Corresponding Author: Matthew Simon

Date: Dec 12, 2017

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work we publish. This form is published with all life science papers and is intended to promote consistency and transparency in reporting. All life sciences submissions use this form; while some list items might not apply to an individual manuscript, all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

   Describe how sample size was determined.

   All targeted sequencing samples were prepared in biological duplicates with the exception of 10uM and 40uM K562 targeted samples (Supplementary Fig. S6e). All TimeLapse-seq samples for both MEF and K562 cells and TT-TimeLapse-seq samples were performed as biological duplicates as is noted in the online methods. No sample size calculation was performed.

2. Data exclusions

   Describe any data exclusions.

   Cases where data were excluded are described in the text and methods. In the targeted sequencing section, data points were excluded if they did not meet a depth threshold of 3000 reads. The criteria for exclusion of reads and mutations in genome wide TimeLapse-seq data are described in the text and methods. Reads were filtered for unique sequences before alignment to either the mouse GRCm38 or human GRCh38 genome. Non-unique aligned reads were removed. Insertion-containing reads were not considered in mutational analyses. Sites of mutations were only considered if their base quality was 45 or above. Reads were removed if they contained greater than five T-to-C mutations and these mutations did not account for at least one third of the observed mutations (NMtag). Mutation data points were also removed if their base position was identified as a SNP. When estimating fraction new, transcripts were only included if at least two samples had more than 100 counts.

3. Replication

   Describe whether the experimental findings were reliably reproduced.

   The efficiency of oxidative-nucleophilic-aromatic-substitution was found to be reproducible in multiple screens assayed by our restriction endonuclease assay. All TimeLapse-seq and TT-TimeLapse-seq samples were performed as biological duplicates and the correlation for each duplicate analysis is presented.

4. Randomization

   Describe how samples/organisms/participants were allocated into experimental groups.

   Randomization was not relevant to our study.

5. Blinding

   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

   Blinding was not relevant to our study. After upstream treatment of cells and isolated RNA, all samples were handled and analyzed with the same protocols and pipelines.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or the Methods section if additional space is needed).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒	A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly.
☐	☒	A statement indicating how many times each experiment was replicated
☐	☒	The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒	A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒	The test results (e.g. *p* values) given as exact values whenever possible and with confidence intervals noted
☐	☒	A summary of the descriptive statistics, including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒	Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	All parameters used to analyze data are described in methods section. Custom scripts implementing these calculations are available upon request as is noted in the revised online methods.

For all studies, we encourage code deposition in a community repository (e.g. GitHub). Authors must make computer code available to editors and reviewers upon request. The *Nature Methods* guidance for providing algorithms and software for publication may be useful for any submission.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	No unique materials were used.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	No antibodies were used.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	MEF cell line was described in Yildirim et al. 2012, Nat. Struct. Mol. Biol. K562 cell line was a generous gift of the Slavoff lab, Yale University Department of Chemistry, New Haven, CT, 06511, USA.
b. Describe the method of cell line authentication used.	No further authentication beyond what is described in Yildirim et al. 2012, Nat. Struct. Mol. Biol. was performed.
c. Report whether the cell lines were tested for mycoplasma contamination.	Cell lines were not tested for mycoplasma contamination.
d. If any of the cell lines used in the paper are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	No commonly misidentified cell lines were used.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	No animals were used.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	Study did not involve human research participants.