Retrotransposable Element Derepression Distinguishes *DNMT3A*- from *TET2*-Mutant Clonal Hematopoiesis

Maroof Hasan*[1], Sylvia Durandeau[1], I Richard Thompson[2], Hanae Roussotte[2], Yi-Ting Tsai[2], Nogayhan Seymen[2], Sila Gerlevik[2], Nicholas Bianchini[2], Khadijeh Alishah[2], Merilyn M Albuquerque[2], Jen Lewis[3], Laarni Bogaanay[3], Niels Asger Jakobsen[10,11], Bernd Zeisig[2], Sheeba Irshad[2,4,5], Paresh Vyas[10,11,12], Eric CW So[2], Alfredo Iacoangeli[6,7,8,9], Xuesen Zheng[2,3], Sarah Mackie[2,3], Kai Yi Mok[2,3], Viktoria Uksaite[2,3], Louis-Francois Handfield[2,3], Giorgio Napolitani[2], James Carmichael[1], Ghulam J Mufti[2,3], Anita K Gandhi[1], Rajasekhar NVS Suragani[1], Lynn Quek[2,3], Mohammad M Karimi*[2]

[1]Bristol Myers Squibb, Hematology Translational Medicine, New Jersey, USA

[2]Comprehensive Cancer Centre, School of Cancer & Pharmaceutical Sciences, Faculty of Life Sciences & Medicine, King's College London, London, UK.

[3]Haematology Department, King's College Hospital NHS Foundation Trust, London, UK.

[4]Breast Cancer Now Research Unit, King's College London, London, UK

[5]Guy's and St Thomas' NHS Foundation Trust, London, UK.

[6]Department of Basic and Clinical Neuroscience, King’s College London, London, UK.

[7]Department of Biostatistics and Health Informatics, King’s College London, London, UK

[8]NIHR BRC SLAM NHS Foundation Trust, London, UK

[9]Perron Institute for Neurological and Translational Science, University of Western Australia Medical School, Perth, Australia

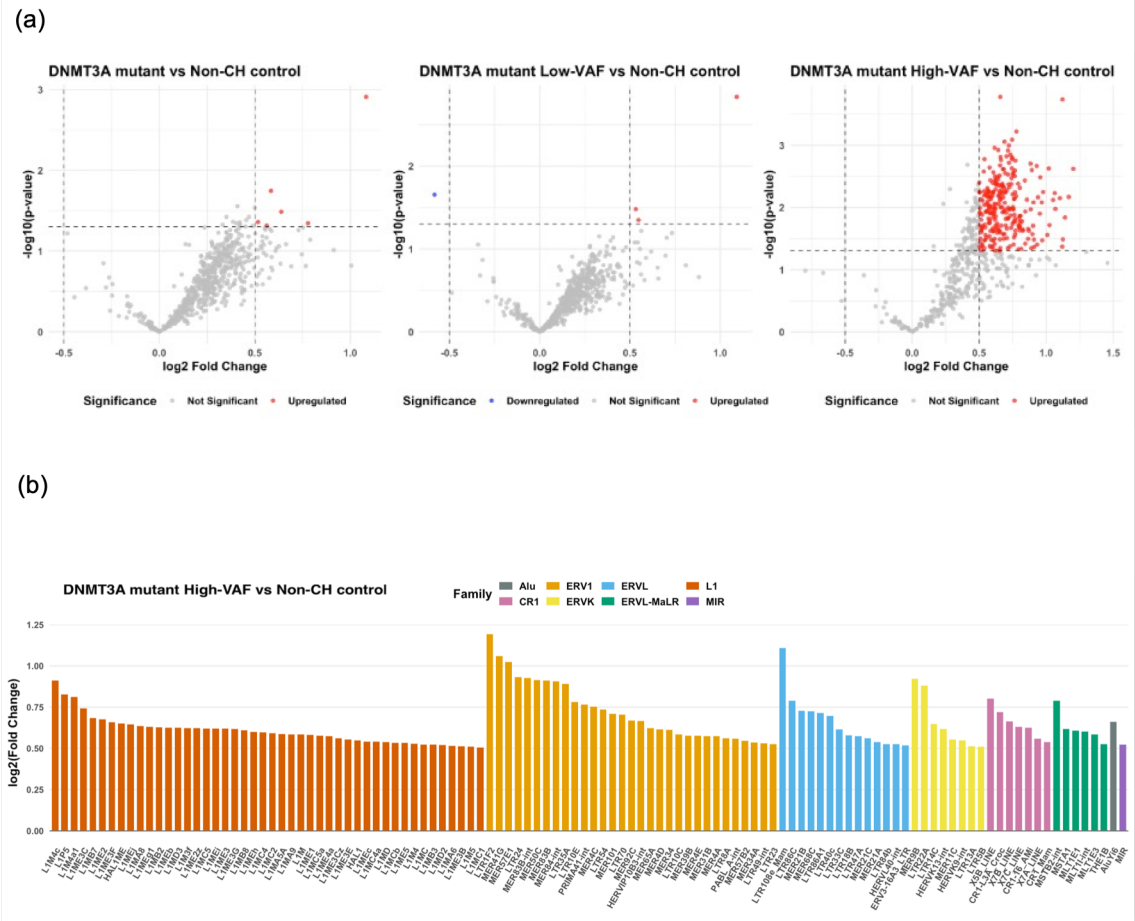
[10]MRC Molecular Haematology Unit, MRC Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK

