Supplemental Appendix
Supplement to: Takahashi K, Wang F, Kantarjian H, et al. Pre-leukemic clonal hematopoiesis increases the risk of therapy-related myeloid neoplasms

Contents

Supplemental Method
Supplemental Table S1
Supplemental Table S2
Supplemental Table S3
Supplemental Table S4
Supplemental Table S5
Supplemental Table S6
Supplemental Table S7
Supplemental Table S8
Supplemental Figure S1
Supplemental Figure S2
Supplemental Figure S3
Supplemental Figure S4
Supplemental Figure S5
Supplemental References
Supplemental Method

Case description of 14 patients with t-MNs

Case UID12766
55 year-old male was diagnosed with T3N2M0 Stage III colon adenocarcinoma and underwent for partial colectomy. He received adjuvant chemotherapy with FOLFOX regimen (5-FU, oxaliplatin and leucovorin) for total of 6 months. Approximately 4 years after diagnosis of his colon cancer, he developed pancytopenia. Bone marrow biopsy revealed therapy-related acute myeloid leukemia (t-AML) with 55% blast. Cytogenetic analysis showed normal diploid karyotype. He received induction therapy with cladribine, idarubicin and cytarabine under clinical trial but on day 19, he developed intracranial hemorrhage and died on day 23.

Case UID984
64 year-old male was diagnosed with T3N1M0 distal esophageal adenocarcinoma and received neoadjuvant concurrent chemoradiation therapy with 5-FU and docetaxel and 50 Gy of radiation then underwent total esophagectomy. Approximately 7 years after diagnosis of esophageal cancer, he developed neutropenia. Bone marrow biopsy revealed therapy-related myelodysplastic syndromes (t-MDS) with 1% blast. Cytogenetic analysis showed complex abnormality including deletion 7q. The patient did not receive any treatment for t-MDS and was observed. He did not require any transfusion. Approximately 2 years after diagnosis of t-MDS, the patient died of unknown cause.

Case UID10164
50 year-old female was diagnosed with malignant peripheral nerve sheath tumor of the left chest wall. She received neoadjuvant chemotherapy with doxorubicin and ifosfamide for 2 cycles, however had progressive disease and underwent mass resection. Approximately 1 year after her initial diagnosis of malignant peripheral nerve sheath tumor, she developed thrombocytopenia. Bone marrow biopsy revealed t-MDS with 4% blast. Cytogenetic analysis showed trisomy 1 and der (1; 7)(q10; p10). The patient died of unknown cause 2 months after diagnosis of t-MDS.

Case UID6982
70 year-old male was diagnosed with limited stage small cell lung cancer (SCLC). He received concurrent chemo-radiation therapy with 4 cycles of cisplatin and etoposide and 45 Gy of radiation. He then received prophylactic cranial radiation. Approximately 3 years after initial diagnosis of SCLC, he developed leukocytosis, anemia and thrombocytopenia. Bone marrow biopsy revealed t-AML with 40% blast. Cytogenetic analysis showed del 5q. He received therapy with azacitidine and vorinostat for 6 cycles with reduction of bone marrow blast but did not achieve complete remission. Approximately 6 months after the diagnosis of t-AML, he suffered from subarachnoid hemorrhage and died.

Case UID488
47 year-old female was diagnosed with stage IIIA non-small cell lung cancer (NSCLC) and received neoadjuvant chemotherapy with carboplatin and paclitaxel for 3 cycles. Then she underwent for right lobectomy followed by 56 Gy of adjuvant radiation therapy. While she was receiving radiation therapy, she was found to have metastasis in the contralateral side of the lung. This metastasis was treated with gemcitabine and vinorelbine for 2 cycles and then she underwent left upper lobe wedge resection. Approximately 8 years from the initial NSCLC diagnosis, she started to develop leucopenia. Bone marrow biopsy revealed t-MDS with 3% blast. Cytogenetic analysis showed normal female karyotype. She was initially treated with GCSF injection but had no response. She was then treated with azacitidine for 3 cycles but her MDS progressed to AML. She then received therapy with idarubicin and cytarabine but had no response. She received salvage therapy with fludarabine, cytarabine and gemtuzumab ozogamicin but developed pneumonia and subsequently died with respiratory failure.

Case UID36491
62 year-old male was diagnosed with limited stage SCLC. He received stereotactic radiation therapy with 50 Gy followed by 4 cycles of cisplatin and etoposide chemotherapy. Approximately 3 years after initial diagnosis of SCLC, he developed a second primary NSCLC. This was treated with concurrent chemo-radiation therapy with 70 Gy of radiation and weekly carboplatin and paclitaxel. Approximately 4 years after initial diagnosis of SCLC, he developed thrombocytopenia. Bone marrow biopsy revealed t-MDS with 3% blast and cytogenetic analysis showed monosomal karyotype. He received therapy with azacitidine and had hematological improvement. However, after 5 cycles of therapy, he lost response and bone marrow blast increased to 13%. He was then treated with clofarabine and low dose cytarabine for 1 cycle but his overall condition deteriorated and patient opted to pursue palliative care.

