LETTER

doi:10.1038/nature11884

Activating RNAs associate with Mediator to enhance chromatin architecture and transcription

Fan Lai[1], Ulf A. Orom[2], Matteo Cesaroni[1], Malte Beringer[3], Dylan J. Taatjes[4], Gerd A. Blobel[5] & Ramin Shiekhattar[1]

Recent advances in genomic research have revealed the existence of a large number of transcripts devoid of protein-coding potential in multiple organisms[1–8]. Although the functional role for long non-coding RNAs (lncRNAs) has been best defined in epigenetic phenomena such as X-chromosome inactivation and imprinting, different classes of lncRNAs may have varied biological functions[8–13]. We and others have identified a class of lncRNAs, termed ncRNA-activating (ncRNA-a), that function to activate their neighbouring genes using a *cis*-mediated mechanism[5,14–16]. To define the precise mode by which such enhancer-like RNAs function, we depleted factors with known roles in transcriptional activation and assessed their role in RNA-dependent activation. Here we report that depletion of the components of the co-activator complex, Mediator, specifically and potently diminished the ncRNA-induced activation of transcription in a heterologous reporter assay using human HEK293 cells. *In vivo*, Mediator is recruited to ncRNA-a target genes and regulates their expression. We show that ncRNA-a interact with Mediator to regulate its chromatin localization and kinase activity towards histone H3 serine 10. The Mediator complex harbouring disease-causing MED12 mutations[17,18] displays diminished ability to associate with activating ncRNAs. Chromosome conformation capture confirmed the presence of DNA looping between the ncRNA-a loci and its targets. Importantly, depletion of Mediator subunits or ncRNA-a reduced the chromatin looping between the two loci. Our results identify the human Mediator complex as the transducer of activating ncRNAs and highlight the importance of Mediator and activating ncRNA association in human disease.

To define the transcriptional complex(es) that orchestrate the responsiveness of activating lncRNAs, we used HEK293 stable cell lines expressing the heterologous TK promoter driving luciferase expression fused to *ncRNA-a7* (also called *LINC00651*) controlled by its own natural promoter and depleted factors known to be involved in transcriptional activation and enhancer function. We also developed stable lines expressing control reporters, in which a DNA fragment devoid of any transcriptional activity was substituted for the *ncRNA-a7* genomic region (Fig. 1a). Importantly, depletion of ncRNA-a7 using two different short interfering RNAs (siRNAs) specifically decreased the transcription of the reporter construct containing ncRNA-a7 (Fig. 1b). Next, we depleted factors known to be involved in transcriptional activation (Supplementary Fig. 1a) and examined the responsiveness of cell lines harbouring the constructs expressing ncRNA-a7 or control constructs devoid of lncRNAs. Interestingly, only depletion of the Mediator subunit MED12 displayed a differential effect on the transcription of the reporter construct containing ncRNA-a7 (Fig. 1c). Depletion of other factors was either ineffective in changing transcriptional output (CDK9, CCNT1, NIPBL or WDR5) or reduced the transcriptional levels for both constructs (GTF2B, p300, SMC1) (Fig. 1c).

Mediator contains a complex and modular subunit composition[19,20]. We depleted different components of Mediator corresponding to the head, middle or the tail modules and assessed their effect on the responsiveness of the cell lines harbouring the constructs driven by ncRNA-a7 or the control construct. Overall, depletion of other subunits of the Mediator complex also reduced the ncRNA-a7-induced activation, although there was also a small reduction in transcription of the control plasmid with depletion of some of the subunits (Fig. 1d and