[11]Oxford Centre for Haematology, NIHR Oxford Biomedical Research Centre, Oxford, UK

[12]Department of Haematology, Oxford University Hospitals NHS Foundation Trust, Oxford, UK.

*Corresponding author:

Dr Mohammad Karimi, Comprehensive Cancer Centre, School of Cancer & Pharmaceutical Sciences, Faculty of Life Sciences & Medicine, King's College London, London, UK. (Mohammad.Karimi@kcl.ac.uk)

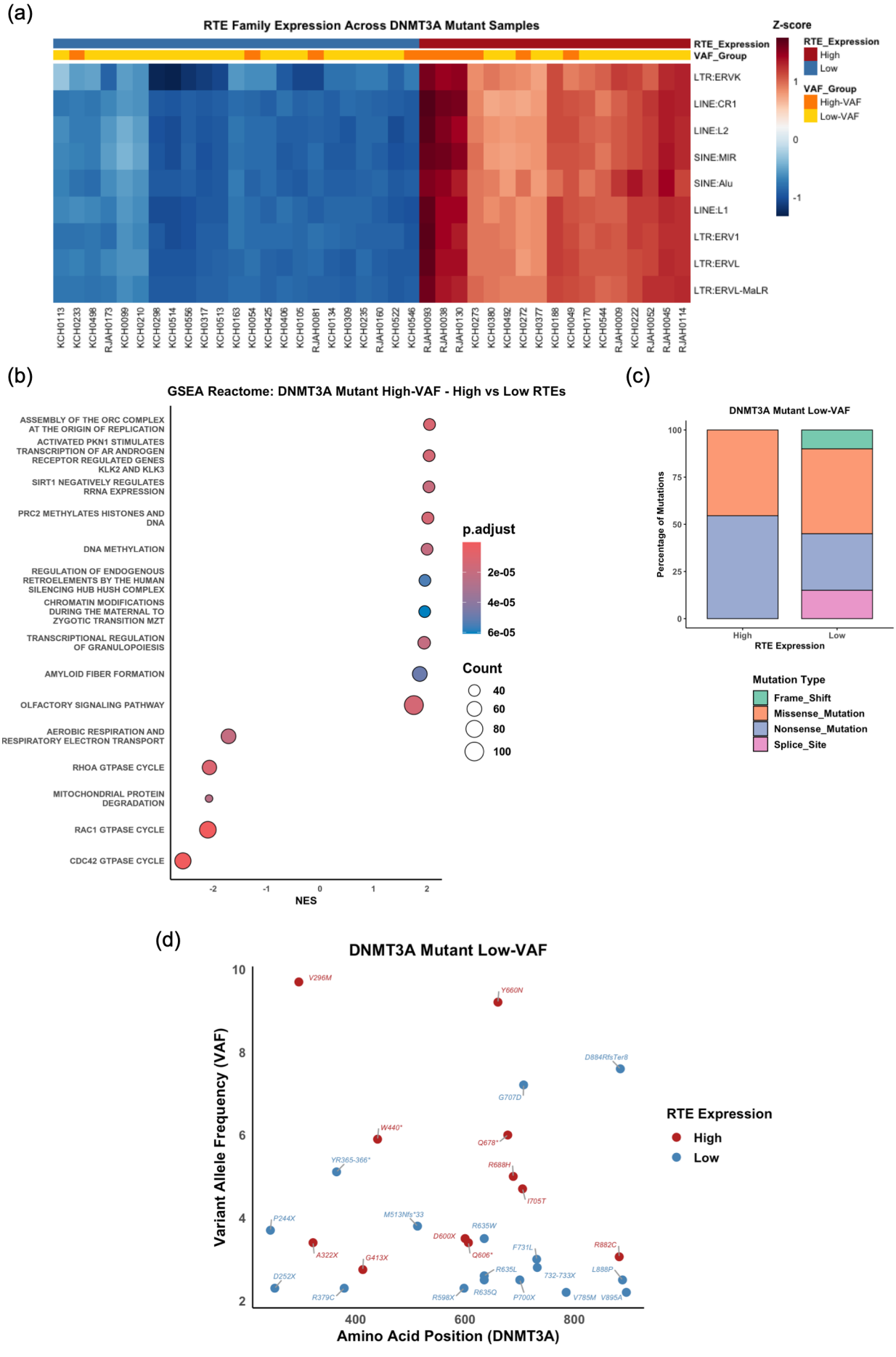


Supplementary Figure 1. Upregulated RTE subfamilies in DNMT3A-mutant clonal haematopoiesis.

(a) Sensitivity analysis using genomic DNA-derived VAF to classify DNMT3A-mutant cases into high- and low-VAF groups. The overall pattern of RTE activation remained consistent with the primary analysis and identified an even greater number of

significantly upregulated RTE subfamilies in the high-VAF group, demonstrating that the findings are robust to the method used for estimating clonal burden.

(b) Lists of RTE subfamilies upregulated in DNMT3A-mutant high-VAF vs. non-CH controls. The DNMT3A-mutant high-VAF group showed the largest number of upregulated subfamilies, particularly within the LINE:L1, LTR:ERV1, and LTR:ERVL classes ($\log_2$ fold change > 0.5, $p < 0.05$).


(a)
RTE Family Expression Across DNMT3A Mutant Samples
Z-score
