**Materials Design Analysis Reporting (MDAR)**

**Checklist for Authors**

The [MDAR framework](https://osf.io/xfpn4/) establishes a minimum set of requirements in transparent reporting mainly applicable to studies in the life sciences.

*eLife* asks authors to **provide detailed information within their article** to facilitate the interpretation and replication of their work. Authors can also upload supporting materials to comply with relevant reporting guidelines for health-related research (see [EQUATOR Network](http://www.equator-network.org/%20)), life science research (see the [BioSharing Information Resource](http://biosharing.org/)), or animal research (see the [ARRIVE Guidelines](http://www.plosbiology.org/article/info:doi/10.1371/journal.pbio.1000412) and the [STRANGE Framework](https://doi.org/10.1038/d41586-020-01751-5); for details, see *eLife*’s [Journal Policies](https://reviewer.elifesciences.org/author-guide/journal-policies)). Where applicable, authors should refer to any relevant reporting standards materials in this form.

For all that apply, please note **where in the article** the information is provided. Please note that we also collect information about data availability and ethics in the submission form.

**Materials:**

|  |  |  |
| --- | --- | --- |
| **Newly created materials** | **Indicate where provided: section/figure legend** | **N/A** |
| The manuscript includes a dedicated "materials availability statement" providing transparent disclosure about availability of newly created materials including details on how materials can be accessed and describing any restrictions on access. | All cell lines containing genetic modifications of ASAR6-141 are freely available upon request (thayerm@ohsu.edu). |  |
|  |  |  |
| **Antibodies** | **Indicate where provided: section/figure legend** | **N/A** |
| For commercial reagents, provide supplier name, catalogue number and [RRID](https://scicrunch.org/resources), if available. | Table S4 |  |
|  |  |  |
| **DNA and RNA sequences** | **Indicate where provided: section/figure legend** | **N/A** |
| Short novel DNA or RNA including primers, probes: Sequences should be included or deposited in a public repository. | shRNA identifiers are in Table S4.  PCR primer and sgRNA sequences are in Table S3. |  |
|  |  |  |
| **Cell materials** | **Indicate where provided: section/figure legend** | **N/A** |
| Cell lines: Provide species information, strain. Provide accession number in repository OR supplier name, catalog number, clone number, OR RRID. | K562: human; ATCC, CCL-243.  HTD114: human; J.A. Tischfield.  C2C12: mouse C3H; ATCC, CRL-1772. |  |
| Primary cultures: Provide species, strain, sex of origin, genetic modification status. |  | N/A |
|  |  |  |
| **Experimental animals** | **Indicate where provided: section/figure legend** | **N/A** |
| Laboratory animals or Model organisms: Provide species, strain, sex, age, genetic modification status. Provide accession number in repository OR supplier name, catalog number, clone number, OR RRID. |  | N/A |
| Animal observed in or captured from the field: Provide species, sex, and age where possible. |  | N/A |
|  |  |  |
| **Plants and microbes** | **Indicate where provided: section/figure legend** | **N/A** |
| Plants: provide species and strain, ecotype and cultivar where relevant, unique accession number if available, and source (including location for collected wild specimens). |  | N/A |
| Microbes: provide species and strain, unique accession number if available, and source. |  | N/A |
|  |  |  |
| **Human research participants** | **Indicate where provided: section/figure legend) or state if these demographics were not collected** | **N/A** |
| If collected and within the bounds of privacy constraints report on age, sex, gender and ethnicity for all study participants. |  | **N/A** |