Case UID4473
44 year-old male was diagnosed with T2N0M0 rectal adenocarcinoma. He was treated with concurrent chemo-radiation therapy with 50 Gy of radiation and capecitabine. Post therapy, he had prolonged neutropenia. Approximately 6 years after the initial diagnosis of rectal cancer, neutropenia worsened. Bone marrow biopsy revealed t-MDS with 5% blast. Cytogenetic analysis showed normal male karyotype. He has been observed since and is alive at the time of this manuscript writing.
Case UID17285
69 year-old male was diagnosed with stage IIIA Hodgkin’s lymphoma. He was treated with 6 cycles of ABVD (doxorubicin, bleomycin, vinblastine, and dacarbazine) chemotherapy and achieved complete response. Approximately 2 years after the initial diagnosis of lymphoma, patient developed mild pancytopenia. Bone marrow biopsy revealed t-MDS with 3% blast. Cytogenetic analysis showed monosomal karyotype including monosomy 5 and 7. He was treated with decitabine for 3 cycles but developed pneumonia and subsequently died with respiratory failure.

Case UID19684
63 year-old male was diagnosed with stage III follicular lymphoma. There was no bone marrow involvement at initial diagnosis. He was treated with R-CHOP chemotherapy (rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone) for 6 cycles and achieved complete response. Approximately 3 years after initial diagnosis of follicular lymphoma, he developed mild cytopenia. Bone marrow biopsy revealed t-MDS with 8% blast and cytogenetic analysis showed monosomy 7. He was treated with lenalidomide and eltrombopag. After 5 months, his bone marrow showed increase in blast count and then he was treated with azacitidine and vorinostat for 3 cycles with stable disease. He then underwent cord blood stem cell transplant but he died at day 20 of transplant because of sepsis and multi-organ failure.

Case UID7394
74 year-old male was diagnosed with squamous cell carcinoma of penis. He was treated with cisplatin, ifosfamide and paclitaxel for 4 cycles followed by 65 Gy of radiation therapy. Approximately 3 years later, he was noted to have pancytopenia. Bone marrow biopsy revealed t-MDS with 3% blast and cytogenetic analysis showed complex abnormality. He was treated with azacitidine and vorinostat for 12 cycles and had hematological improvement. He then received single agent azacitidine for additional 4 cycles but lost response and progressed to AML. Subsequently, he developed respiratory failure and died approximately 1.5 years after the diagnosis of t-MDS.

Case UID49278
63 year-old male was diagnosed with stage I NSCLC. He was treated with concurrent chemo-radiation therapy with 66 Gy of radiation and weekly carboplatin and pemetrexed. Approximately 1 year after the initial diagnosis of NSCLC, he developed leukocytosis. Bone marrow biopsy revealed t-AML with 88% blast. Cytogenetic analysis showed –Y. He was treated with cladribine and low dose cytarabine and achieved complete remission after 1 cycle. He received total 9 cycles of this regimen and continues to maintain complete remission.

Case UID12484
64 year-old male was diagnosed with stage IV mantle cell lymphoma (MCL) and was treated with R-Hyper-CVAD chemotherapy (rituximab, cyclophosphamide, vincristine, doxorubicin, and dexamethasone alternating with methotrexate and cytarabine) for 5 cycles and achieved complete response. One year later, he had relapse and was treated with ibrutinib and rituximab. He again achieved complete response and received this therapy for 5 cycles. Approximately 2 years after the initial diagnosis of MCL, he developed pancytopenia. Bone marrow biopsy revealed t-AML with 28% blast. Cytogenetic analysis revealed der (7;17)(p10;q10). He was treated with azacitidine and vorinostat for 3 cycles but with no response. He developed pneumonia and died approximately 4 months after the diagnosis of t-AML.

Case UID19304
25 year-old female was diagnosed with glioblastoma multiforme (GBM). She had tumor resection and received adjuvant chemotherapy with temozolamide for 12 months. Approximately 3 years from the initial diagnosis of GBM, she developed pancytopenia. Bone marrow biopsy revealed t-MDS with 13% blast and cytogenetic analysis showed inversion 3q. She was treated with fludarabine, idarubicin and cytarabine and achieved complete remission after 1 cycle of therapy and received 1 cycle of consolidation therapy with the same chemotherapy. She then underwent matched unrelated donor stem cell transplant. At the time of this manuscript writing, she is day 84 of stem cell transplant and is doing well.