RTE_Expression
VAF_Group
High
Low
High-VAF
Low-VAF
LTR:ERVK
LINE:CR1
LINE:L2
SINE:MIR
SINE:Alu
LINE:L1
LTR:ERV1
LTR:ERVL
LTR:ERVL-MaLR
(b)
GSEA Reactome: DNMT3A Mutant High-VAF - High vs Low RTEs
ASSEMBLY OF THE ORC COMPLEX AT THE ORIGIN OF REPLICATION
ACTIVATED PKN1 STIMULATES TRANSCRIPTION OF AR ANDROGEN RECEPTOR REGULATED GENES KLK2 AND KLK3
SIRT1 NEGATIVELY REGULATES RRNA EXPRESSION
PRC2 METHYLATES HISTONES AND DNA
DNA METHYLATION
REGULATION OF ENDOGENOUS RETROELEMENTS BY THE HUMAN SILENCING HUB HUSH COMPLEX
CHROMATIN MODIFICATIONS DURING THE MATERNAL TO ZYGOTIC TRANSITION MZT
TRANSCRIPTIONAL REGULATION OF GRANULOPOIESIS
AMYLOID FIBER FORMATION
OLFACTORY SIGNALING PATHWAY
AEROBIC RESPIRATION AND RESPIRATORY ELECTRON TRANSPORT
RHOA GTPASE CYCLE
MITOCHONDRIAL PROTEIN DEGRADATION
RAC1 GTPASE CYCLE
CDC42 GTPASE CYCLE
NES
p.adjust
Count
(c)
DNMT3A Mutant Low-VAF
Percentage of Mutations
RTE Expression
High
Low
Mutation Type
Frame_Shift
Missense_Mutation
Nonsense_Mutation
Splice_Site
(d)
DNMT3A Mutant Low-VAF
Variant Allele Frequency (VAF)
Amino Acid Position (DNMT3A)
RTE Expression
High
Low
V296M
Y660N
D884RfsTer8
G707D
W440*
Q678*
YR365-366*
R688H
I705T
P244X
M513Nfs*33
R635W
D600X
Q606*
A322X
G413X
F731L
R882C
D252X
R379C
R635L
732-733X
L888P
R598X
R635Q
P700X
V785M
V895A