**Design:**

|  |  |  |
| --- | --- | --- |
| **Study protocol** | **Indicate where provided: section/figure legend** | **N/A** |
| If the study protocol has been pre-registered, provide DOI. For clinical trials, provide the trial registration number OR cite DOI. |  | N/A |
|  |  |  |
| **Laboratory protocol** | **Indicate where provided: section/figure legend** | **N/A** |
| Provide DOI OR other citation details if detailed step-by-step protocols are available. | Replication timing assay:   doi: 10.3791/4400. |  |
|  |  |  |
| **Experimental study design (statistics details) \*** | | |
| **For in vivo studies: State whether and how the following have been done** | **Indicate where provided: section/figure legend. If it could have been done, but was not, write “not done”** | **N/A** |
| Sample size determination |  | N/A |
| Randomisation |  | **N/A** |
| Blinding |  | N/A |
| Inclusion/exclusion criteria |  | N/A |
|  |  |  |
| **Sample definition and in-laboratory replication** | **Indicate where provided: section/figure legend** | **N/A** |
| State number of times the experiment was replicated in the laboratory. | RNA-DNA FISH: Figure 1A and 1B legend, 2 replicates in each cell type. Figure 2G-J, 12 independent clones assayed for sense or antisense transgenes. Figure 3, each shRNA treatment of K562 was done 2 times on different days. Quantitation of XIST RNA clouds were on >20 cells for each shRNA assay (Figure 3W). Figure 4A, 50 clones were screened for colocalization of Xist and ASAR6-141 RNA. One clone was used for this figure. Figure 5, each shRNA treatment of HTD114 was done 2 times on different days. Figure 7, mouse cells containing the ASAR6 BAC integrated into chromosome 3 were previously published (Donley, et al. 2013). C0T-1 RNA FISH in combination with ASAR6 DNA was done 2 times.  Replication timing assays: Figure 2A-E legend, 2 independent clones from each knockout genotype, were assayed. Figure 2F-2L legend, 3 different clones with integrations into different chromosomes were assayed 2 different times for the sense and antisense transgenes. Figure 4, one clone, out of 62 screened, representing the sense transgene integrated into the X chromosome was identified and assayed. Figure 6, 2 independent replicates for each shRNA treatment were carried out. |  |
| Define whether data describe technical or biological replicates. | Figure 1A and 1B, biological replicates in two different cell lines. Figure 1C-1G, the ENCODE data represents two biological replicates for eCLIP assay on 2 different cell types, K562 and HepG2.  Figure 2C-2E, two biological replicates were assayed for each genotype. Figure 2J-2L, 12 clones from the sense or antisense transgenes were screened for expression and 3 clones each were assayed in biological replicates for replication timing.  Figure 3 and 5, two biological replicates for each shRNA treatment of K562 (Figure 3) or HTD114 (Figure 5) were assayed for protein expression and for RNA FISH.  Figure 4, one biological replicate of the transgene integrated into the X chromosome was assayed. 2 biological replicates of the replication timing assay were carried out.  Figure 6, two biological replicates of the C0T-1 RNA FISH were carried out with similar results. |  |
|  |  |  |
| **Ethics** | **Indicate where provided: section/submission form** | **N/A** |
| Studies involving human participants: State details of authority granting ethics approval (IRB or equivalent committee(s), provide reference number for approval. |  | N/A |
| Studies involving experimental animals: State details of authority granting ethics approval (IRB or equivalent committee(s), provide reference number for approval. |  | N/A |
| Studies involving specimen and field samples: State if relevant permits obtained, provide details of authority approving study; if none were required, explain why. |  | N/A |
|  |  |  |
| **Dual Use Research of Concern (DURC)** | **Indicate where provided: section/submission form** | **N/A** |
| If study is subject to dual use research of concern regulations, state the authority granting approval and reference number for the regulatory approval. |  | N/A |