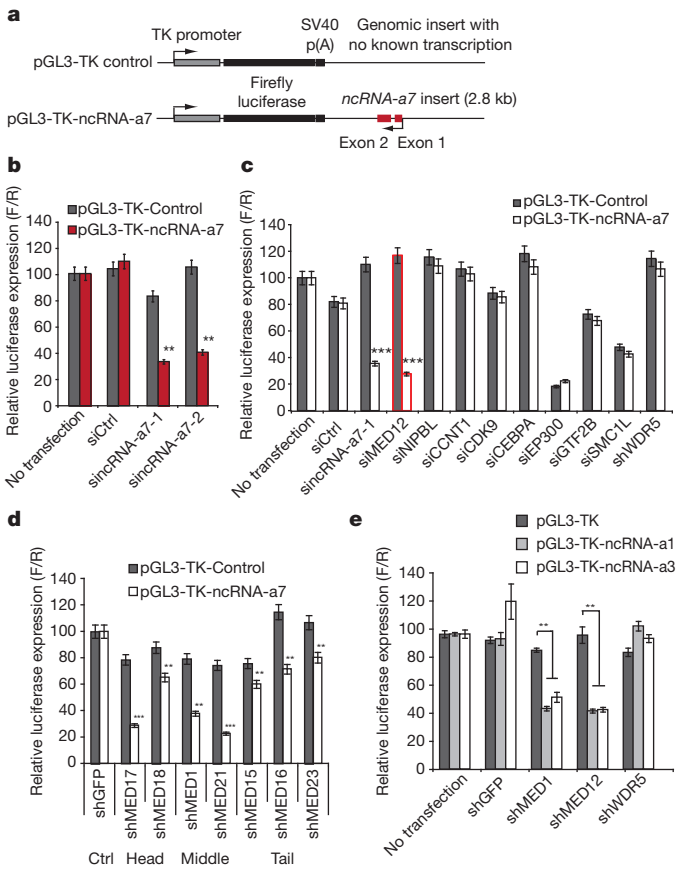

Figure 1 | Mediator confers the ncRNA-a-dependent activation of a heterologous reporter. **a**, Schematic representation of genomic control insert (top) or *ncRNA-a7* insert (bottom) in luciferase reporter vector driven by a TK promoter. **b**, Depletion of ncRNA-a7 reduces the luciferase activity in the ncRNA-a7 luciferase reporter cell lines. **c**, Depletion of the transcription factors or enhancers in the control or ncRNA-a7 reporter cell lines. The red bar indicates reduced transcription using siRNA against *MED12*, comparable with siRNA against *ncRNA-a7*. **d**, Depletion of different Mediator subunits (head, middle and tail) using the ncRNA-a7 reporter cell lines. **e**, Depletion of the Mediator subunits in ncRNA-a1 or ncRNA-a3 luciferase reporter cell lines. All data shown are mean ± s.e.m. of three independent experiments. $^{**}P < 0.01$, $^{***}P < 0.001$ by two-tailed Student's *t*-test. F/R, firefly luciferase/*Renilla* luciferase.

[1]The Wistar Institute, 3601 Spruce Street, Philadelphia, Pennsylvania 19104, USA. [2]Max Planck Institute for Molecular Genetics, Ihnestr, 14195 Berlin, Germany. [3]Center de Regulacio Genomica and UPF, Barcelona 08003, Spain. [4]Department of Chemistry and Biochemistry, University of Colorado, 215 UCB, Boulder, Colorado 80309, USA. [5]Division of Hematology, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA.

Supplementary Fig. 1b). Similar to its effects on ncRNA-a7, depletion of the Mediator subunits decreased the RNA-induced activation of two other ncRNA-a, which were previously shown to regulate *ECM1* and *TAL1* genes (Fig. 1e). Taken together, these results identify the Mediator complex as a transducer of ncRNA-a function.

We next depleted the Mediator subunits and assessed their effect on ncRNA-a targets *in vivo*. We previously found that ncRNA-a7 and ncRNA-a3 regulate *SNAI1* and *TAL1* genes, respectively[5]. Moreover, using gene expression arrays, we had observed a profound decrease in *AURKA* transcription when ncRNA-a7 was depleted[5]. We confirmed these results by depleting ncRNA-a7 and ncRNA-a3, which led to a decrease in *SNAI1*, *AURKA* and *TAL1* transcript levels using real-time PCR (Fig. 2a, b). We next depleted two different subunits of the Mediator complex, MED1 and MED12, and assessed the *SNAI1*, *AURKA* and *TAL1* transcript levels (Supplementary Fig. 1b). Importantly, depletion of Mediator subunits decreased transcription of all three targets similar to that shown with depletion of ncRNA-a (Fig. 2a, b, the affected genes are depicted by a black bar). Interestingly, depletion of Mediator subunits did not affect the transcript levels of genes that are not regulated by ncRNA-a (*UBE2V1*, *CSTF*, *STIL*, depicted by a white bar in Fig. 2a, b). However, whereas ncRNA-a7 transcription was not affected after Mediator depletion, there was a decrease in ncRNA-a3 levels (Fig. 2a, b).

To ensure that such transcriptional effects after Mediator depletion were direct, we performed chromatin immunoprecipitation (ChIP) using antibodies against the MED12 and MED1 subunits of Mediator. This analysis revealed that the Mediator complex occupied the promoters of ncRNA-a3, ncRNA-a7 and their targets (Fig. 2c, d). Moreover, we found RNA polymerase II (RNAPII) at these promoters, consistent with Mediator occupancy (Fig. 2c, d). Importantly, depletion of ncRNA-a7 or ncRNA-a3 decreased the occupancy of Mediator and RNAPII at their target genes without affecting the occupancy at genes not regulated by ncRNA-a (Fig. 2c, d). These results indicate that the Mediator complex occupies the promoters of genes regulated by ncRNA-a, and depletion of ncRNA decreases the occupancy of Mediator at these sites.