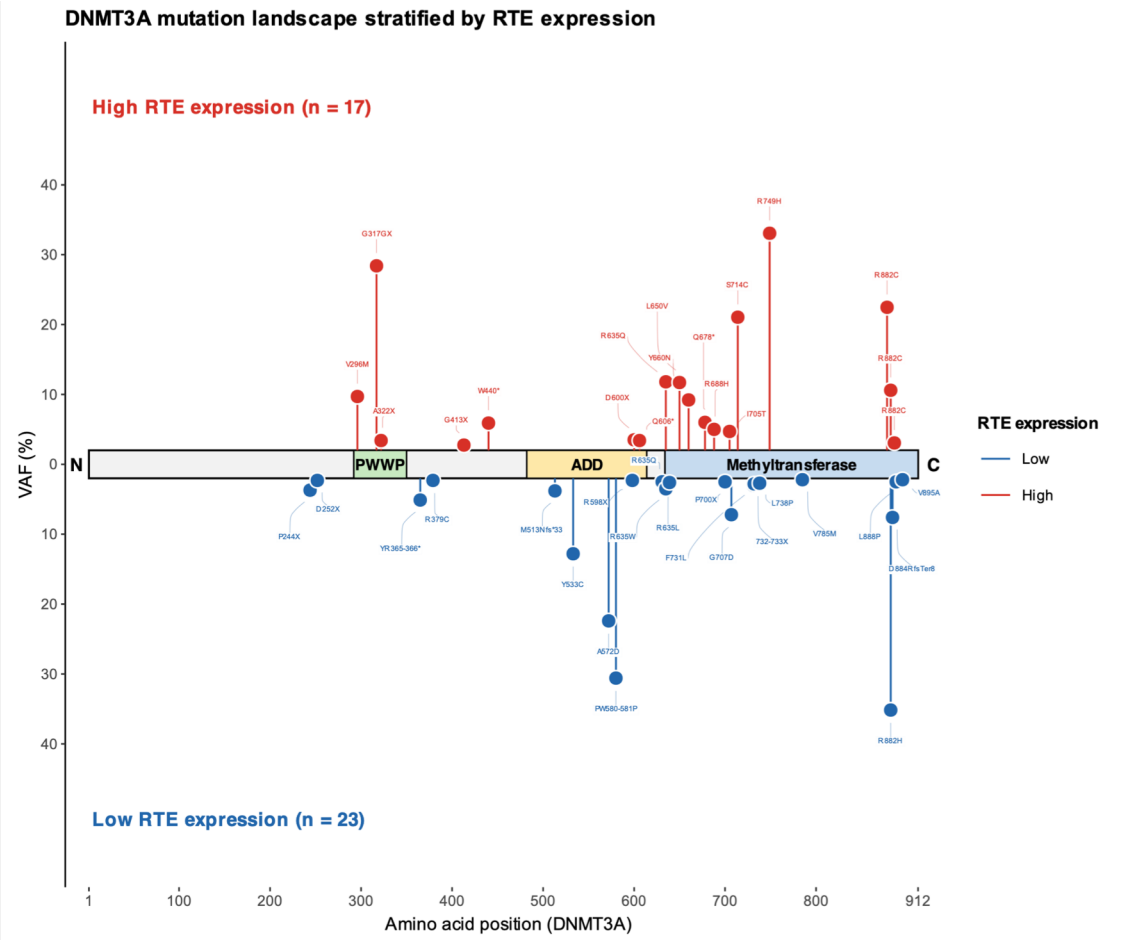
Supplementary Figure 2. Stratification of *DNMT3A*-mutant cases and additional analyses of low-VAF and high-VAF subgroups.

(a) *DNMT3A*-mutant cases were stratified into high vs. low RTE expression groups based on RTE family expression.

(b) Gene set enrichment analysis (GSEA) of high VAF *DNMT3A* clones comparing high vs. low RTE expression showing enrichment of pathways involved in HUSH & PRC2-mediated silencing and SIRT1-mediated rRNA regulation.

(c) Enrichment analysis of mutation types in *DNMT3A*-mutant low-VAF samples comparing high vs. low RTE expression, showing an overrepresentation of nonsense mutations.

