 to 250 bp were selected and subjected to 8 cycles of PCR. The library was then hybridized to the SureSelect Human All Exon Kit (Agilent catalog #G7540D). Sequencing was performed on the SOLiD 3plus or SOLiD 4. Variant detection was performed as described in detail in **Supplemental Figure 1A**. Briefly, samples were aligned with BFAST[2,3] and processed with Picard (http://picard.sourceforge.net) and GATK tools[4,5]. SNP and indel calling was done with the GATK UnifiedGenotyper[4,5], samtools mpileup[6], and SomaticSniper (http://gmt.genome.wustl.edu/somatic-sniper/current/). The union of all putative somatic calls from all three callers was taken and only variants that altered coding sequence and were not in dbSNP129 or 1000 genomes (www.1000genomes.org/) were then validated with Sanger sequencing **(Supplemental Figure 1B)**. Candidate frameshift alterations were validated by Topo-TA cloning (Invitrogen K4500-01). Sanger sequencing primers available upon request.

Clonality Assay: Determination of X-inactivation ratios was performed as previously described[1]. Polymerase chain reaction (PCR) amplification of the polymorphic CAG repeat at the *HUMARA* locus was performed in tandem on undigested (HEX-labeled primer) and on *Hpa*II-digested (6-FAM–labeled primer) DNA. PCR products were analyzed on an ABI Prism 3100 genetic analyzer (Applied Biosystems, Foster City, CA) to determine the area under the curve (AUC) for each allele. The ratio between the X-linked alleles was expressed as the degree of skewing (DS). To determine DS, the proportion of the superior allele *(Psup)* is calculated as follows:

$$1-\left[\frac{\frac{\frac{A}{A+a}}{\frac{A'}{A'+a'}}}{\frac{\frac{A}{A+a}}{\frac{A'}{A'+a'}}+\frac{\frac{a}{A+a}}{\frac{a'}{A'+a'}}}\right]$$

A and *a* represent the AUC for the upper and lower alleles from the digested sample; *A*' and *a*' represent the AUC for the upper and lower alleles in the undigested sample. *DS* = |*Psup* – 0.5|. Allelic skewing consistent with clonal granulopoiesis was defined as a 3:1

ratio between X-linked alleles, which is equivalent to DS at least 0.25. Patients with "clonal granulopoiesis" have at least 50% clonally derived granulocytes but may have 50% or fewer admixed polyclonal cells; conversely, patients with "polyclonal granulopoiesis" may have 50% or fewer admixed clonal granulocytes.

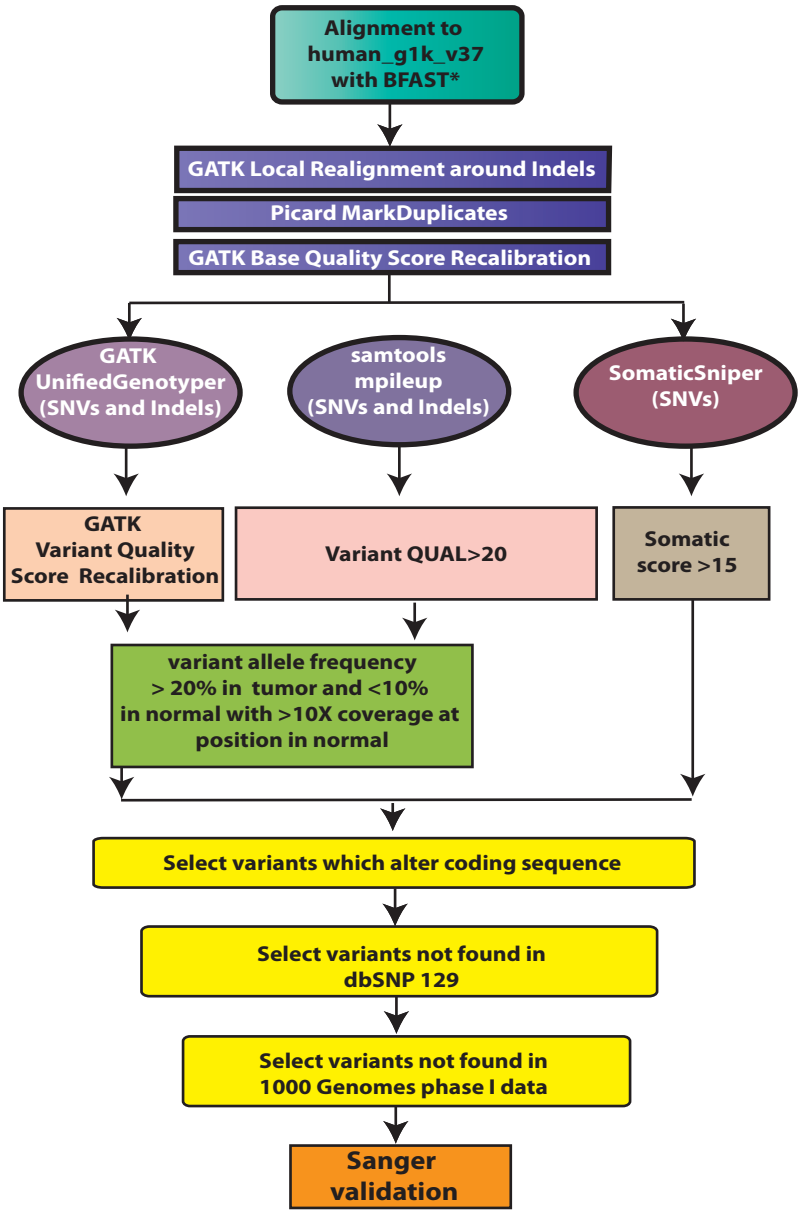
Statistical Analyses: NCSS2004 was used for descriptive statistics, T-test, box-plot, Mann-Whitney non parametric test, etc. Fischer exact test was used instead of chi square test. Logistic regression was done with SAS9.1. Online Fisher's Exact Test was used at http://www.langsrud.com/fisher.htm