Case UID31000
40 year-old male was diagnosed with rhabdomyosarcoma in the right gluteus. He received neo-adjuvant chemotherapy consists of doxorubicin and ifosfamide for total 6 cycles and then received radiation therapy of 50 Gy in 25 fractions. He then underwent for radical resection of the mass. Approximately 2.5 years from the initial diagnosis of rhabdomyosarcoma, he developed pancytopenia. Bone marrow biopsy revealed t-AML with 31% myeloblast and cytogenetic analysis showed complex karyotype. He was treated with induction regimen consists of cladribine, idarubicin, and cytarabine and achieved complete remission. He received 1 cycle of consolidation therapy and underwent for matched unrelated donor stem cell transplant. He received 1 cycle of post-transplant azacitidine but at day 60, his bone marrow confirmed relapse of t-AML. He received salvage therapy with SGN-CD33a under clinical trial but with no response. He then received 1 cycle of decitabine but with no response. He developed pneumonia and died approximately 9 months after the diagnosis of t-AML. Of particular note, he was evaluated at familial cancer clinic when he was diagnosed with t-AML because his father and sister died from unknown cancer. He was confirmed to have Li-Fraumeni Syndrome because of the findings with 2 germline TP53 mutations.
Targeted gene sequencing of t-MN bone marrow samples and detection of high-confidence driver mutations

Genomic DNA was extracted from diagnostic BM aspirate samples using an Autopure extractor (QIAGEN/Gentra, Valencia, CA). DNAs were fragmented and bait-captured in solution as previously described according to manufacturer protocols.(1) Captured DNA libraries were then sequenced using a HiSeq 2000 sequencer (Illumina, San Diego, CA) with 76 basepair (bp) paired-end reads. Raw sequencing data from the Illumina platform were converted to a fastq format and aligned to the reference genome (hg19) using the Burroughs-Wheeler Aligner (BWA).(2) The aligned BAM files were subjected to mark duplication, re-alignment, and recalibration using Picard and GATK (https://www.broadinstitute.org/gatk/guide/best-practices?bpm=DNAseq, last accessed 9/29/2016). Preprocessed BAM files were then analyzed to detect single nucleotide variants (SNV) and small insertions and deletions (indels) using MuTect(3) and Pindel(4) algorithms, respectively, against virtual normal sequence developed in-house.

We modified an approach described by Pappamanuil et al. to identify high-confidence driver mutations in the bone marrow samples without matched germline control.(5) First, variants with low quality supporting sequencing data were filtered out. Specifically, variants matching one or more of the following criteria were considered of low quality and therefore filtered out from further analysis: 1) tumor coverage < 15x and 20x for single nucleotide variants (SNVs) and insertions/deletions (indels), respectively, 2) tumor allele frequency < 5%, and 3) normal allele frequency >= 1% and 0% for SNVs and INDELs, respectively. Second, only variants which would introduce an obvious protein-coding change were kept for further analysis. Specifically, variant with an ANNOVAR annotation of non-synonymous, stop-gain, stop-loss, splicing, frameshift insertion, frameshift deletion, non-frameshift insertion or non-frameshift deletion was considered to be able to introduce an obvious protein-coding change and were therefore kept for further analysis. Third, common polymorphisms were removed to reduce the load of possible germline contamination due to the absence of matched normal. Specifically, a series of public variant database including the 1000 Genome Database (http://www.1000genomes.org/, last accessed 9/29/2016), ESP6500 Database (http://evs.gs.washington.edu/EVS/, last accessed 9/29/2016), dbSNP ver.132 (https://www.ncbi.nlm.nih.gov/projects/SNP/snp_summary.cgi?build_id=132, last accessed 9/29/2016), and Exome Aggregation Consortium database (http://exac.broadinstitute.org/, last accessed 9/29/2016), were utilized. Variant with a population frequency of 0.14% or more in any of the databases was considered possible germline polymorphism and was therefore removed from further analysis. Finally, hierarchical classification system was developed to assign confidence level for each remaining variant in order to facilitate the identification of putative driver mutations. Specifically, each variant was classified based on the following hierarchical order and was assigned a confidence level corresponding to its rank in the system: 1) Confirmed somatic mutation based on COSMIC database (version 72), 2) loss-of-function mutation such as splicing, stop-gain, stop-loss and frameshift mutation in known tumor suppressor genes, 3) recurrent variant which resides within three amino acids away from a confirmed somatic mutation according to COSMIC database (version 72), 4) variant which resides within three amino acids away from a confirmed somatic mutation according to COSMIC database (version 72), 5) variant which was predicted to be damaging by in-silico function prediction algorithms, and 6) variant with unknown significance. The first three groups were considered high-confidence tier while the remaining were considered of low-confidence. The final annotated variant list was passed out of the pipeline and was further analyzed by manual inspection and literature mining in order to identify high-confidence driver mutations.

Molecular barcode sequencing using Haloplex High Sensitivity Assay

Illumina compatible libraries were prepared using Agilent’s HaloPlex HS Target Enrichment System. Briefly, 50 ng of genomic DNA was fragmented using 16 restriction enzymes, in a double digestion format. The restriction fragments were pooled and hybridized to our custom biotinylated HaloPlex HS probe library. Molecular barcodes, Illumina index sequences and Illumina adapters were incorporated into the targeted fragments during hybridization. Nicks in the circularized probe-target hybridization products were then closed by DNA ligation. The circularized target-DNA-HaloPlex HS probe hybrids were captured on streptavidin beads and non-circular fragments were removed. The circular libraries were amplified by 24 cycles of PCR, purified using AMPure XP beads (Beckman Coulter), quantified fluorometrically using the Qubit™ dsDNA HS Assay (ThermoFisher). Enrichment was validated using the Agilent 2100 Bioanalyzer. The libraries were multiplexed 14 samples per pool and sequenced on the Illumina HiSeq4000 using the 150nt paired end format.