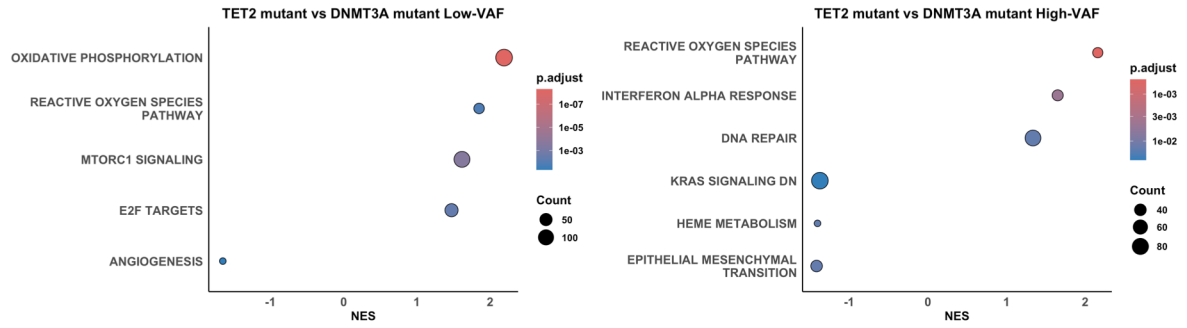
(d) Scatterplot of amino acid position vs. variant allele frequency (VAF) in *DNMT3A*-mutant low-VAF samples, colour-coded by RTE expression group.



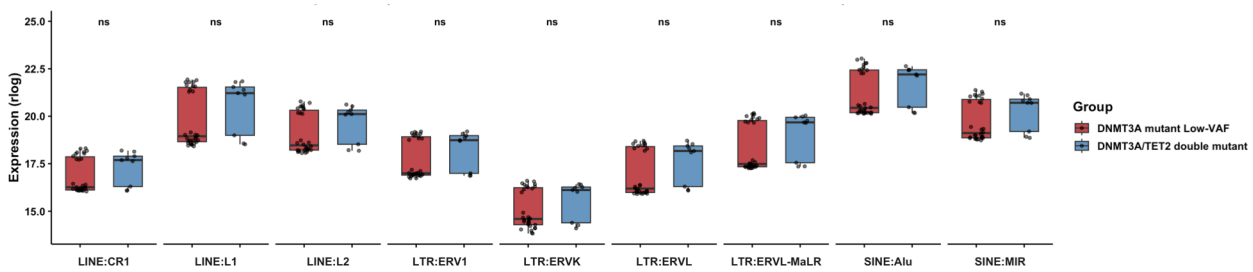
Supplementary Figure 3. Distribution of DNMT3A mutations according to RTE expression status.

Lollipop plot showing the distribution of DNMT3A mutations across the functional domains of the DNMT3A protein. Mutations are stratified according to high (red, above the protein) and low (blue, below the protein) RTE expression groups. Each lollipop represents an individual patient, with the position indicating the amino acid affected by the mutation. The vertical position of each lollipop corresponds to the DNMT3A VAF, serving as a measure of clonal burden. Functional protein domains (PWWP, ADD and DNA methyltransferase [MTase]) are shown along the protein schematic.

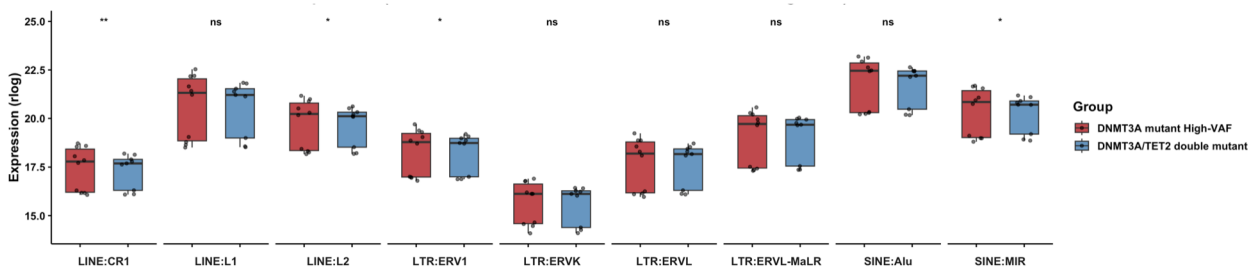
(a)