Epigenetic and expression analysis: Global 5-hydroxymethylcytosine (5-hmC) and 5-methylcytosine (5-mC) levels in PMN DNA from normal elderly individuals was also assessed by liquid chromatography-electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS) as previously described[7]. DNA from 10 young/clonal/*TET2* wildtype, 15 elderly/non-clonal/*TET2* wildtype, and 6 elderly/clonal/*TET2* mutant individuals was used for LC/MS analysis. The amount of global 5-hmC and 5-mC was expressed as a proportion of total cytosine. Next, a normalized value was calculated for each sample by dividing the raw value of 5-hmC or 5-mC of each sample by the average value for that experimental group. Finally, 5-mC/5-hmC data is displayed as values normalized to the values in aged-clonal, *TET2* wildytpe individuals. HELP assay originally performed on *TET2* wildtype and mutant patients with acute myeloid leukemia (AML) to identify genetic loci of differential hypermethylation as previously described[8]. These differentially methylated loci from AML patient data were then investigated in 6 normal elderly individuals wildtype and 6 mutant for *TET2* using MALDI-TOF mass spectrometry EpiTYPER by MassARRAY (Sequenom, San Diego, CA) on bisulfite-converted PMN DNA as previously described[9]. MassARRAY primers were designed as previously described and all primer sequences are available upon request. The expression of genes with differential methylation as validated by Epityper was analyzed by qRT-PCR analysis using SYBR green quantification and ABI 7500 sequence detection system. Primer sequences available upon request.

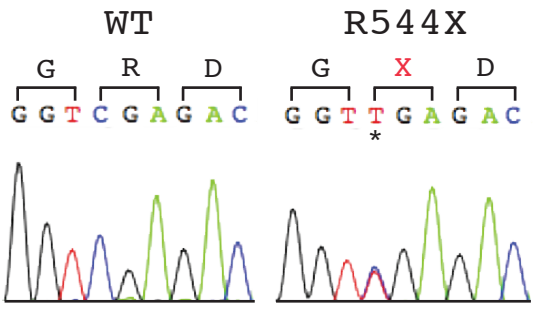
Supplementary Figure 1: Exome sequencing analysis and Sanger sequencing validation of *TET2* mutations in normal individuals with clonal hematopoiesis. Variant detection was performed as described in **(A).** Sanger sequencing validates and reveals clonal dominance of the *TET2* mutation identified by exome sequencing **(B).** Additional examples of somatic *TET2* mutations found in elderly normal individuals with clonal hematopoiesis **(C)**.

