Mechanism-guided quantification of LINE-1 reveals p53 regulation of both retrotransposition and transcription

ROME THERAPEUTICS

Visit our website to view this poster and our preprint

Alexander Solovyov[1*], Julie M. Behr[2*], David Hoyos[1], Eric Banks[2], Alexander W. Drong[2], Jimmy Z. Zhong[2], Enrique Garcia-Rivera[2], Wilson McKerrow[2], Chong Chu[2], Dennis Zaller[2], Menachem Fromer[2**], Benjamin Greenbaum[1,3**]

Affiliations: [1]Computational Oncology, Memorial Sloan Kettering Cancer Center, New York, NY, USA [2]ROME Therapeutics, Inc., Boston, MA, USA [3]Physiology, Biophysics & Systems Biology, Weill Cornell Medical College, New York, NY, USA

*These authors contributed equally to this work **These authors contributed equally to this work

Background

Somatic activity of LINE-1 (L1) mobile elements has been implicated in cancer etiology. Yet, detecting L1 genomic reintegration remains challenging. We developed Total ReCall to detect L1 insertions from short-read whole-genome sequencing and applied the algorithm to high-quality data from >750 paired tumor and normal samples from The Cancer Genome Atlas (TCGA). By integrating L1 insertional and transcriptional data with TP53 mutational status, we show that TP53 mutations facilitate retrotransposition (RT) both by disinhibiting L1 expression and deregulating its reintegration.

Schematic of insertion of a non-LTR retrotransposon

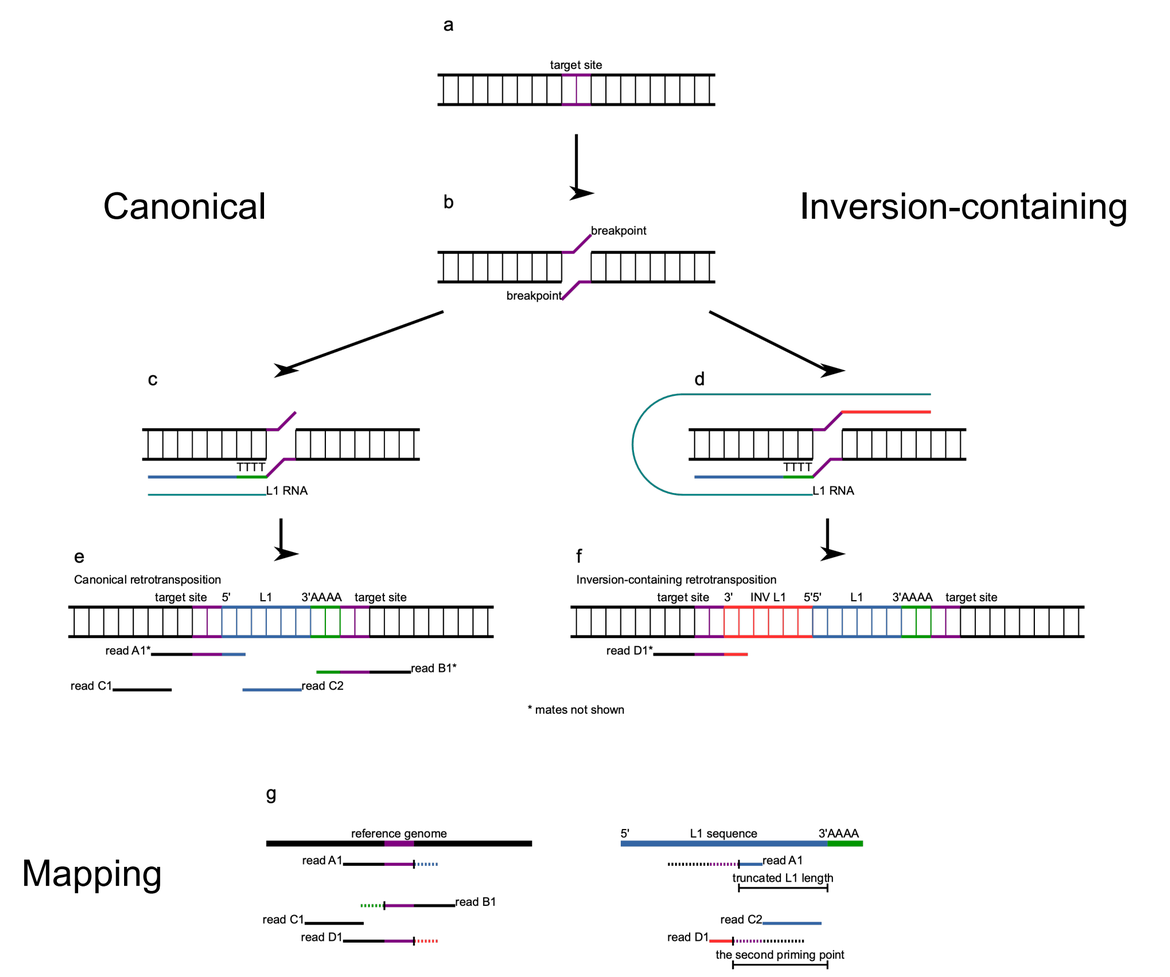
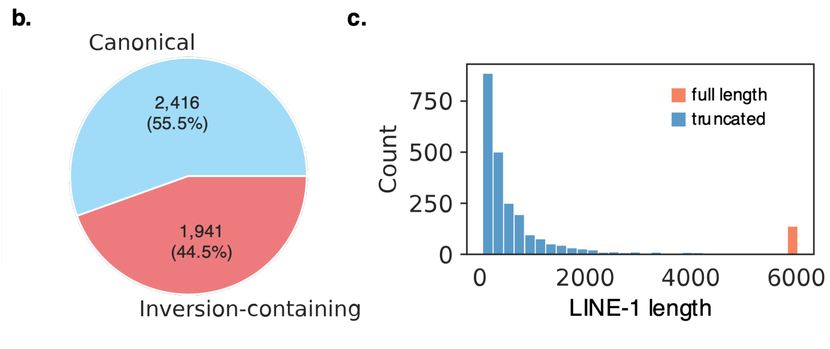