TET2 mutant vs DNMT3A mutant Low-VAF
OXIDATIVE PHOSPHORYLATION
REACTIVE OXYGEN SPECIES PATHWAY
MTORC1 SIGNALING
E2F TARGETS
ANGIOGENESIS
-1
0
1
2
NES
p.adjust
1e-07
1e-05
1e-03
Count
50
100
TET2 mutant vs DNMT3A mutant High-VAF
REACTIVE OXYGEN SPECIES PATHWAY
INTERFERON ALPHA RESPONSE
DNA REPAIR
KRAS SIGNALING DN
HEME METABOLISM
EPITHELIAL MESENCHYMAL TRANSITION
-1
0
1
2
NES
p.adjust
1e-03
3e-03
1e-02
Count
40
60
80

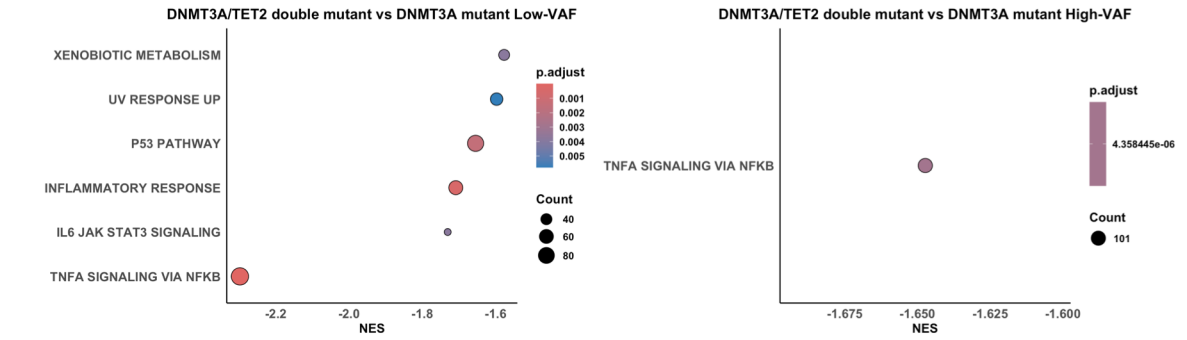
(b)


Expression (rlog)
25.0
22.5
20.0
17.5
15.0
ns
LINE:CR1
LINE:L1
LINE:L2
LTR:ERV1
LTR:ERVK
LTR:ERVL
LTR:ERVL-MaLR
SINE:Alu
SINE:MIR
Group
DNMT3A mutant Low-VAF
DNMT3A/TET2 double mutant

(c)


Expression (rlog)
25.0
22.5
20.0
17.5
15.0
**
ns
*
LINE:CR1
LINE:L1
LINE:L2
LTR:ERV1
LTR:ERVK
LTR:ERVL
LTR:ERVL-MaLR
SINE:Alu
SINE:MIR
Group
DNMT3A mutant High-VAF
DNMT3A/TET2 double mutant

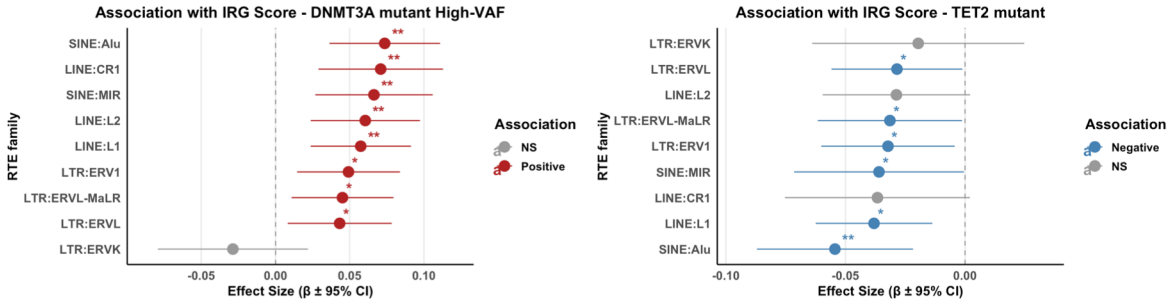
(d)


DNMT3A/TET2 double mutant vs DNMT3A mutant Low-VAF
XENOBIOTIC METABOLISM
UV RESPONSE UP
P53 PATHWAY
INFLAMMATORY RESPONSE
IL6 JAK STAT3 SIGNALING
TNFA SIGNALING VIA NFKB
-2.2
-2.0
-1.8
-1.6
NES
p.adjust
0.001
0.002
0.003
0.004
0.005
Count
40
60
80
DNMT3A/TET2 double mutant vs DNMT3A mutant High-VAF
TNFA SIGNALING VIA NFKB
-1.675
-1.650
-1.625
-1.600
NES
p.adjust
4.358445e-06
Count
101