**Analysis:**

|  |  |  |
| --- | --- | --- |
| **Attrition** | **Indicate where provided: section/figure legend** | **N/A** |
| Describe whether exclusion criteria were pre-established. Report if sample or data points were omitted from analysis. If yes, report if this was due to attrition or intentional exclusion and provide justification. | No data points were excluded from any analysis reported. |  |
|  |  |  |
| **Statistics** | **Indicate where provided: section/figure legend** | **N/A** |
| Describe statistical tests used and justify choice of tests. | Figure 1: P-values were generated for eCLiP data using the “region-based” method described in 26. For 10kb sliding windows across the genome the log2-Ratio was calculated between the number of reads in the eCLiP sample and matched control, and P-values were generated using the Python scipy.stats function “zscore”. FDR correction was performed using the Python Statsmodels function “fdrcorrection”.  Figure 2E, 2I 3K, 3W, 6E-F: P values were calculated using the Kruskal-Wallis test. |  |
|  |  |  |
| **Data availability** | **Indicate where provided: section/submission form** | **N/A** |
| For newly created and reused datasets, the manuscript includes a data availability statement that provides details for access (or notes restrictions on access). |  |  |
| When newly created datasets are publicly available, provide accession number in repository OR DOI and licensing details where available. |  |  |
| If reused data is publicly available provide accession number in repository OR DOI, OR URL, OR citation. |  |  |
|  |  |  |
| **Code availability** | **Indicate where provided: section/figure legend** | **N/A** |
| For any computer code/software/mathematical algorithms essential for replicating the main findings of the study, whether newly generated or re-used, the manuscript includes a data availability statement that provides details for access or notes restrictions. | We used standard code available from the SAMtools [PMC7931819](http://www.ncbi.nlm.nih.gov/pmc/articles/pmc7931819/), and BEDtools [doi.org/10.1093/bioinformatics/btq033](https://doi.org/10.1093/bioinformatics/btq033) software packages. |  |
| Where newly generated code is publicly available, provide accession number in repository, OR DOI OR URL and licensing details where available. State any restrictions on code availability or accessibility. |  | N/A |
| If reused code is publicly available provide accession number in repository OR DOI OR URL, OR citation. | SAMtools [PMC7931819](http://www.ncbi.nlm.nih.gov/pmc/articles/pmc7931819/)  BEDtools: [doi.org/10.1093/bioinformatics/btq033](https://doi.org/10.1093/bioinformatics/btq033) |  |

**Reporting:**

The MDAR framework recommends adoption of discipline-specific guidelines, established and endorsed through community initiatives.

|  |  |  |
| --- | --- | --- |
| **Adherence to community standards** | **Indicate where provided: section/figure legend** | **N/A** |
| State if relevant guidelines (e.g., ICMJE, MIBBI, ARRIVE, STRANGE) have been followed, and whether a checklist (e.g., CONSORT, PRISMA, ARRIVE) is provided with the manuscript. |  | N/A |

\* We provide the following guidance regarding transparent reporting and statistics; we also refer authors to [Ten common statistical mistakes to watch out for when writing or reviewing a manuscript](https://doi.org/10.7554/eLife.48175).

**Sample-size estimation**

* You should state whether an appropriate sample size was computed when the study was being designed
* You should state the statistical method of sample size computation and any required assumptions
* If no explicit power analysis was used, you should describe how you decided what sample (replicate) size (number) to use

**Replicates**

* You should report how often each experiment was performed
* You should include a definition of biological versus technical replication
* The data obtained should be provided and sufficient information should be provided to indicate the number of independent biological and/or technical replicates
* If you encountered any outliers, you should describe how these were handled
* Criteria for exclusion/inclusion of data should be clearly stated
* High-throughput sequence data should be uploaded before submission, with a private link for reviewers provided (these are available from both GEO and ArrayExpress)

**Statistical reporting**

* Statistical analysis methods should be described and justified
* Raw data should be presented in figures whenever informative to do so (typically when N per group is less than 10)
* For each experiment, you should identify the statistical tests used, exact values of N, definitions of center, methods of multiple test correction, and dispersion and precision measures (e.g., mean, median, SD, SEM, confidence intervals; and, for the major substantive results, a measure of effect size (e.g., Pearson's r, Cohen's d)
* Report exact p-values wherever possible alongside the summary statistics and 95% confidence intervals. These should be reported for all key questions and not only when the p-value is less than 0.05.

**Group allocation**

* Indicate how samples were allocated into experimental groups (in the case of clinical studies, please specify allocation to treatment method); if randomization was used, please also state if restricted randomization was applied
* Indicate if masking was used during group allocation, data collection and/or data analysis