The Mediator complex, via its CDK8 module, displays kinase activity towards histone H3 serine 10 (H3S10), a histone modification with strong ties to transcriptional activation[21,22]. We next assessed whether ncRNA-a regulate Mediator kinase activity towards histone H3. Addition of ncRNA-a7 or ncRNA-a3 led to a significant stimulation of the Mediator kinase activity towards histone H3, whereas the addition of primary let7b ncRNA or ncRNA HOTAIR had no effect (Fig. 2e, f and Supplementary Fig. 2a). Such stimulation of Mediator kinase activity by ncRNA-a was specific towards histone H3 as the substrate. We did not observe any stimulation when GST–CTD (corresponding to the C-terminal domain of RPB1) or cyclin H was used as substrates (Supplementary Fig. 2b, c). Moreover, consistent with a role for ncRNA-a in stimulation of histone H3 phosphorylation, depletion of ncRNA-a7 or ncRNA-a3 led to a specific decrease in H3S10 levels at *SNAI1*, *AURKA* and *TAL1* (Supplementary Fig. 2d and e).

We next asked whether ncRNA-a could associate with the Mediator complex. We isolated Mediator using stable cell lines expressing Flag-epitope-tagged MED12 and examined its interaction with *in-vitro*-transcribed ncRNA-a7. An eluate from a stable Flag–GFP cell line was used as a control (Fig. 3a). Whereas the control primary let7 transcript (pri-let7) or the ncRNA HOTAIR (a 429-nucleotide RNA with a similar base composition as ncRNA-a7) did not specifically interact with the Mediator complex, we detected a robust and specific association between ncRNA-a7 and the Mediator complex (Fig. 3a). Notably, Mediator also associated with ncRNA-a1 and ncRNA-a3, two other ncRNA-a that were shown to activate transcription of their neighbouring genes[5] (Fig. 3b).