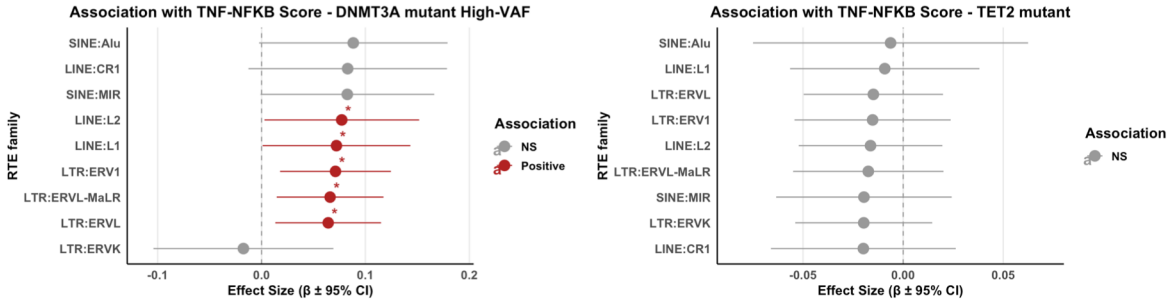
Supplementary Figure 4. Pathway and RTE expression analyses in *TET2* and *DNMT3A/TET2* double mutants.

(a) Gene set enrichment analysis (GSEA) of Hallmark gene sets comparing (1) *TET2-mutant*s vs. non-CH controls, (2) *TET2-mutant*s vs. *DNMT3A*-mutant low-VAF, and (3) *TET2-mutant*s vs. *DNMT3A*-mutant high-VAF. Oxidative phosphorylation was enriched in comparisons 1 and 2, whereas reactive oxygen species (ROS) and interferon-α response pathways were enriched in comparison 3.
(b–c) Boxplots of RTE family expression comparing (a) *DNMT3A*-mutant low-VAF vs. *DNMT3A/TET2* double mutants, and (b) *DNMT3A*-mutant high-VAF vs. *DNMT3A/TET2* double mutants. CR1, L2, ERV1, and MIR families were significantly downregulated in panel b.
(d) GSEA of Hallmark gene sets comparing (1) *DNMT3A/TET2* double mutants vs. *DNMT3A*-mutant low-VAF, and (2) *DNMT3A/TET2* double mutants vs. *DNMT3A*-mutant high-VAF. Pathway analysis revealed attenuation of TNF–NFκB and interferon programs compared with *DNMT3A-mutant* cases.

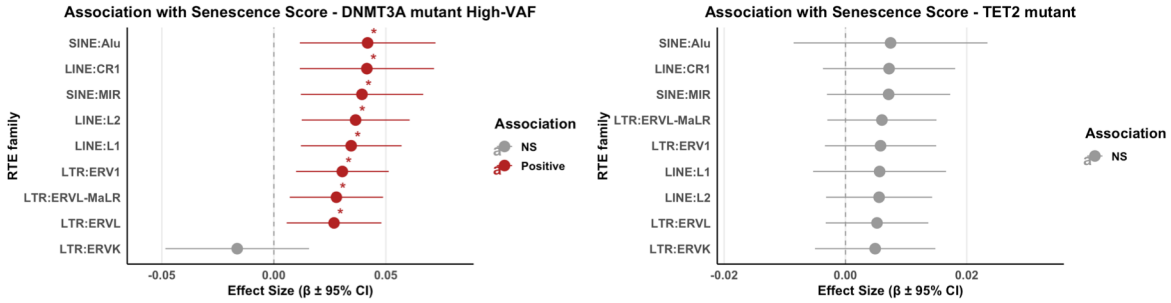
(a)


Association with IRG Score - DNMT3A mutant High-VAF
RTE family
SINE:Alu
LINE:CR1
SINE:MIR
LINE:L2
LINE:L1
LTR:ERV1
LTR:ERVL-MaLR
LTR:ERVL
LTR:ERVK
-0.05
0.00
0.05
0.10
Effect Size (β ± 95% CI)
Association
NS
Positive
Association with IRG Score - TET2 mutant
LTR:ERVK
LTR:ERVL
LINE:L2
LTR:ERVL-MaLR
LTR:ERV1
SINE:MIR
LINE:CR1
LINE:L1
SINE:Alu
-0.10
-0.05
0.00
Effect Size (β ± 95% CI)
Association
Negative
NS