Figure 1: a) Genome before RT. b) Endonuclease breaks each strand of DNA. c) L1 RNA is reverse transcribed directly into the genome d) In some cases, double priming occurs. e-f) Genome after synthesis of the second strand of DNA and repair. g) Mapping of reads A-D to the reference genome (left) and the transposon sequence (right).

Pan-cancer survey reveals highest L1 RT burden in lung squamous cell carcinoma

a.

	Tumor specific	Normal specific
Total (N=765)	4,357	100
Mean per sample	5.7	0.13



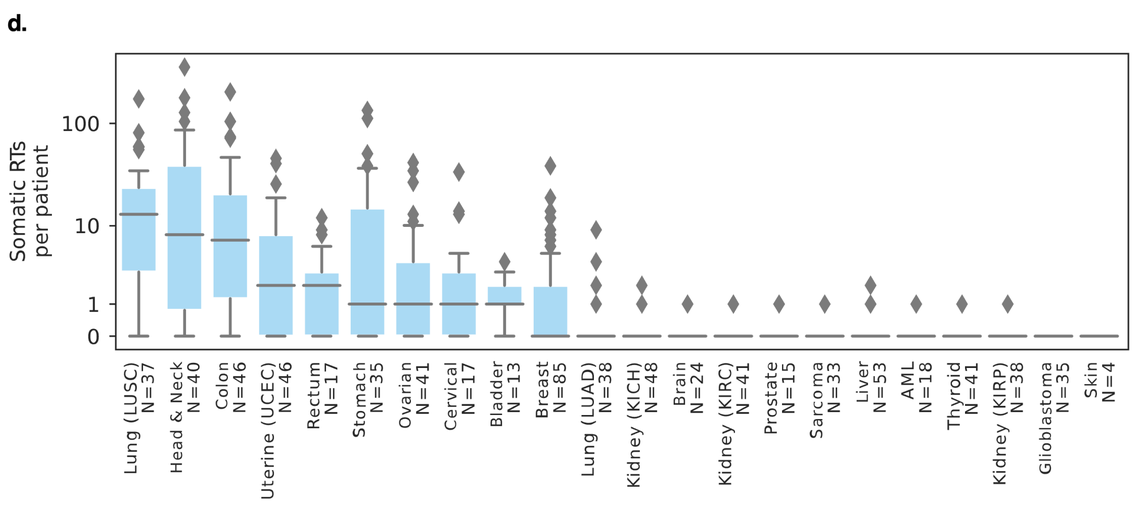


Figure 2: a) Total number of RTs identified across 765 tumor and paired normal samples. b) Breakdown of total canonical or inversion-containing insertions tumor-specific somatic RTs. c) Estimated length of inserted L1 within the canonical, tumor-specific RTs. d) Somatic tumor-specific L1 RTs in each sample ("RT burden") grouped by tumor type.

L1 RNA levels differ by tumor types but are related to RT burden

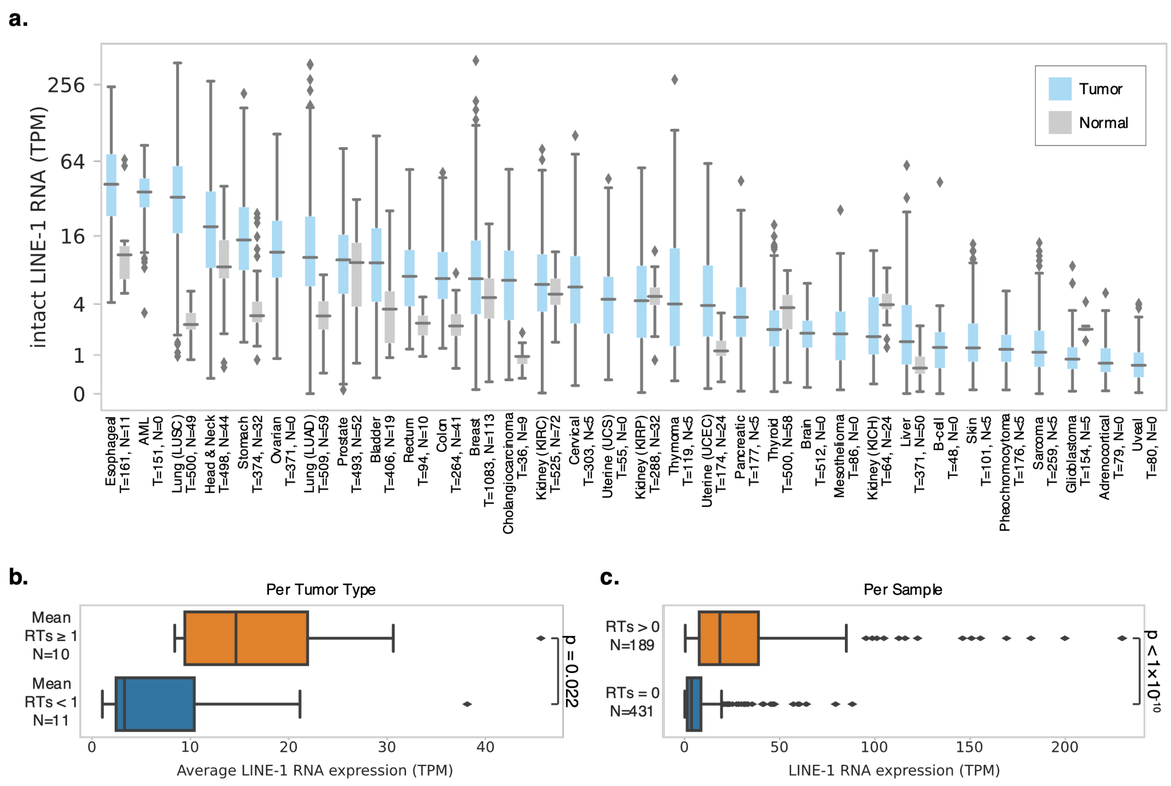
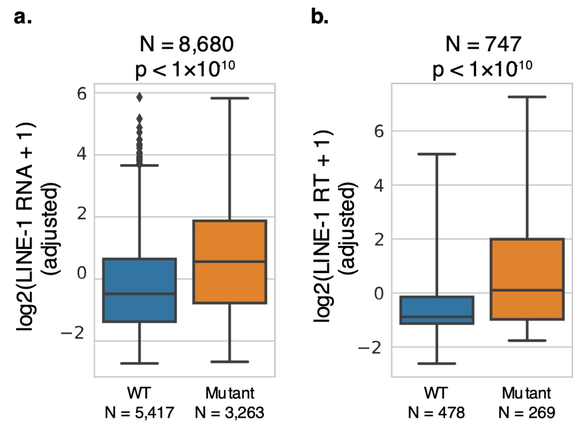