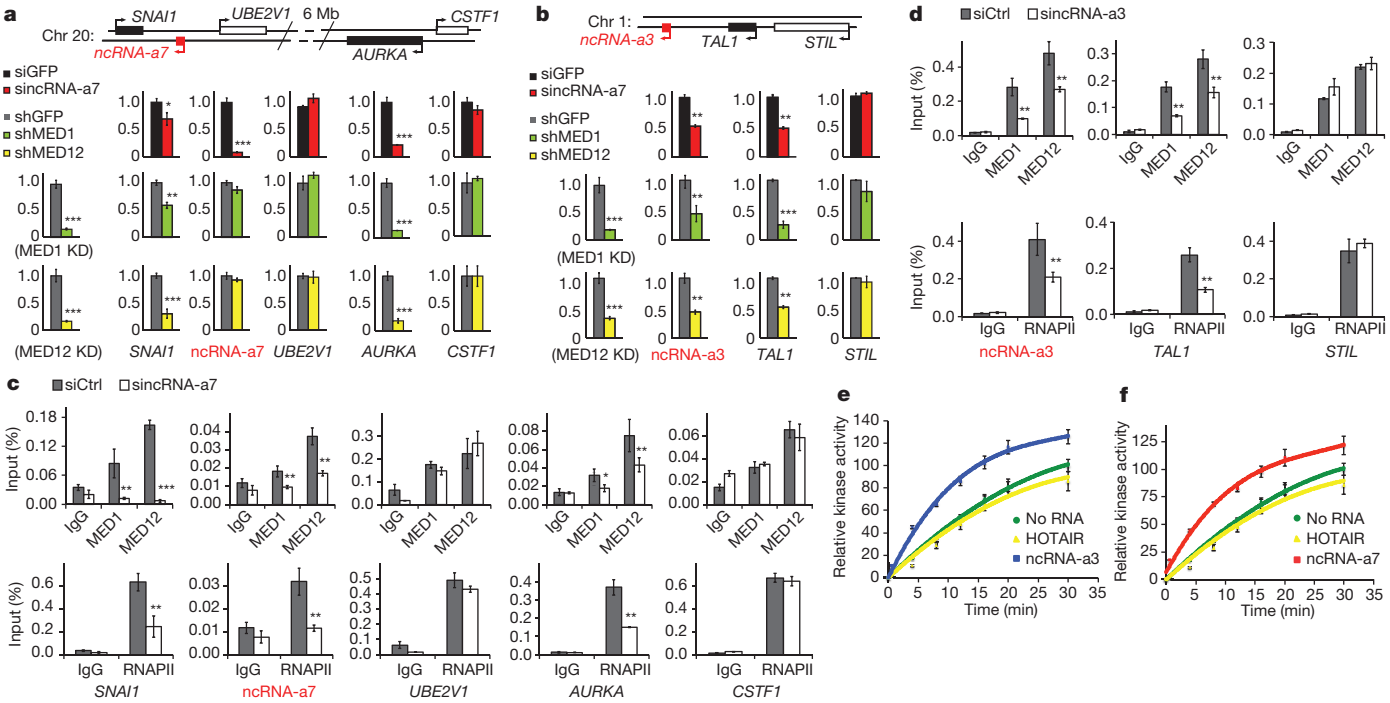

Figure 2 | Functional association of Mediator and activating ncRNAs. **a**, Knockdown of ncRNA-a7 or Mediator subunits decreases expression of *SNAI1* and *AURKA*. Depletion of ncRNA-a7 diminished *SNAI1* and *AURKA* expression as detected by real-time PCR (top row, red bars). Similarly, knockdown (KD) of MED1 (green bars) or MED12 (yellow bars) reduced expression of *SNAI1* or *AURKA*. **b**, Knockdown of ncRNA-a3 or Mediator subunits decreased *TAL1* expression. Expression of *TAL1* or control *STIL* after knockdown of ncRNA-a3 (red bars), MED1 (green bars) or MED12 (yellow bars) is shown. **c**, **d**, Knockdown of ncRNA-a7 (**c**) or ncRNA-a3 (**d**) reduces the genomic occupancy of MED1, MED12 (top row) or RNAPII (bottom row) on *SNAI1*, *AURKA* (**c**) or *TAL1* (**d**) in A549 cells. **e**, **f**, Activating ncRNAs specifically stimulate Mediator kinase activity towards histone H3.1 substrate *in vitro*. Quantification of kinase assay after addition of ncRNA-a7 (**e**) or ncRNA-a3 (**f**) is shown. Error bars represent ± s.e.m. ($n = 3$), $P < 0.01$ by two-tailed Student's *t*-test. The mean ± s.e.m. for all results represents three independent experiments. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ by two-tailed Student's *t*-test.

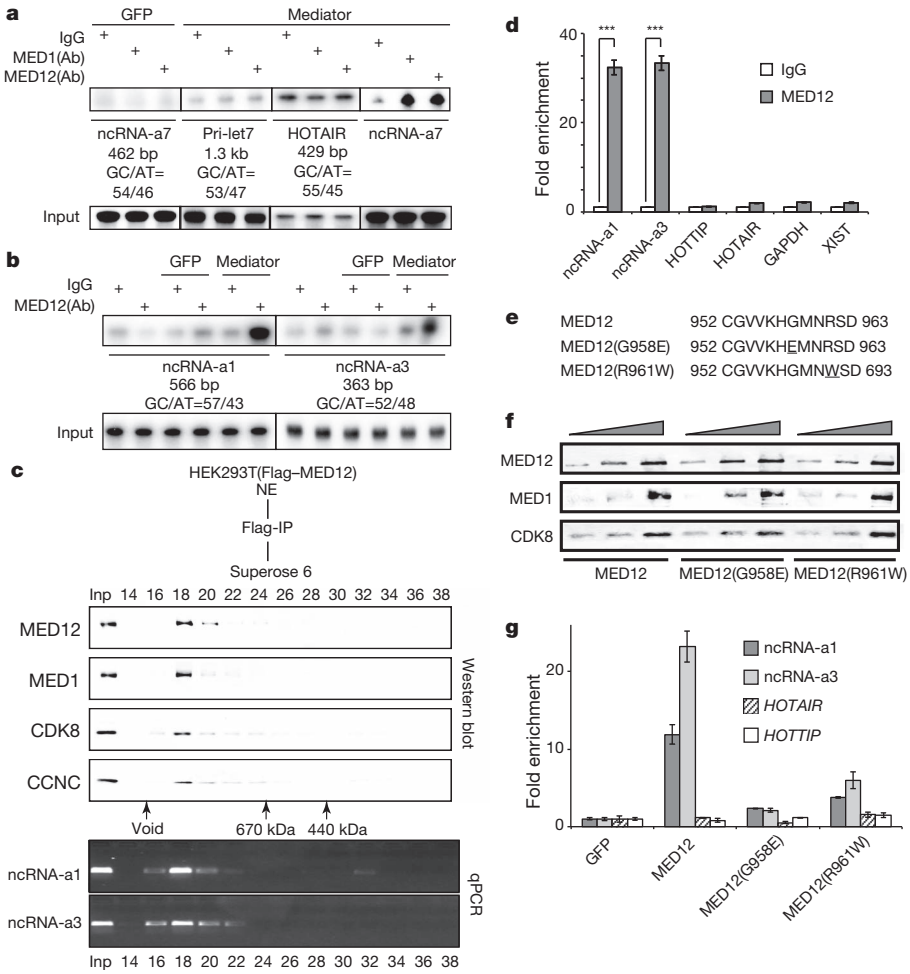

Figure 3 | Interaction of Mediator and activating ncRNAs is disrupted by FG syndrome mutations of MED12. **a**, **b**, ncRNA-a associate with the Mediator complex *in vitro*. RNA immune precipitation (RIP) was performed using protein A dynabeads coupled with IgG, MED1 or MED12 antibodies. Identical results were obtained in three independent experiments. **c**, Purification of Mediator protein complex and associated RNAs. The purified Flag–MED12 affinity elution was subjected to Superose 6 size-exclusion chromatography and the Mediator subunits were detected by western blot. The Superose 6 fractions (14–38) are shown on top and molecular mass markers (kilodaltons) are at the bottom. The RNAs in each fraction were extracted followed by RT–PCR using specific primers. **d**, Ultraviolet crosslink followed by RIP using HEK293T whole-cell lysates. The associated RNAs were analysed by RT–qPCR with individual specific RNA primers. The mean ± s.e.m. are from three independent experiments. $***P < 0.001$ by two-tailed Student's *t*-test. **e**, Protein sequences of MED12 mutants causing FG syndrome. Part of the protein sequence of MED12 exon 21 is shown (top tow). The mutated sites indicate the G958E mutation (underlined) and the R961W mutation (underlined). **f**, HEK293T cells were transfected with Flag–MED12 or mutant constructs. After ultraviolet crosslinking, total cell extracts were immunoprecipitated and analysed by western blotting. **g**, After UV-RIP of wild-type and mutant MED12 samples shown in **f**, RT–qPCR was performed using transcript-specific primers from the lncRNAs shown. The data are representative of three independent experiments. The two-tailed Student's *t*-test indicated a $P < 0.001$ for ncRNA-a1 and ncRNA-a3, whereas there was no significant difference for *HOTAIR* and *HOTTIP*.