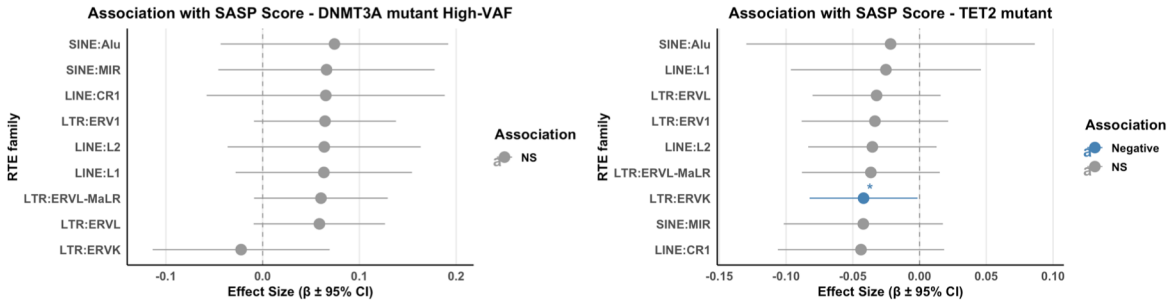
(b)


Association with TNF-NFKB Score - DNMT3A mutant High-VAF
RTE family
SINE:Alu
LINE:CR1
SINE:MIR
LINE:L2
LINE:L1
LTR:ERV1
LTR:ERVL-MaLR
LTR:ERVL
LTR:ERVK
-0.1
0.0
0.1
0.2
Effect Size (β ± 95% CI)
Association
NS
Positive
Association with TNF-NFKB Score - TET2 mutant
SINE:Alu
LINE:L1
LTR:ERVL
LTR:ERV1
LINE:L2
LTR:ERVL-MaLR
SINE:MIR
LTR:ERVK
LINE:CR1
-0.05
0.00
0.05
Effect Size (β ± 95% CI)
Association
NS

(c)


Association with Senescence Score - DNMT3A mutant High-VAF
RTE family
SINE:Alu
LINE:CR1
SINE:MIR
LINE:L2
LINE:L1
LTR:ERV1
LTR:ERVL-MaLR
LTR:ERVL
LTR:ERVK
-0.05
0.00
0.05
Effect Size (β ± 95% CI)
Association
NS
Positive
Association with Senescence Score - TET2 mutant
SINE:Alu
LINE:CR1
SINE:MIR
LTR:ERVL-MaLR
LTR:ERV1
LINE:L1
LINE:L2
LTR:ERVL
LTR:ERVK
-0.02
0.00
0.02
Effect Size (β ± 95% CI)
Association
NS

(d)


Association with SASP Score - DNMT3A mutant High-VAF
RTE family
SINE:Alu
SINE:MIR
LINE:CR1
LTR:ERV1
LINE:L2
LINE:L1
LTR:ERVL-MaLR
LTR:ERVL
LTR:ERVK
-0.1
0.0
0.1
0.2
Effect Size (β ± 95% CI)
Association
NS
Association with SASP Score - TET2 mutant
SINE:Alu
LINE:L1
LTR:ERVL
LTR:ERV1
LINE:L2
LTR:ERVL-MaLR
LTR:ERVK
SINE:MIR
LINE:CR1
-0.15
-0.10
-0.05
0.00
0.05
0.10
Effect Size (β ± 95% CI)
Association
Negative
NS

Supplementary Figure 5. Validation of associations between RTE family expression and gene signatures in an independent cohort.

Validation cohort (*n* = 13) from the study by Smith *et al.*, including 7 *DNMT3A*-mutant high-VAF and 6 *TET2*-mutant cases, supporting the findings shown in Figure 4. Forest plots show associations between RTE families and gene signature scores for *DNMT3A*-mutant high-VAF (left) and *TET2*-mutant (right) cases across four panels: (a) interferon regulatory genes (IRG), (b) TNF–NFκB, (c) senescence, and (d) senescence-associated secretory phenotype (SASP). Positive associations were observed for IRG, TNF–NFκB, and senescence signatures in *DNMT3A*-mutant high-VAF cases, while SASP showed only a trend.