Supplementary Figure 1

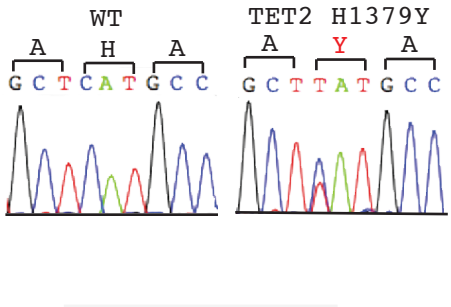
(A)

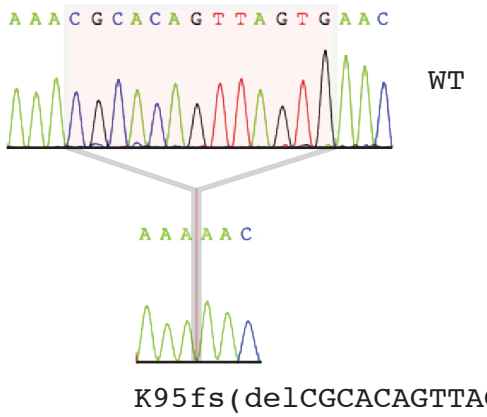


(B)



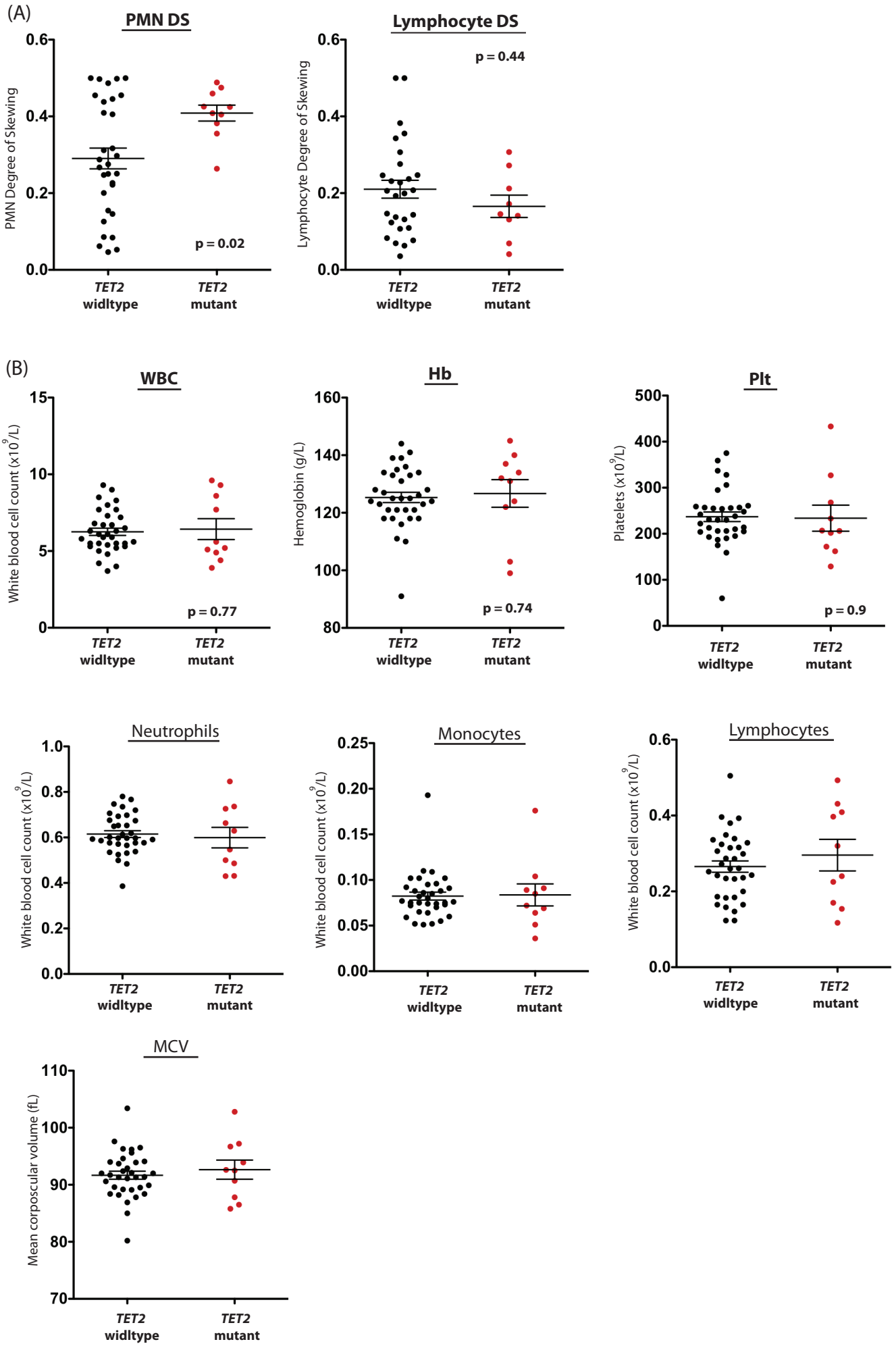
(C)