To assess the association of endogenous ncRNA-a and Mediator, the Flag–MED12 affinity-purified Mediator was fractionated by gel filtration and column fractions were subjected to western blot analysis and polymerase chain reaction with reverse transcription (RT–PCR) to detect the associated ncRNA-a (Fig. 3c). Whereas the HEK293 nuclear extract contained similar levels of ncRNA-a1, ncRNA-a3, HOTAIR and HOTTIP (HEK293 cells do not express ncRNA-a7; Supplementary Fig. 3a), we could only detect ncRNA-a1 and ncRNA-a3 in association with the affinity-purified Mediator (Fig. 3c). Indeed, ncRNA-a1 and ncRNA-a3 displayed a co-elution with components of the Mediator complex on gel filtration, further supporting their association with Mediator (Fig. 3c). Next, we performed RNA immunoprecipitation following ultraviolet crosslinking (UV-RIP) with anti-MED12 antibodies or IgG, as control. Activating ncRNA-a1 and ncRNA-a3 were significantly enriched in the MED12 UV-RIP, whereas other abundant control messenger RNAs or non-coding RNAs such as *GAPDH*, *HOTAIR*, *HOTTIP* and *XIST* were not (Fig. 3d). Two mutations in MED12 proteins have been linked to the development of the human genetic disorder Opitz–Kaveggia syndrome (also know as FG syndrome), causing physical anomalies and developmental delays[17,18] (Fig. 3e). Notably, whereas such mutations in MED12 do not affect its association with other Mediator subunits, they significantly diminish its association with ncRNA-a1 and ncRNA-a3 as measured by UV-RIP (Fig. 3f, g).

We next asked whether we could detect long-range chromatin looping between the *ncRNA-a7* locus and its targets, *SNAI1* and *AURKA* and whether ncRNA-a7 and the Mediator complex have a role in such an association. We performed chromosome conformation capture (3C) to assess the association of ncRNA-a7 with *AURKA* as well as *SNAI1*. We also performed a similar analysis between ncRNA-a3 and the *TAL1* gene. We used anchoring points near the 3′ untranslated region of ncRNA-a7 or ncRNA-a3 to measure the extent of chromatin looping between the *ncRNA-a7* or *ncRNA-a3* locus and their targets as shown in Fig. 4a, d and Supplementary Fig. 3b, c. This analysis revealed a strong association between the *ncRNA-a7* locus and the region encompassing the promoter region of *SNAI1* (Fig. 4b, upper panel).