Figure 3: a) Estimated expression of intact L1 RNA in each sample by tumor type. b) Average expression of intact L1 RNA per tumor type grouped by average RT count per tumor type. c) Expression of intact L1 RNA grouped by RT count per sample.

p53 mutant groups exhibited significantly higher L1 RNA and L1 RT burden than WT

Figure 4. a) Expression of intact L1 RNA in samples with mutant or wild-type p53. b) Count of somatic L1 RTs in samples with mutant or wild-type p53.



Pan-cancer statistical model finds that p53 limits RT by repression of LINE- 1 transcription and regulation of integration

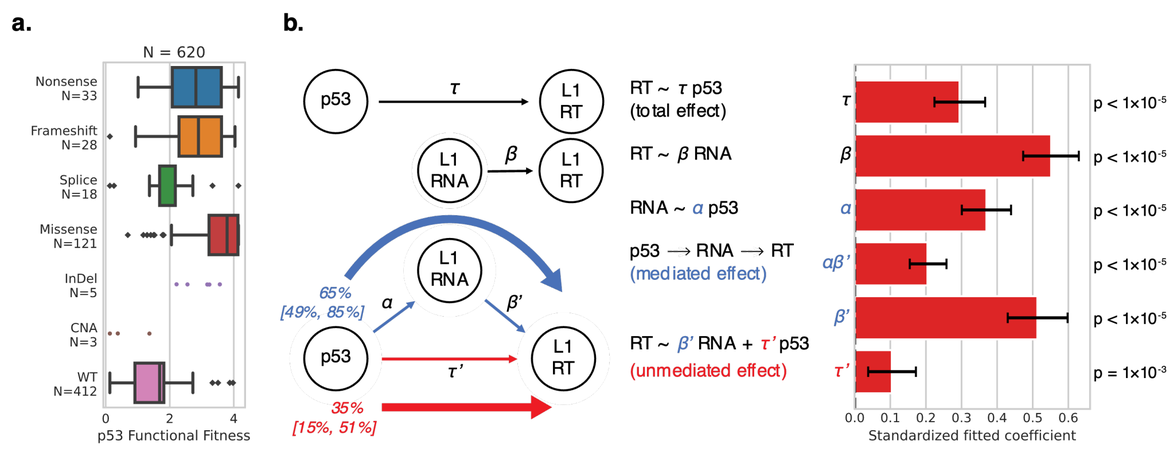


Figure 5: a) p53 functional fitness score per sample by mutation category. b) Mediation model taking p53 functional fitness as the independent variable, adjusted log2 of intact L1 RNA expression as the mediating variable, and adjusted log2 of somatic L1 RT count as the dependent variable. Left, schematic showing the linear regressions performed, as well as the resulting estimated weights for the mediated and unmediated pathways. Right, standardized fitted values for each coefficient within the mediation model and corresponding likelihood.

Conclusions

- Applying Total ReCall to high-quality data for >750 paired tumor and normal samples from The Cancer Genome Atlas (TCGA) shows high heterogeneity among tumor types, with increased RT burden in lung squamous cell carcinoma, head and neck, and colon cancers.
- We assessed active RNA expression of intact L1 in >9,000 TCGA tumor samples, revealing a clear correlation between L1 expression and RT.
- Leveraging these measurements, we show that TP53 mutations facilitate RT both by disinhibiting L1 expression and deregulating its reintegration.