Rare variant calling and detection of clonal hematopoiesis

De-multiplexed sequence reads were first analyzed using the SureCall pipeline ver. 3.5 available from Agilent’s website (http://www.genomics.agilent.com/en/NGS-Data-Analysis-Software/SureCall/?cid=AG-PT-154&tabld=AG-PR-1196, last accessed 9/29/2016). Briefly, reads were first trimmed and aligned to the human hg19 reference genome using BWA. Reads with the same molecular barcode sequence and mapped to the same genomic coordinates were considered to be derived from the same read family. For each read family, the consensus sequence was generated after removing random errors, which occurred in the minority of the reads within the read family. Read families with less than three unique reads were removed from further analysis. The aligned consensus reads were then subjected to variants calling using VarScan (http://varscan.sourceforge.net/, last accessed 9/29/2016). A binomial test was performed to calculate P-value for each variant by considering observed VAF and expected substitution rate at the given nucleotide position based on the entire sequenced population. After removing possible germline variants, the following filtering criteria were used: 1) alternative read count (Read2) ≥ 5, 2) binomial P-value < 1.0 x 10-6, 3) both strands called the variant, and 4) observed VAF is higher than the mean VAF for entire population at the given position. To be consistent with the prior definition of clonal hematopoiesis in healthy individuals, variants that passed the above filtering criteria were then annotated as clonal hematopoiesis if the variant matched with the list specified by Jaiswal et al.(6)
**Table S1.** List of 295 genes targeted by next generation sequencing.

| Gene name | Gene name | Gene name | Gene name | Gene name | Gene name | Gene name | Gene name | Gene name | Gene name | Gene name | Gene name | Gene name | Gene name | Gene name | Gene name | Gene name | Gene name | Gene name | Gene name | Gene name | Gene name | Gene name |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| ABCC9     | CALR      | CUL5      | FANCD2    | HIST1H2BF | LEF1      | NBN       | PLA2G2D   | SF3B1     | TINF2     |          |          |          |          |          |          |          |          |          |          |          |          |
| ABL1      | CARD11    | CUX1      | FANCE     | HIST1H3D  | LRP1B     | NCO2      | POT1      | SFRS1     | TLR2      |          |          |          |          |          |          |          |          |          |          |          |          |
| ACTG1     | CBL       | CYLD      | FANCG     | HIST1H4D  | LTB       | NCO2      | POT1      | SFRS7     | TLR9      |          |          |          |          |          |          |          |          |          |          |          |          |
| AKT1      | CBLB      | DAXX      | FANC1     | HNRNPK    | LUC7L2    | NF1       | POU2AF1   | SGK1      | TNFAIP3   |          |          |          |          |          |          |          |          |          |          |          |          |
| ANKRD11   | CCND1     | DCLRE1C   | FANCL     | HRAS      | LYN       | NFE2      | PRDM1     | SH2B3     | TNRFSF14  |          |          |          |          |          |          |          |          |          |          |          |          |
| ARID1A    | CCND3     | DDX3X     | FAS       | ICOS      | MALT1     | NFKB1     | PRKCB     | SHH       | TNKS      |          |          |          |          |          |          |          |          |          |          |          |          |
| ARID1B    | CD200     | DUS3      | FAT1      | ID3       | MAP2K1    | NFKB2     | P TEN     | SMAD2     | TOX       |          |          |          |          |          |          |          |          |          |          |          |          |
| ARID2     | CD274     | DCK1      | FAT3      | IDH1      | MAPK1     | NFKB1A    | PTPN1     | SMC1A     | TP53      |          |          |          |          |          |          |          |          |          |          |          |          |
| ARID5B    | CD58      | DLM1      | FBXW7     | IDH2      | MAX       | NFKBIE    | PTPN11    | SMC3      | TRAF3     |          |          |          |          |          |          |          |          |          |          |          |          |
| ARPP21    | CD79A     | DNM2      | FGFR3     | IKBKA     | MDM2      | NOTCH1    | RAD21     | SMC5      | TRAF6     |          |          |          |          |          |          |          |          |          |          |          |          |
| ASXL1     | CD79B     | DNMT1     | FLI1      | IKZF1     | MED12     | NOTCH2    | RAD51C    | SNX7      | TYK2      |          |          |          |          |          |          |          |          |          |          |          |          |
| ATF7IP    | CDK4      | DNTM3A    | FLT3      | IKZF2     | MEF2B     | NPM1      | RAG1      | SOCS1     | TYK3      |          |          |          |          |          |          |          |          |          |          |          |          |
| ATM       | CDKN2A    | DNM3B     | FNDC3A    | IKZF3     | MEF2C     | NR3C2     | RAG2      | SOX5      | U2AF1     |          |          |          |          |          |          |          |          |          |          |          |          |
| ATRX      | CDKN2B    | EBF1      | FOXP1     | IL7R      | MGA       | N RAS     | RAS42     | SP140     | U2AF2     |          |          |          |          |          |          |          |          |          |          |          |          |
| B2M       | CDKN2C    | ECT2L     | FYN       | IRAK1     | mir125a   | NSD2      | RB1       | SPEN      | UBR5      |          |          |          |          |          |          |          |          |          |          |          |          |
| BCL10     | CEBPA     | EED       | G6PC3     | IRAK4     | mir142    | NT5C2     | REL       | SPIB      | USP29     |          |          |          |          |          |          |          |          |          |          |          |          |
| BCL12     | CEBPE     | EGR1      | GAB2      | IRF1      | mir155    | PAG1      | RELA      | SRSF2     | VPREB1    |          |          |          |          |          |          |          |          |          |          |          |          |
| BCL6      | CHD2      | EGR2      | GATA1     | IRF4      | mir15a    | PALB2     | RELB      | STAG1     | WHSC1L1   |          |          |          |          |          |          |          |          |          |          |          |          |
| BCL7A     | CHK2      | ELANE     | GATA2     | IRF7      | mir16-1   | PAX5      | RELN      | STAG2     | WHSC1L1   |          |          |          |          |          |          |          |          |          |          |          |          |
| BCR       | CIITA     | EP300     | GATA3     | ITPKB     | MIR17HG   | PDCD1     | RHOA      | STAT1     | WT1       |          |          |          |          |          |          |          |          |          |          |          |          |
| BIRC3     | CREBBP    | EPOR      | GF11B     | JAK2      | mir34b    | PDGFRB    | ROBO1     | SUZ12     | ZAP70     |          |          |          |          |          |          |          |          |          |          |          |          |
| BLK       | CRLF2     | ERG       | GNA13     | JAK3      | mir34c    | PEG3      | ROR1      | SYK       | ZMYM2     |          |          |          |          |          |          |          |          |          |          |          |          |
| BM11      | CSF2RA    | ETV6      | GNAS      | JARID2    | MLL       | PHF6      | RPL10     | TBL1XR1   | ZMYM3     |          |          |          |          |          |          |          |          |          |          |          |          |
| Braf      | CSF3R     | EZH2      | GNB1      | KDM4C     | MLL2      | PHIP      | RPL5      | TCF3      | ZKSR2     |          |          |          |          |          |          |          |          |          |          |          |          |
| BRIP1     | CTBP1     | FAM46C    | GPBC5A    | KDM6A     | MLL3      | PIGA      | RUNX1     | TERC      |          |          |          |          |          |          |          |          |          |          |          |          |          |
| BTG1      | CTBP2     | FAM5C     | HAX1      | KIT       | MPL       | PIK3CA    | RUNX2     | TERT      |          |          |          |          |          |          |          |          |          |          |          |          |          |
| BTK       | CTCF      | FANC4     | HIST1H1E  | KLHL6     | MS4A1     | PIK3CB    | SAMHD1    | TET1      |          |          |          |          |          |          |          |          |          |          |          |          |          |
| BTLA      | CTLA4     | FANCB     | HIST1H2AD | KRAS      | MYB       | PIK3CG    | SETBP1    | TET2      |          |          |          |          |          |          |          |          |          |          |          |          |          |
| C22orf194 | CTNNA1    | FANCC     | HIST1H2BE | LAMB4     | MYD88     | PIK3R1    | SETD2     | TGDS      |          |          |          |          |          |          |          |          |          |          |          |          |          |
| C2orf63   | CTNNA2    | FANCG     | HIST1H2F  | LAMB5     | MYD89     | PIK3R2    | SETD3     | TGFS      |          |          |          |          |          |          |          |          |          |          |          |          |          |
Table S2. List of 32 genes targeted by Haloplex HS molecular barcode sequencing.