Supplementary Figure 2: **Hematologic parameters in normal elderly individuals with *TET2* mutations are indistinguishable from those of *TET2* wildtype age-matched controls.** Elderly subjects with *TET2* mutations (n=10) were compared with age-matched controls (n=34) in a 1:3 ratio in terms of polymorphonuclear neutrophil (PMN) and lymphocyte degree of skewing (DS) **(A)** and peripheral blood white blood cell count (WBC) (reference range 4.5-10.8 $x10^9$/L), hemoglobin (Hb) (reference range 118-158 g/L), platelets (Plt) (reference range 140-440 $x10^9$/L), peripheral blood neutrophils (reference range: 1.8-7$x10^9$/L), monocytes (reference range: 0.1-0.8$x10^9$/L), lymphocytes (reference range: 1.3-3.5$x10^9$/L) and mean corpuscular volume (MCV) (reference range: 81-98 fL) **(B)**. Skewing within the myeloid compartment was the only statistically significant difference (p=0.02, two-tailed Mann-Whitney U test).

Supplementary Figure 2



Supplementary Table 1: Somatic mutations found by exome sequencing in a normal elderly individual.

Mutation	Chromosome	Position
TET2 p.R544X	4	106,156,729
DNMT3A p.L508V	2	25,468,152
KIAA1919 p.E219D	6	111,587,422
ERCC6 p.Q1413R	10	50,667,105
SLC39A12 p.V469I	10	18,280,215

Supplementary Table 2: Primer sequences used for genomic DNA resequencing and for quantitative real-time PCR (qRT-PCR) analysis[1].