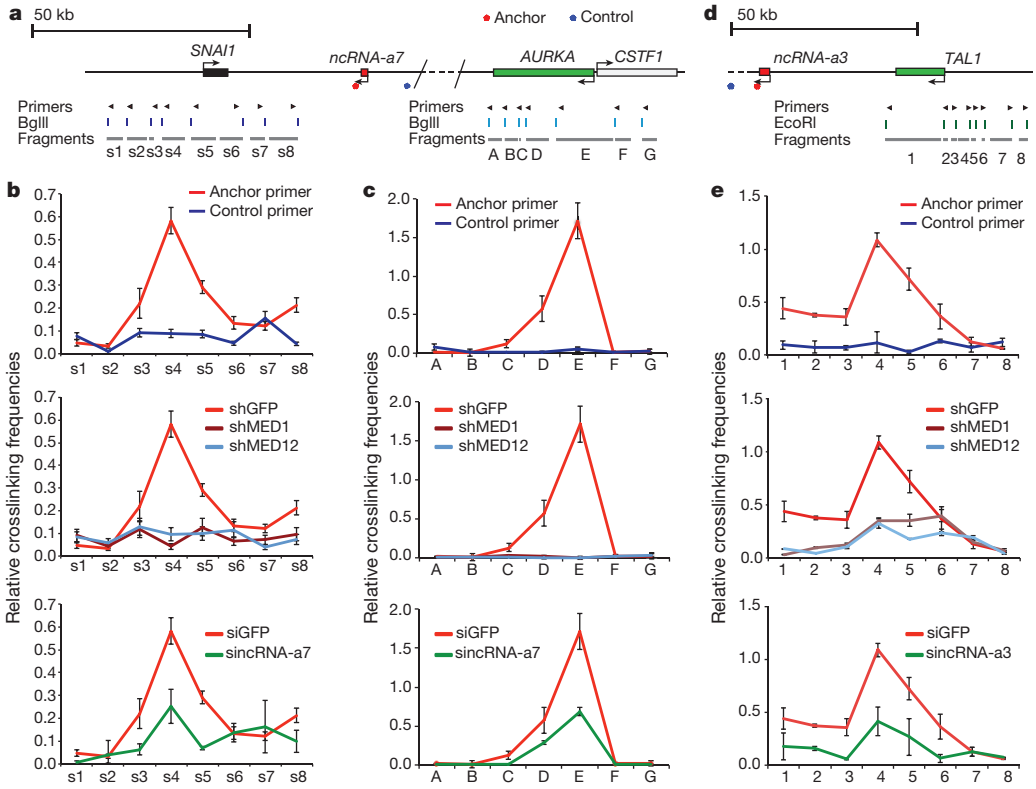

Figure 4 | Mediator complex and ncRNA-a promote chromatin looping. The schematic diagrams represent the genomic locus between the *ncRNA*-a7, *SNAI1* and *AURKA* loci (**a**), and the *ncRNA*-a3 and *TAL1* loci (**d**). The top arrows show the position of primers, the digestion sites are shown in the middle, and the s1–s8, A–G or 1–8 fragments are presented below. The looping events between ncRNAs and their targets were detected between ncRNA-a7 and *SNAI1* (**b**), ncRNA-a7 and *AURKA* (**c**) and ncRNA-a3 and *TAL1* (**e**) using chromosome conformation capture (3C). Depletion of MED1 or MED12 abolished the loop interaction (**b**, **c**, **e**; middle panels). Knockdown of ncRNA-a7 or ncRNA-a3 reduced the chromosomal looping events (**b**, **c**, **e**; lower panels). The interaction frequency between the anchoring points and distal fragments was determined by real-time PCR and normalized to BAC templates and control anchors. Each error bar represents ± s.e.m. from three independent experiments, $P < 0.01$ by two-tailed Student's *t*-test. Representative gel images of the 3C experiments for *AURKA* are presented in Supplementary Fig. 3.

Moreover, ncRNA-a7 also displayed an interaction with regions near the 5′ end of the *AURKA* gene that extended into the body of the gene (Fig. 4c, upper panel). There was no detectable association when a control anchor was placed in the genomic region between *ncRNA-a7* and *AURKA* (Fig. 4b, c).

Next we depleted the Mediator subunits MED1 or MED12 and assessed their potential role in mediating the chromosomal looping between ncRNA-a7, *SNAI1* and *AURKA*. Depletion of either MED1 or MED12 completely abrogated the chromosomal looping between ncRNA-a7 and *SNAI1* (Fig. 4b, middle panel). A similar result was obtained after depletion of MED1 or MED12 at the *AURKA* locus (Fig. 4c, middle panel). Importantly, depletion of ncRNA-a7 had a comparable effect, reducing the chromosomal looping between ncRNA-a7 and both *SNAI1* and *AURKA* loci (Fig. 4b, c, lower panels). Because depletion of ncRNA-a7 using siRNAs resulted in a partial decrease of ncRNA-a7 concentration (about 60% depletion, Supplementary Fig. 1), the residual ncRNA-a7 may be sufficient to promote smaller levels of DNA looping seen after ncRNA-a7 depletion.

We also performed a similar analysis measuring the extent of DNA looping between ncRNA-a3 and that of the *TAL1* locus. Similar to results obtained with ncRNA-a7, we detected a specific and robust DNA looping between ncRNA-a3 and *TAL1* (Fig. 4e). Finally, depletion of Mediator subunits MED1 and MED12 or ncRNA-a3 nearly abolished the DNA looping between ncRNA-a3 and *TAL1* (Fig. 4e, middle and lower panels). Notably, Mediator and ncRNA-a association with their target genes extend beyond the transcription start sites to include the more distal sites. This may reflect the engagement of proximal regulatory elements in ncRNA-a and Mediator chromatin looping (Fig. 4b, e). These results suggest a mechanism of action for ncRNA-a and Mediator that may involve additional chromatin contacts beyond that which has been observed for distal regulatory elements and transcriptional start sites.

We propose a new mechanism of action for a class of lncRNAs termed ncRNA-activating (ncRNA-a) involved in long-range transcriptional activation through their association with the Mediator complex (Supplementary Fig. 4). It is noteworthy that such Mediator–ncRNA-a-dependent chromatin loops extend beyond the transcription start sites of the target promoters, which may reflect the inclusion of promoter proximal regulatory elements in such associations. Future studies could also assess whether such ncRNA-a–Mediator chromatin loops augment, or function exclusively of, Mediator–Cohesin complexes[23].