| Gene  | Gene  | Gene  | Gene  |
|-------|-------|-------|-------|
| ASXL1 | GATA2 | KRAS  | SRSF2 |
| ATM   | GNAS  | MYD88 | STAG2 |
| BCOR  | GNB1  | NRAS  | STAT3 |
| CBL   | IDH1  | PHF6  | TET2  |
| CUX1  | IDH2  | PTPN11| TP53  |
| DNMT3A| JAK2  | RAD21 | U2AF1 |
| EZH2  | KDM6A | RUNX1 | WT1   |
| FLT3  | KIT   | SF3B1 | ZRSR2 |
Table S3 Comparison of clinical characteristics of 14 patients who developed t-MNs (cases) and 54 patients with lymphoma who did not develop t-MNs after at least 5 years of follow up (control).

|                      | Cases (N = 14) | Control (N = 54) | P-value |
|----------------------|---------------|------------------|---------|
|                      | N (IQR or %)  | N (IQR or %)     |         |
| Median age, years    | 62 (46 - 65)  | 58 (49 - 63)     | 0.62    |
| Female               | 4 (29)        | 24 (44)          | 0.21    |
| Follow up duration, years | 8.9 (95% CI: 6.3-11.5) | 6.1 (95% CI: 6.0-6.3) | 0.001 |
| Primary Cancer Types | NA            |                  |         |
| GI                   | 3 (21)        | 0                |         |
| Sarcoma              | 2 (14)        | 0                |         |
| Lung                 | 3 (21)        | 0                |         |
| Brain                | 1 (7)         | 0                |         |
| GU/Prostate          | 2 (14)        | 0                |         |
| Lymphoma             | 3 (21)        | 54 (100)         |         |
| Frontline chemotherapy |               |                  |         |
| Alkylating agent containing | 10 (71) | 54 (100) | 0.001 |
| Topoisomerase II inhibitor containing | 6 (43) | 3 (6) | 0.002 |
| Radiation            | 9 (64)        | 14 (26)          | 0.01    |
| Auto SCT             | 0             | 6 (11)           | 0.24    |