Gene	Forward Primer Sequence	Reverse Primer Sequence
qRT-PCR primer sequences		
SCAND2	TCCTGGGAGCTGCTCACTAT	GTTAAAGAACGGGATGCCAA
BAIAP2	CTGGATCTGCCTGTGGACTT	AAAGGCTACTTTGACGCCCT
GPR77	TCCGAGAGGTCGCTGTAATC	TCAAGGACTCCCAAAACCAG
Genomic DNA primer sequences		
DNMT3a	GTAAAACGACGGCCAGTCCTCTCTCCCACCTTTCCTC	CAGGAAACAGCTATGACCCTGAGTGCCGGGTTGTTTAT
	GTAAAACGACGGCCAGTGGAAAACAAGTCAGGTGGGA	CAGGAAACAGCTATGACCTGGATCTAAGATTGGCCAGG
	GTAAAACGACGGCCAGTccacactagctggagaagca	CAGGAAACAGCTATGACCggggctcttaccctgtgaac
	GTAAAACGACGGCCAGTcatggcagagcagctagtca	CAGGAAACAGCTATGACCtgtgtggctcctgagagaga
	GTAAAACGACGGCCAGTAATACCCAACCCCAGGAGTC	CAGGAAACAGCTATGACCCTTCCTGTCTGCCTCTGTCC
	GTAAAACGACGGCCAGTGAAGCCATTAGTGAGCTGGC	CAGGAAACAGCTATGACCCAACTTGGTCCCGTTCTTGT
	GTAAAACGACGGCCAGTTTGCCAAAAGTATTGGGAGG	CAGGAAACAGCTATGACCCCAGTTGGATCCAGAAAGGA
	GTAAAACGACGGCCAGTaagcttccccttgggataa	CAGGAAACAGCTATGACCcagggtgtgtgggtctagga
	GTAAAACGACGGCCAGTAGGGTCCTAAGCAGTGAGCA	CAGGAAACAGCTATGACCCGGTCTTTCCATTCCAGGTA
	GTAAAACGACGGCCAGTaggtgtgctacctggaatgg	CAGGAAACAGCTATGACCcagggcttaggctctgtgag
	GTAAAACGACGGCCAGTATCTGGGGACTAAAATGGGG	CAGGAAACAGCTATGACCCCTGGACTCTTTTCTGGCTG
	GTAAAACGACGGCCAGTAGCAAAGGTGAAAGGCTGAA	CAGGAAACAGCTATGACCAGCCCAAGGTCAAGGAGATT
	GTAAAACGACGGCCAGTTCCCAGGCAACAAACTTACC	CAGGAAACAGCTATGACCGAACAAGTTGGAGACCAGGC
	GTAAAACGACGGCCAGTTCTTCTGGAGGAGGAAAGCA	CAGGAAACAGCTATGACCCCTGTGCCACCCTCACTACT
	GTAAAACGACGGCCAGTAGTAGTGAGGGTGGCACAGG	CAGGAAACAGCTATGACCCTCCTCTTTGCATCGGGTAA
	GTAAAACGACGGCCAGTCTTACACTTGCAAGCACCCA	CAGGAAACAGCTATGACCGCCTCGTGACCACTGTGTAA
	GTAAAACGACGGCCAGTCATCCACCAAGACACAATGC	CAGGAAACAGCTATGACCCTGTCACTGTTCCGGGTTTT
	GTAAAACGACGGCCAGTTCTTCTCCACAATTCCCCTG	CAGGAAACAGCTATGACCAGGGCCGTGTTTCCTAGATT
	GTAAAACGACGGCCAGTCACTCTTTTCAAACCCGGAG	CAGGAAACAGCTATGACCgcgcTAATCTCTTCCAGAGC
	GTAAAACGACGGCCAGTactgaggcccatcacttctg	CAGGAAACAGCTATGACCcattgtgtttgaggcgagtg
	GTAAAACGACGGCCAGTCTTCCCACAGAGGGATGTGT	CAGGAAACAGCTATGACCgaaCAGCTAAACGGCCAGAG
	GTAAAACGACGGCCAGTTACAATCACCCAGCCCTCTC	CAGGAAACAGCTATGACCAGCGGTCAATGATCCAAAAC
	GTAAAACGACGGCCAGTAGCCAAGTCCCTGACTCTCA	CAGGAAACAGCTATGACCAGCGGTCAATGATCCAAAAC
	GTAAAACGACGGCCAGTTTGAAGAATGGGGTACCTGC	CAGGAAACAGCTATGACCGGTGGGGGCATATTACACAG
	GTAAAACGACGGCCAGTtgcggtcatgcaCTCAGTAT	CAGGAAACAGCTATGACCGATCCTCTTCTCTCCCCCAC
