

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
<input type="checkbox"/>	<input checked="" type="checkbox"/> The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
<input type="checkbox"/>	<input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
<input type="checkbox"/>	<input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
<input type="checkbox"/>	<input checked="" type="checkbox"/> A description of all covariates tested
<input type="checkbox"/>	<input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
<input type="checkbox"/>	<input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
<input type="checkbox"/>	<input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
<input checked="" type="checkbox"/>	<input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
<input checked="" type="checkbox"/>	<input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
<input type="checkbox"/>	<input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	Code is available at https://github.com/bfairkun/ChromatinSplicingQTLs . This includes a reproducible pipeline with specific software versions and all new code. More description, including the exact version for all softwares used, is available in the Methods
Data analysis	Code is available at https://github.com/bfairkun/ChromatinSplicingQTLs . This includes a reproducible pipeline with specific software versions and all new code. More description, including the exact version for all softwares used, is available in the Methods

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Sequence data (naRNA-seq and H3K36ME3 Cut&Tag) is publicly available and has been deposited in Gene Expression Omnibus (GEO, www.ncbi.nlm.nih.gov/geo/)

under accession GSE252006. Data aligned to GRCh38 and transcript release v34 annotations from GenCode (<https://www.encodegenes.org/human/>). Some analyses (see Supplemental Methods) also utilized v37 annotations.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	We aimed for a sex-balanced study design. naRNA-seq dataset contained 41 female, 46 male. H3K36me3 dataset contained 45 female, 48 male
Reporting on race, ethnicity, or other socially relevant groupings	We utilized lymphoblastoid cell lines (LCLs) in the NIGMS cell repository from Yoruban ancestry donors (YRI collection), as this collection of cell lines has been extensively molecularly characterized for molecular QTL studies. Genetic ancestry was not explicitly accounted for in QTL mapping analyses, though all YRI lines utilized are unrelated and from similar ancestry.
Population characteristics	See above
Recruitment	Cell lines were obtained from Coriell, under the MTA agreement: MTA000215_Li_Coriell_MTA_3_3_2021
Ethics oversight	n/a

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size is typical for similar molecular QTL studies: at least n>50, with greater sample sizes typically being better powered for more discoveries. We were able to achieve a sample size of n=93 lines for H3K36me3 Cut&Tag, and n=86 cell lines for naRNA-seq. No statistical method was used to predetermine sample size for novel data (naRNA-seq, H3K36ME3 CUT&Tag, RNA-seq and naRNA-seq at various risdiplam doses). We note that our sample sizes are similar to previous QTL studies using this panel of cell lines (ref 37,38,40) and risdiplam transcriptomics studies (ref 45,48)
Data exclusions	As described in Statistics and Reproducibility Methods section: For eQTL and sQTL mapping using naRNA, we excluded line NA18855. This was not a predetermined decision. Rather, principal component analysis revealed this sample as an outlier, and differential expression analysis against other lines identified differentially expressed genes clustered along large sections of chromosomes, suggestive of chromosomal abnormalities.
Replication	For various analyses, we replicated our initial analyses by reanalysis of published datasets (e.g., the conclusion that ~2.3% of splice junctions are unproductive in our initial naRNA-seq analysis was replicated in published RNA-seq of cell lines subject to knockdown of NMD-factors). Each of our experiments was performed once, with no additional replication. Risdiplam experiment was performed once, with biological replicates (separate dishes) for each treatment dose for naRNA-seq, and a single replicate for each dose in steady-state RNA. Analyses utilized replicates for maximum power, and discoveries were not further replicated in independent datasets
Randomization	The experiments were not randomized, as molQTL studies typically do not require randomization for confounders, which are assumed to already be randomized with respect to genotype. Experiments with risdiplam used a single isogenic cell line, grown in a single well-controlled batch, such that randomization is not needed.
Blinding	The Investigators were not blinded to allocation during experiments and outcome assessment, as cell lines (for QTL studies) and treatment were required knowledge for carrying out the experiments

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used	H3K36Me3 CUT&Tag used 1uL of the following antibody per sample: Abcam catalog no. ab9050, lot GR3386101-2.
Validation	We validated by analyzing meta gene profile of CUT&Tag sequencing data (Figure S2). Further, manufacturer states "Our Abpromise guarantee covers the use of ab9050 in the following tested applications: ICC/ICF, WB, ChIP" (https://www.abcam.com/products/primary-antibodies/histone-h3-tri-methyl-k36-antibody-chip-grade-ab9050.html?productWallTab=Abreviews)

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)	LCL line GM12878. Coriell cell repository, NIGMS collection, YRI sub-collection LCL lines: NA18486,NA18497,NA18498,NA18499,NA18502,NA18504,NA18505,NA18507,NA18508,NA18510,NA18511,NA18516,NA18519,NA18520,NA18522,NA18523,NA18852,NA18853,NA18855,NA18858,NA18861,NA18862,NA18864,NA18867,NA18868,NA18870,NA18876,NA18877,NA18879,NA18881,NA18907,NA18909,NA18910,NA18912,NA18913,NA18915,NA18917,NA18923,NA18924,NA18934,NA19092,NA19093,NA19095,NA19096,NA19098,NA19099,NA19101,NA19102,NA19107,NA19108,NA19114,NA19117,NA19118,NA19119,NA19121,NA19122,NA19127,NA19128,NA19130,NA19131,NA19137,NA19138,NA19140,NA19141,NA19143,NA19146,NA19147,NA19150,NA19152,NA19153,NA19159,NA19160,NA19171,NA19184,NA19190,NA19193,NA19195,NA19196,NA19198,NA19200,NA19201,NA19203,NA19206,NA19207,NA19209,NA19210,NA19213,NA19214,NA19223,NA19225,NA19236,NA19238,NA19239,NA19247,NA19257
Authentication	Identity of all cell lines was verified from molecular profiling sequencing data (eg assessment of SNP genotypes in H3K36me# CUT&Tag or naRNA-seq data)
Mycoplasma contamination	cell lines were not tested for mycoplasma
Commonly misidentified lines (See ICLAC register)	No commonly misidentified lines were used in this study

Plants

Seed stocks	n/a
Novel plant genotypes	n/a
Authentication	n/a