GI: Gastrointestinal, GU: Genitourinal, SCT: stem cell transplant.
Table S4. Mutations detected as clonal hematopoiesis in pre-treatment PB samples from 54 patients with lymphoma who did not develop t-MNs.

| UID   | Age | Lymphoma Diagnosis | Gene       | AA change     | VAF (%) |
|-------|-----|-------------------|------------|---------------|---------|
| UID12 | 58  | DLBCL             | CBL        | p.R420G       | 0.35%   |
| UID19 | 79  | DLBCL             | TET2       | p.Q632fs      | 0.59%   |
| UID41 | 59  | HL                | TET2       | p.D1402fs     | 0.74%   |
| UID51 | 49  | DLBCL             | DNMT3A     | p.F751fs      | 0.34%   |
| UID59 | 49  | FL                | DNMT3A     | p.R771X       | 0.39%   |
| UID81 | 58  | FL                | DNMT3A     | p.R366P       | 1.80%   |
| UID82 | 43  | DLBCL             | DNMT3A     | p.R882H       | 0.92%   |
| UID95 | 60  | DLBCL             | ASXL1      | p.T880fs      | 0.90%   |
|        |     |                   | PHF6       | p.G186X       | 1.23%   |
| UID108| 71  | MCL               | STAT3      | p.D661V       | 1.12%   |
| UID113| 52  | DLBCL             | DNMT3A     | p.M513fs      | 0.45%   |
| UID117| 69  | DLBCL             | DNMT3A     | p.Q237fs      | 1.20%   |
|        |     |                   | DNMT3A     | c.1554+1G>A (splice site) | 1.32% |
| UID130| 51  | DLBCL             | DNMT3A     | p.Y735C       | 1.29%   |
| UID190| 43  | FL                | TET2       | p.K188fs      | 0.74%   |
|        |     |                   | DNMT3A     | p.Y735C       | 1.79%   |
| UID191| 76  | HL                | DNMT3A     | p.V441fs      | 0.65%   |
|        |     |                   | ASXL1      | p.W796X       | 0.83%   |
| UID222| 66  | DLBCL             | STAG2      | p.E950X       | 0.42%   |
| UID247| 48  | MZL               | DNMT3A     | p.R771X       | 0.43%   |
| UID272| 64  | DLBCL             | DNMT3A     | p.Q415X       | 0.65%   |

AA: Amino acid, VAF: Variant allele frequency, DLBCL: Diffuse large B cell lymphoma, HL: Hodgkin's lymphoma, FL: Follicular lymphoma, MCL: Mantle cell lymphoma, MZL: Marginal zone lymphoma
Table S5. Clinical characteristics of an independent cohort of 74 patients with lymphoma treated with frontline chemotherapy protocol.

| N = 74 (IQR or %) |
|--------------------|
| **Median age, years** | 56 (44 - 64) |
| **Female** | 18 (27) |
| **Male** | 48 (73) |
| **Diagnosis** |
| **DLBCL** | 64 (86) |
| **High grade FL** | 7 (9) |
| **ALCL** | 2 (3) |
| **PBL** | 1 (2) |
| **Radiation** | 35 (47) |
| **Auto SCT** | 16 (22) |
| **Developed t-MNs** | 5 (7) |

DLBCL: Diffuse large B cell lymphoma, FL: Follicular lymphoma, ALCL: Anaplastic large cell lymphoma, PBL: Plasmablastic lymphoma, SCT: stem cell transplant. *All patients received CHOP (cylophosphamide, doxorubicin, vincristine, and prednisone) chemotherapy with or without melatonin as a frontline therapy.
Table S6. Clinical characteristics of 5 patients in an independent cohort who developed t-MNs after therapy.

| UID      | Age at primary cancer | Gender | Primary Cancer | Chemotherapy | Radiation therapy | Auto SCT | Latency to t-MN (years) | Age at t-MN | t-MN diagnosis | Cytogenetics in t-MN bone marrow | Antecedent clonal hematopoiesis | VAF |
|----------|-----------------------|--------|----------------|--------------|-------------------|----------|------------------------|-------------|----------------|----------------------------------|---------------------------------|-----|
| UID800699 59 | Male | DLBCL | CHOP | 50 Gy | Yes | 12.8 | 71 | t-MDS | 45 XY, -7 | ND | NA | |
| UID842047 50 | Female | High grade FL | CHOP, R-FND | 35 Gy | Yes | 7.0 | 57 | t-MDS | 46 XX, del20q | TET2 p. E478X | 0.38% | |
| UID915445 62 | Male | DLBCL | CHOP, RICE | NA | Yes | 3.0 | 65 | t-MDS | 46XY, del7q | DNMT3As p. S714C | 0.23% | |
| UID10010 59 | Female | DLBCL | CHOP, RICE | NA | Yes | 1.5 | 61 | t-MDS | 45XX, -7 | CBL p. C381R | 1.64% | |
| UID942741 56 | Male | DLBCL | CHOP, MINE, FND | NA | No | 3.0 | 59 | t-MDS | 42,XY;del(5)(q11.2;q33),del(14)(q24.3q32),del(12)(q13.13q14.13),del(19)(q13.2q13.3),-14,-16,-20 | ASXL1 p. R117fs | 0.50% | |