ERCC6	GTAAAACGACGGCCAGTCCTGACAGCATCTTTTGGGT	CAGGAAACAGCTATGACCGCATCATGAAAAAGGAGGGA
	GTAAAACGACGGCCAGTCTCCTGCAGTATCTCCCTGG	CAGGAAACAGCTATGACCCTTGTGACCCTTCACAGCCT
	GTAAAACGACGGCCAGTTCTGCTCTAGCAGCCAGTGA	CAGGAAACAGCTATGACCcccagccAGCAGTTACTTTC
	GTAAAACGACGGCCAGTTATAAGCCTCCACACCTGCC	CAGGAAACAGCTATGACCGGAGAAAAGTGTCTTTCTTGCAT
	GTAAAACGACGGCCAGTGCCCTCTATGCACCATCAGT	CAGGAAACAGCTATGACCTCCAAAACGCAAGAAGTTCC
	GTAAAACGACGGCCAGTTTCCTTCAAACTTGGCGTCT	CAGGAAACAGCTATGACCTTTAGCCATTCATGCAGCAG
	GTAAAACGACGGCCAGTATCATGGTCTGCTCCAAAGG	CAGGAAACAGCTATGACCTTTAGCCATTCATGCAGCAG
	GTAAAACGACGGCCAGTCACACCATAGCAGGCAGAAA	CAGGAAACAGCTATGACCAGTGCTTGTCATTGGGAAGG
	GTAAAACGACGGCCAGTAACGTATGCTGAAGCCAACC	CAGGAAACAGCTATGACCGGAACAGCCCAGTAATTCCA
	GTAAAACGACGGCCAGTTGGAATTACTGGGCTGTTCC	CAGGAAACAGCTATGACCCAAACTCCTATCCCCCACCT
	GTAAAACGACGGCCAGTCCACCTCAGCATCAGTGGTA	CAGGAAACAGCTATGACCCCTCCTTGCCTAGGGAATCT
	GTAAAACGACGGCCAGTGGCTGACTGCACATCTACCA	CAGGAAACAGCTATGACCGAGCAGGAAGCCAGTCTTTG
	GTAAAACGACGGCCAGTGGGGCTTAAGGAGAATGGAG	CAGGAAACAGCTATGACCTTGCAATCATTTCCCACTGA
	GTAAAACGACGGCCAGTCCAGGAACAATGCCAAAACT	CAGGAAACAGCTATGACCTCTCTTGTAGGGGCCAGTTG
	GTAAAACGACGGCCAGTCAACTGGCCCCTACAAGAGA	CAGGAAACAGCTATGACCTCCTTCCCCTAATAACGGCT
	GTAAAACGACGGCCAGTTGTTTGAAGCCACTAATGCG	CAGGAAACAGCTATGACCTGTTTGTTGTTTTGGGGGAT
	GTAAAACGACGGCCAGTCCTTTGATGTCCCCCACTAA	CAGGAAACAGCTATGACCTGTGGTGTTCTCTGTCCTGC
	GTAAAACGACGGCCAGTTAGTCTGCTGAGGCAGGTCC	CAGGAAACAGCTATGACCGGTATGCCAGAAGAGGAGCA
	GTAAAACGACGGCCAGTTCAGCCCTCTGTATGCCTCT	CAGGAAACAGCTATGACCTGGCAGATCAAGCAAAACTG
	GTAAAACGACGGCCAGTATCTCCTGCACTGGCACTTT	CAGGAAACAGCTATGACCTTTCATGCTAGTGCGAAACG
	GTAAAACGACGGCCAGTGGATTTGGTTCTGCAGGTGT	CAGGAAACAGCTATGACCACGGCCCCTTTACTCCTAGA
	GTAAAACGACGGCCAGTTGGGATGTTAAAGGAAAACAGC	CAGGAAACAGCTATGACCtgaaGCAGCTTACAAATGGG
	GTAAAACGACGGCCAGTAAGGGCTGGGACTCAAAGAT	CAGGAAACAGCTATGACCGAGAATGCCAAATGAGGGAA
	GTAAAACGACGGCCAGTTCATAGACGTCCACACCCAA	CAGGAAACAGCTATGACCGTGGCTGGCTTTCTTAGCTG
KIAA1919	GTAAAACGACGGCCAGTCGGTCTCCTAGGCAACCATA	CAGGAAACAGCTATGACCaaGCTTGCTGGAGCTGAAAG
	GTAAAACGACGGCCAGTTTGTAAATCCCAGACCCTCC	CAGGAAACAGCTATGACCCAACCAGTGTGTGATCCAGG
	GTAAAACGACGGCCAGTCCTGATTTTGCATGTTTGGA	CAGGAAACAGCTATGACCTCTGTGTCATTTTAGGGGGC
	GTAAAACGACGGCCAGTGCGTCTTGCTAATGTGCTTG	CAGGAAACAGCTATGACCCATGGATGGTCGTGTACTGC