Recently, experiments in zebrafish uncovered a fundamental role for lncRNAs in developmental control[24]. Moreover, recent reports have linked a number of lncRNAs to human diseases such as cancer and neurological disorders[11,25–27]. Importantly, our studies indicated that Mediator complexes containing disease-causing mutant MED12 proteins corresponding to FG syndrome fail to associate with ncRNA-a. These results underscore the significance of Mediator–ncRNA-a association and suggest the loss of such Mediator–ncRNA-a interactions as a possible contributing factor in such developmental disorders.

METHODS SUMMARY

Depletion experiments using siRNA/DNA transfection and ncRNA luciferase assays. pGL3-TK-ncRNA-a and pGL3-TK-control constructs were generated as described previously. On-TargetPlus siRNAs and non-targeting control siRNA were purchased from Dharmacon. The modified siDNA oligonucleotides were purchased from Integrated DNA Technologies. siRNA/DNA transfections were

performed with Lipofectamine RNAiMAX (Life Technologies) according to the manufacturer's instructions, using 50 nM siRNA/DNAs in a final volume of 2 ml of culture medium without antibiotics. Cells were collected 48–72 h after two rounds of transfection.

Chromatin immunoprecipitation analysis was performed using 1×10^7 to 3×10^7 A549 or MCF7 cells were collected for each independent immunoprecipitation. After chemically crosslinking, cells were fixed in 1% formaldehyde 1×PBS with 10% serum for 15 min at room temperature. After immunoprecipitation, bound chromatin was eluted at 65 °C for 10 min in 60 µl fresh elution buffer twice (TE, 1% SDS) and 80 µl of elution buffer was added to the final volume of 200 µl. Crosslinking samples were reversed overnight at 65 °C. After proteinase K treatment (0.5 mg ml^{-1}), DNA was purified using a QIAquick PCR purification kit (Qiagen).

Protein affinity purification and RNA immunoprecipitation. The affinity-purified mediator complex and the GFP protein control were generated from the Flag-tagged MED12 or GFP HEK293T stable cell lines. RNA probes, internally labelled with [α-^{32}P]CTP, were transcribed using T7 or Sp6 RNA polymerase (Riboprobe System, Promega). RNA immunoprecipitations were performed using Dynabeads Protein A (Invitrogen) according to the manufacturer's recommendations. Transient transfected HEK293T cells were UV-crosslinked at 254 nm (200 mJ cm^{-2}) in 10 ml ice-cold PBS and collected by scraping. Cells were incubated in lysis solution (0.1% SDS, 0.5% NP40, 0.5% sodium deoxycholate, 400 U ml^{-1} RNase inhibitor (Roche)), and protease inhibitor at 4 °C for 25 min with rotation, followed by DNase treatment (30 U of DNase, 15 min at 37 °C).

Full Methods and any associated references are available in the online version of the paper.

Received 29 March; accepted 31 December 2012.
Published online 17 February 2013.

1. Birney, E. *et al.* Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447,** 799–816 (2007).
2. Bertone, P. *et al.* Global identification of human transcribed sequences with genome tiling arrays. *Science* **306,** 2242–2246 (2004).
3. The FANTOM Consortium. The transcriptional landscape of the mammalian genome. *Science* **309,** 1559–1563 (2005).
4. Okazaki, Y. *et al.* Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* **420,** 563–573 (2002).
5. Ørom, U. A. *et al.* Long noncoding RNAs with enhancer-like function in human cells. *Cell* **143,** 46–58 (2010).
6. Guttman, M. *et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458,** 223–227 (2009).
7. Kapranov, P. *et al.* RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **316,** 1484–1488 (2007).
8. Wang, K. C. & Chang, H. Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **43,** 904–914 (2011).
9. Lee, J. T. The X as model for RNA's niche in epigenomic regulation. *Cold Spring Harb. Perspect. Biol.* **2,** a003749 (2010).
10. Rinn, J. L. *et al.* Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129,** 1311–1323 (2007).
11. Gupta, R. A. *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464,** 1071–1076 (2010).
12. Yap, K. L. *et al.* Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol. Cell* **38,** 662–674 (2010).
13. Zhao, J. *et al.* Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Mol. Cell* **40,** 939–953 (2010).
14. Kim, T. K. *et al.* Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465,** 182–187 (2010).
15. De Santa, F. *et al.* A large fraction of extragenic RNA pol II transcription sites overlap enhancers. *PLoS Biol.* **8,** e1000384 (2010).
16. Wang, K. C. *et al.* A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472,** 120–124 (2011).
17. Risheg, H. *et al.* A recurrent mutation in MED12 leading to R961W causes Opitz-Kaveggia syndrome. *Nature Genet.* **39,** 451–453 (2007).
18. Rump, P. *et al.* A novel mutation in MED12 causes FG syndrome (Opitz-Kaveggia syndrome). *Clin. Genet.* **79,** 183–188 (2011).
19. Malik, S. & Roeder, R. G. The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nature Rev. Genet.* **11,** 761–772 (2010).
20. Taatjes, D. J. The human Mediator complex: a versatile, genome-wide regulator of transcription. *Trends Biochem. Sci.* **35,** 315–322 (2010).
21. Knuesel, M. T., Meyer, K. D., Donner, A. J., Espinosa, J. M. & Taatjes, D. J. The human CDK8 subcomplex is a histone kinase that requires Med12 for activity and can function independently of mediator. *Mol. Cell. Biol.* **29,** 650–661 (2009).
22. Nowak, S. J. & Corces, V. G. Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends Genet.* **20,** 214–220 (2004).
23. Kagey, M. H. *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467,** 430–435 (2010).
24. Ulitsky, I., Shkumatava, A., Jan, C. H., Sive, H. & Bartel, D. P. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* **147,** 1537–1550 (2011).
25. Yu, W. *et al.* Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. *Nature* **451,** 202–206 (2008).
26. Sopher, B. L. *et al.* CTCF regulates ataxin-7 expression through promotion of a convergently transcribed, antisense noncoding RNA. *Neuron* **70,** 1071–1084 (2011).
27. Cabianca, D. S. *et al.* A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. *Cell* **149,** 819–831 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank Shiekhattar laboratory members for support and discussions. We thank W. Dang (Blobel laboratory) and I. Tempera for suggestions and technical help on 3C experiments. R.S. was partially supported by a cancer core grant from NIH (P30 CA 010815). D.J.T. was supported by a grant from the NCI (R01 CA127364). G.A.B. was supported by 5R37DK058044.