*DLBCL: Diffuse large B cell lymphoma, FL: Follicular lymphoma, CHOP: Cyclophosphamide, doxorubicin, vincristine, prednisone, RFND: Rituximab, fludarabine, mitoxantrone, dexamethasone, RICE: Rituximab, ifosfamide, carboplatin, etoposide, MINE: Mesna, ifosfamide, mitoxantrone, etoposide; SCT: stem cell transplant, t-MN: therapy-related myeloid neoplasm, ND: not detected, NA: not applicable
## Table S7. Mutations detected as clonal hematopoiesis in pre-treatment PB samples from 74 patients in an independent cohort.

| UID      | Age | Lymphoma diagnosis | Gene   | AA change       | VAF (%) | Developed t-MNs |
|----------|-----|--------------------|--------|-----------------|---------|-----------------|
| UID808859 | 83  | DLBCL              | TP53   | p.W91X          | 0.10%   |                 |
| UID812597 | 63  | DLBCL              | TET2   | p.H762fs        | 1.11%   |                 |
| UID827609 | 71  | DLBCL              | ASXL1  | p.Q760X         | 2.03%   |                 |
| UID827765 | 50  | FL                 | DNMT3A | c.1554+1G>A     | 3.26%   |                 |
| UID831007 | 76  | DLBCL              | DNMT3A | p.R736C         | 0.84%   |                 |
|           |     |                    |        | p.W330X         | 1.80%   |                 |
| UID834919 | 65  | DLBCL              | TP53   | p.Y163C         | 0.45%   |                 |
| UID835179 | 40  | DLBCL              | TET2   | p.Q897X         | 4.66%   |                 |
| UID836815 | 57  | DLBCL              | DNMT3A | p.W860R         | 1.90%   |                 |
| UID837041 | 55  | DLBCL              | DNMT3A | p.G746fs        | 0.35%   |                 |
| UID842047 | 50  | FL                 | TET2   | p.E478X         | 0.38%   | Yes             |
| UID853091 | 64  | DLBCL              | TP53   | p.S241fs        | 0.16%   |                 |
| UID915445 | 62  | DLBCL              | DNMT3A | p.S741C         | 0.23%   | Yes             |
| UID921701 | 30  | DLBCL              | ASXL1  | p.G804X         | 0.15%   |                 |
| UID942741 | 56  | DLBCL              | ASXL1  | p.R1171fs       | 0.50%   | Yes             |
| UID100100 | 59  | DLBCL              | CBL    | p.C381R         | 1.64%   | Yes             |
Table S8. Association between clonal hematopoiesis and auto-SCT. Correlation was tested by Fisher’s exact test.

|                        | No auto SCT | Auto SCT | P-value |
|------------------------|-------------|----------|---------|
| No clonal hematopoiesis| 48          | 11       | 0.29    |
| Clonal hematopoiesis   | 10          | 5        |         |
Supplemental References

1. Zhang J, Fujimoto J, Zhang J, Wedge DC, Song X, Zhang J, et al. Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. Science. 2014;346(6206):256-9.

2. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9.

3. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nature biotechnology. 2013;31(3):213-9.

4. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics. 2009;25(21):2865-71.

5. Papaemmanuil E, Gerstung M, Malcovati L, Tauro S, Gundem G, Van Loo P, et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood. 2013;122(22):3616-27; quiz 99.

6. Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman PV, Mar BG, et al. Age-related clonal hematopoiesis associated with adverse outcomes. The New England journal of medicine. 2014;371(26):2488-98.
| Patient ID | Gene    | AA Change             | VAF Change (%) | Depth (X) | Binomial Pvalue Primary Cancer |
|-----------|---------|-----------------------|----------------|-----------|------------------------------|
| UID12766  | W71     | p.S381X               | ND             | 2197      | NA                           |
| UID10164  | RUNX1   | p.N249fs              | 9.91           | 629       | 2.2e-16                      |
|           | RUNX1   | p.L189fs              | 5.40           | 1463      | 2.2e-16                      |
|           | TP53    | p.A271P               | 9.6            | 1654      | 2.2e-16                      |
|           | TP53    | p.M540V               | 5.06           | 2088      | 2.2e-16                      |
|           | IDH2    | p.R140Q               | 10.1           | 2502      | 2.2e-16                      |
|           | SRSF2   | p.R95delinRR          | 15.2e           | 872       | 2.2e-16                      |
|           | DNMTF3A | p.R880P               | 19.83          | 1612      | 2.2e-16                      |
|           | IDH2    | p.R172K               | ND             | 2300      | NA                           |
|           | ASXL1   | p.R209X               | 0.92           | 2459      | 2.2e-16                      |
|           | DNMTF3A | p.c.1475-2A+G         | 7.03           | 2622      | 2.2e-16                      |
|           | TP53    | p.C242Y               | 0.62           | 4206      | 2.2e-16                      |
|           | TP53    | p.I193R               | 23.11          | 2627      | 2.2e-16                      |
|           | DNMTF3A | p.I809_R89del        | 0.92           | 4143      | 7.83e-14                     |
|           | TET2    | p.L1212X              | 5.23           | 1286      | 2.2e-16                      |
|           | TET2    | p.S1444fs             | 9.40           | 2103      | 2.2e-16                      |
|           | TET2    | p.Y1265X              | 2.85           | 568       | 2.2e-16                      |
|           | TP53    | p.Y203C               | 0.67           | 2008      | 2.2e-16                      |
|           | U2AF1   | p.Q157P               | 4.92           | 2578      | 2.2e-16                      |
|           | DNMTF3A | p.R880C               | 19.83          | 1920      | 2.2e-16                      |
|           | KDM6A   | p.I1231fs             | 5.15           | 274       | 2.2e-16                      |
|           | KDM5A   | p.A378fs              | 7.25           | 1231      | 2.2e-16                      |
|           | NFRAS   | p.G131V               | 7.35           | 1424      | 2.2e-16                      |
|           | PPP1F11 | p.G500V               | 4.23           | 881       | 2.2e-16                      |
|           | TET2    | p.R581fs              | 0.92           | 949       | 2.2e-16                      |
|           | TP53    | p.R280G               | 0.48           | 1866      | 2.2e-16                      |
|           | KRAS    | p.G12A                | 2.45           | 389       | NA                           |
|           | NFRAS   | p.G131R               | 5.35           | 1715      | NA                           |
|           | TP53    | p.Y107X               | 0.32           | 3058      | 2.2e-16                      |
|           | ASXL1   | p.S1391fs             | 9.22           | 2194      | 2.2e-16                      |
|           | ASXL1   | p.Y991X               | 9.4            | 3531      | 2.2e-16                      |
|           | CUX1Y   | p.G816fs              | 0.16           | 4185      | 3.99e-11                     |
|           | CUX1Y   | p.E308fs              | 8.7            | 4971      | 2.2e-16                      |
|           | FL73    | D593delinsEAP007      | 0.68           | 1606      | 2.2e-16                      |
|           | GNB1    | p.K57E                | 23.63          | 1623      | 2.2e-16                      |
|           | KDM6A   | p.R650X               | 11.55          | 672       | 6.36e-12                     |
|           | MAD21   | p.E553X               | 4.43           | 191       | 2.2e-16                      |
|           | RUNX1   | p.G1665fs             | 9.36           | 939       | NA                           |
|           | RUNX1   | p.R264X               | 9.61           | 1629      | 2.2e-16                      |
|           | RUNX1   | p.c.909-1G+T          | 2.67           | 3697      | 2.2e-16                      |
|           | RUNX1   | p.O351A               | 0.93           | 2521      | 2.2e-16                      |
|           | SRSF2   | p.P95H                | 30.84          | 1154      | 2.2e-16                      |
|           | TP53    | p.L194H               | 0.41           | 1723      | NA                           |
|           | QATA2   | p.Y322_M0325delinsW   | 0.33           | 3172      | NA                           |
|           | TP53    | p.G242Y               | 0.63           | 3156      | 2.2e-16                      |
|           | TP53    | p.H193R               | 3.2            | 1909      | 2.2e-16                      |
|           | TET2    | p.H1360Y              | 9.74           | 2093      | 2.2e-16                      |
|           | TP53    | p.R156H               | 3.43           | 1620      | 2.2e-16                      |
|           | TP53    | p.R264Q               | 42.01          | 1859      | 2.2e-16                      |
Figure S2. VAF difference of mutations in prior PB samples that became drivers and did not become drivers.

P = 0.001
Figure S3A. Cumulative incidence of t-MNs between patients who underwent autologous SCT and who did not.
Figure S3B. Cumulative incidence of t-MNs between patients who were age 60 years or older and younger.

Number at risk

| Age          | 0 | 5 | 10 | 15 |
|--------------|---|---|----|----|
| Age < 60     | 48| 41| 31 | 12 |
| Age ≥ 60     | 26| 20| 16 | 8  |

P = 0.46
Figure S3C. Cumulative incidence of t-MNs between patients who received radiation therapy and who did not.
Figure S4. Patients with clonal hematopoiesis in the entire cohort of 142 patients. Primary axis shows actual number of cases with clonal hematopoiesis. Secondary axis shows proportion of patients with clonal hematopoiesis within the age group defined in x axis.
Figure S5. Comparison of the frequency of the mutated genes as clonal hematopoiesis between patients who developed t-MNs (N = 19) and patients who did not develop t-MNs (N = 123). One asterisk indicates $P < 0.05$ and two asterisks indicate $P < 0.01$. 