

Reporting Summary

Nature Research 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 Research policies, see [Authors & Referees](#) 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

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

The code for the GENerative adversarial Tensorial Reinforcement Learning (GENTRL) system is provided with the manuscript. It will be freely available at GitHub upon the acceptance of the manuscript.

Data analysis

Maestro Release 2018-3; Phoenix WinNonlin 6.3; custom python scripts using PyTorch 0.4.1 and RDKit 2018.03.4 libraries

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/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [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 list of figures that have associated raw data
- A description of any restrictions on data availability

The 30,000 structures generated by GENTRL for the DDR1 kinase are available in supplementary materials.

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

Life sciences study design

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

Sample size	No applicable.
Data exclusions	Data outside 3xSD were excluded.
Replication	Results are verified by independent biological experiments.
Randomization	Not applicable.
Blinding	Not applicable.

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		Methods	
n/a	Involved in the study	n/a	Involved in the study
☐	☒ Antibodies	☒	☐ ChIP-seq
☐	☒ Eukaryotic cell lines	☒	☐ Flow cytometry
☒	☐ Palaeontology	☒	☐ MRI-based neuroimaging
☐	☒ Animals and other organisms		
☒	☐ Human research participants		
☒	☐ Clinical data		

Antibodies

Antibodies used	Mouse anti- human COL1A1 antibody (3G3) (Santa Cruz, sc-293182); mouse anti-human CTGF antibody (E-5) (Santa Cruz, sc-365970); mouse anti human α -smooth muscle actin antibody (SPM332) (Santa Cruz, sc-56499); mouse anti-human GLYCERALDEHYDE-3-PDH, (Merck-Millipore, Merck-MAB374); rabbit anti-Phospho-DDR1 (Tyr513) (E1N8F) (Cell Signaling, 14531S); mouse anti-HA antibody (Sigma, H9658-200UL); goat Anti-Mouse IgG H&L (HRP) (Abcam, ab205719); goat Anti-Rabbit IgG H&L (HRP) (Abcam, ab205718)
Validation	The antibodies were validated by the manufacturers for Western blot.

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	MRC-5 was from ATCC (CCL-171); LX-2 was from Merck Millipore (SCC064); U-2 OS-DDR1 stable cell line was made at RSD Biology Department, WuXi Apptec, Inc. from U-2 OS parent cell line from ATCC® (HTB96™);
Authentication	The cells were authenticated by STR.
Mycoplasma contamination	All cell lines used have been tested free of mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	None.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	Mouse, C57BL/6, male, 7- 9 weeks
Wild animals	Study did not involve wild animals
Field-collected samples	Following arrival at WuXi AppTec animals were assessed as to their general health by a member of the veterinary staff or other authorized personnel. Animals were acclimated for at least 3 days (upon arrival at WuXi AppTec) before being placed on study.

Animals were group housed during acclimation and individually housed during the study. The animal room environment was controlled (target conditions: temperature 18 to 26°C, relative humidity 30 to 70%, 12 hours artificial light and 12 hours dark). Temperature and relative humidity was monitored daily. The animals were overnight fasted. They had access to Certified Rodent Diet ad libitum 4 hr post dose. The lot number and specifications of each lot used were archived at WuXi AppTec. Water was autoclaved before provided to the animals ad libitum. Periodic analyses of the water was performed and the results archived at WuXi AppTec. There are no known contaminants in the diet or water that, at the levels of detection, is expected to interfere with the purpose, conduct or outcome of the study.

Ethics oversight

Study was permitted by IACUC (Institutional Animal Care and Use Committee).

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