	GTAAAACGACGGCCAGTCGACAGGTGGTAACGTCCTT	CAGGAAACAGCTATGACCTCCATTTTTCTGCATCCTCC
	GTAAAACGACGGCCAGTTGCTGGCATGAAAGAAAGTG	CAGGAAACAGCTATGACCCTGAGGGGTGGAGACCATAA
	GTAAAACGACGGCCAGTGGAGGATGCAGAAAAATGGA	CAGGAAACAGCTATGACCAACAATGCCTTCTTTGGGTG
SLC39A12	GTAAAACGACGGCCAGTCTGCCATTTGGGGTTAGAAA	CAGGAAACAGCTATGACCAAAAGTTCGCGCTCACAAAT
	GTAAAACGACGGCCAGTAAGCAATGAAGACCCTGGTG	CAGGAAACAGCTATGACCAAACCCACAAATACTTACCTCCC
	GTAAAACGACGGCCAGTAAGGACCAGTGAGATGAGGG	CAGGAAACAGCTATGACCTCAAGAGCTTTGGCCTTTGT
	GTAAAACGACGGCCAGTGGGTGGAAATGTCCATTCTG	CAGGAAACAGCTATGACCTCAATCTGAACCCAAGGGAG
	GTAAAACGACGGCCAGTAGCTGAAAGTTTGCAGGGAA	CAGGAAACAGCTATGACCGCATACATTGTGCACATCAGC
	GTAAAACGACGGCCAGTAATTCTTCTGGCATTTTCGC	CAGGAAACAGCTATGACCAGAGAGGCCTTGGGGAAATA
	GTAAAACGACGGCCAGTTGGAAGCATTCACATACCCA	CAGGAAACAGCTATGACCCCCTTACACAGGCAACTTCA
	GTAAAACGACGGCCAGTCTTGCTTTGCCCACAGACTT	CAGGAAACAGCTATGACCTCAATTTGAGAGCTGGATAGGA
	GTAAAACGACGGCCAGTACAGACGGGAAGGTCATCAG	CAGGAAACAGCTATGACCGCCCTGGGAAACATCATAAA
	GTAAAACGACGGCCAGTTGGTTCCAGTGCTGACATTC	CAGGAAACAGCTATGACCCCAGAATGCACTTTCTTGGAG
	GTAAAACGACGGCCAGTACAGTCGCTCATGTTTTCCC	CAGGAAACAGCTATGACCTCTCCCAAAATGGCTTGTTC
TET2	GTAAAACGACGGCCAGTCACCCTTGTTCTCCATGACC	CAGGAAACAGCTATGACCTGGTTGACTGCTTTCACCTG
	GTAAAACGACGGCCAGTAAATGGAGACACCAAGTGGC	CAGGAAACAGCTATGACCGAGGTATGCGATGGGTGAGT
	GTAAAACGACGGCCAGTATGAGCAGGAGGGGAAAAGT	CAGGAAACAGCTATGACCTGGTGTGGTAGTGGCAGAAA
	GTAAAACGACGGCCAGTACTCACCCATCGCATACCTC	CAGGAAACAGCTATGACCAGATAGTGCTGTGTTGGGGG
	GTAAAACGACGGCCAGTTTCCACAGGTTCCTCAGCTT	CAGGAAACAGCTATGACCGAGAAGTGCACCTGGTGTGA
	GTAAAACGACGGCCAGTAAGGCAAGCTTACACCCAGA	CAGGAAACAGCTATGACCGGTTCCACCTTAATTGGCCT
	GTAAAACGACGGCCAGTAATGTCCAAATGGGACTGGA	CAGGAAACAGCTATGACCACTGGCCCTGACATTTCAAC
	GTAAAACGACGGCCAGTCCCCAGAAGGACACTCAAAA	CAGGAAACAGCTATGACCCAAATTGCTGCCAGACTCAA
	GTAAAACGACGGCCAGTACTTGATAGCCACACCCCAG	CAGGAAACAGCTATGACCTTCCCCCAACTCATGAAGAC
	GTAAAACGACGGCCAGTtgcacaaaaggtagaatgcaa	CAGGAAACAGCTATGACCacgtgggatttcacacaaca
	GTAAAACGACGGCCAGTTTTCCCATTTTCACCCACAT	CAGGAAACAGCTATGACCACCCAATTCTCAGGGTCAGA
	GTAAAACGACGGCCAGTAGGGTCAAAGCCCACTTTTT	CAGGAAACAGCTATGACCTGAGGCCATGTGGTTACAGA
	GTAAAACGACGGCCAGTGTGTGGTTATGCCACAGCTT	CAGGAAACAGCTATGACCCCAAAGAGGAAGTTTTTGTTGC
	GTAAAACGACGGCCAGTACCATACGGCTTAATTCCCC	CAGGAAACAGCTATGACCTGTTACAATTGCTGCCAATGA
	GTAAAACGACGGCCAGTTGTCATTCCATTTTGTTTCTGG	CAGGAAACAGCTATGACCCTGCTAAGCTGTCCTCAGCC
	GTAAAACGACGGCCAGTTCTGGATCAACTAGGCCACC	CAGGAAACAGCTATGACCGGGGGCAAAACCAAAATAAT
	GTAAAACGACGGCCAGTTCAAGCAGAGGCATGTTCAG	CAGGAAACAGCTATGACCTATTTCCAAACCTTGGCTGG
	GTAAAACGACGGCCAGTAATCCCATGAACCCTTACCC	CAGGAAACAGCTATGACCACCAGACCTCATCGTTGTCC
	GTAAAACGACGGCCAGTATCAGTGGACAACTGCTCCC	CAGGAAACAGCTATGACCATGAAACGCAGGTAAGTGGG
	GTAAAACGACGGCCAGTATTGGCACTAGTCCAGGGTG	CAGGAAACAGCTATGACCACTGTGACCTTTCCCCACTG

1- Reference sequences used for primer design were as follows: *BAIAP2* NM_006340, *DNMT3A* NM_175629, ERCC6 NM_000124, KIAA1919 NM_153369, *SCAND2* NR_004859*,* SLC39A12 NM_001145195, *TET2* NM_001127208

Supplementary Table 3. Comparison of *TET2* mutant allele burden and degree of polymorphonuclear cell (PMN) skewing in normal individuals with clonal hematopoiesis and somatic *TET2* mutations.

TET2 mutation	*TET2* mutant allele burden[1]	Degree of PMN skewing[2]
TET2 p.Phe1104Leufs*3	40%	0.41
TET2 p.Arg544*	50%	0.49
TET2 p.Glu1909*	38%	0.48
TET2 p.Thr1331Pro	22%	0.41

1- Quantification of mutant *TET2* allele burden was determined using Ion Torrent PGM sequencing.
2- Degree of PMN skewing was performed by determination of X-inactivation ratios at the *HUMARA* locus as explained in the Supplemental Methods.

Supplementary Table 4: Comparison of clinical characteristics and hematological parameters between *TET2* mutated elderly individuals (n=10) and age-matched *TET2* wildtype controls (n=34).

Characteristic	*TET2* genotype	Mean	Std Dev.	Std Error	95% LCL of Mean[1]	95% UCL of Mean[2]	*P* value[3]
Age	WT	79.00	4.94	0.85	77.28	80.72	0.553
	mut	80.10	5.74	1.82	75.99	84.21	
PMN skewing(DS)[4]	WT	0.29	0.15	0.03	0.24	0.35	**0.022**
	mut	0.41	0.07	0.02	0.36	0.46	
Neutrophils (%)	WT	0.61	0.09	0.01	0.58	0.65	0.672
	mut	0.60	0.14	0.05	0.50	0.70	
Neutrophils, absolute count	WT	3.86	1.10	0.19	3.47	4.24	0.999
	mut	3.86	1.55	0.49	2.75	4.97	
Hemoglobin	WT	125.32	10.30	1.77	121.73	128.92	0.742
	mut	126.70	15.19	4.80	115.84	137.56	
Platelets	WT	236.97	60.10	10.31	216.00	257.94	0.900
	mut	233.90	89.44	28.28	169.92	297.88	
White blood cell count	WT	6.25	1.38	0.24	5.77	6.73	0.756
	mut	6.43	2.15	0.68	4.89	7.97	
Lymphocytes (%)	WT	0.27	0.09	0.01	0.23	0.30	0.401
	mut	0.30	0.13	0.04	0.20	0.39	
Lymphocytes, absolute count	WT	1.65	0.63	0.11	1.43	1.87	0.382
	mut	1.88	1.00	0.32	1.16	2.60	
Monocytes (%)	WT	0.08	0.03	0.00	0.07	0.09	0.897
	mut	0.08	0.04	0.01	0.06	0.11	
Monocyte, absolute count	WT	0.52	0.23	0.04	0.44	0.60	0.732
	mut	0.55	0.34	0.11	0.30	0.80	
Eosinophils (%)	WT	0.03	0.02	0.00	0.02	0.03	0.125
	mut	0.02	0.02	0.01	0.01	0.03	
Eosinophil, absolute count	WT	0.17	0.13	0.02	0.13	0.22	0.201
	mut	0.11	0.13	0.04	0.02	0.20	
MCV	WT	91.67	4.12	0.71	90.24	93.11	0.540
	mut	92.65	5.31	1.68	88.85	96.45	

1) LCL – lower confidence limit.
2) UCL – upper confidence limit.
3) *P*-values represent 2-tailed p-values as calculated by Student's T-test.
4) DS represents degree of skewing as described in the Supplemental Methods.

References:

1. Bolduc, V. et al. No evidence that skewing of X chromosome inactivation patterns is transmitted to offspring in humans. *J Clin Invest* **118**, 333-41 (2008).
2. Homer, N., Merriman, B. & Nelson, S.F. BFAST: an alignment tool for large scale genome resequencing. *PLoS One* **4**, e7767 (2009).
3. Homer, N., Merriman, B. & Nelson, S.F. Local alignment of two-base encoded DNA sequence. *BMC Bioinformatics* **10**, 175 (2009).
4. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297-303 (2010).
5. DePristo, M.A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* **43**, 491-8 (2011).
6. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-9 (2009).
7. Figueroa, M.E. et al. DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia. *Cancer Cell* **17**, 13-27 (2010).
8. Figueroa, M.E. et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell* **18**, 553-67 (2010).
9. Ehrich, M. et al. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. *Proc Natl Acad Sci U S A* **102**, 15785-90 (2005).