Author Contributions R.S., F.L. and U.A.O. conceived and designed the overall project with help from G.A.B., M.B., D.J.T. and M.C. F.L. and U.A.O. performed the RNAi screens; F.L. received advice from G.A.B. to execute the 3C experiments. F.L. and M.C. performed the ChIP experiments for Mediator. F.L. performed all experiments regarding the ncRNA association with Mediator and the kinase assays. F.L., M.C. and R.S. analysed the data and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.S. (shiekhattar@wistar.org).

METHODS

siRNA/DNA transfection and ncRNA luciferase assays. pGL3-TK-ncRNA-a and pGL3-TK-control constructs were generated as described previously. Firefly luciferase constructs, pRL-CMV and a selectable marker for puromycin resistance were co-transfected in HEK293T cells. Transfected cells were grown in the presence of 2.5 μg ml^{-1} puromycin (Sigma) for selection. Individual colonies were isolated and screened for luciferase expression. Luciferase assays were performed in 96-well white plates using Dual-Glo luciferase assay system (Promega) according to the manufacturer's protocol.

On-TargetPlus siRNAs and non-targeting control siRNA were purchased from Dharmacon. The modified siDNA oligonucleotides were purchased from Integrated DNA Technologies. siRNA or siDNA transfections were performed with Lipofectamine RNAiMAX (Life Technologies) according to the manufacturer's instructions, using 50 nM siRNA/DNAs in a final volume of 2 ml of culture medium without antibiotics. Cells were collected 48–72 h after two rounds of transfection.

Corresponding ncRNA-a HEK293T cell lines growing in 12-well dishes were transfected with corresponding siRNA oligonucleotides, using Metafectene Pro (Biontex Laboratories). Twenty-four hours after transfection, the cells were lysed in 200 μl passive lysis buffer (Promega), and firefly and Renilla luciferase activities were measured on 30 μl lysate using the Dual-Luciferase reporter assay kit (Promega). Individual experiments were realized in triplicate, and the ratios presented are the averages of three independent experiments.

siRNA sequences. siControl1: ON-TARGET non-targeting siRNA 1 (D-001810-01); siControl2: ON-TARGET Non-targeting siRNA 2 (D-001810-02); siGFP: GFP Duplex I (P-002048-01); sincRNA-a3-1: 5′-CUAUGGAAUUUAAGCCCAAUU-3′; sincRNA-a3-2: 5′-GCAAGAUCACGGUGCGUCAUU-3′; sincRNA-a7-1: 5′-CCGAUUUGAGAGAGUGAGAUU-3′; sincRNA-a7-2: 5′-GAAGGGAACCAGUGCUAAAUU-3′; siMED12: 5′-UCACUCAUCUCAUGUUAUAUU-3′; siNIPBL: 5′-GGGAAUAUGAAGAGCGUGAUU-3′; siCCNT1: 5′-GAACAAACGUCCUGGUGAUUU-3′; siCDK9: 5′-UGACGUCCAUGUUCGAGUAUU-3′; siCEBPA: 5′-ACAAGAACAGCAACGAGUAUU-3′; siEP300: 5′-GGACUACCCUAUCAAGUAAUU-3′; siGTF2B: 5′-ACAAUCAGACAGUCCUAUAUU-3′; siSMC1L: 5′-CAUCAAAGCUCGUAACUUCUU-3′.

siDNA sequences. siGFP: 5′-T*C*A*C*C*T*T*C*A*C*C*C*T*C*T*C*C*A*C*T-3′; sincRNA-a1: 5′-C*C*T*G*C*C*C*T*C*T*T*T*C*C*T*C*C*T*G*T*A-3′; sincRNA-a3: 5′-G*C*T*C*G*T*C*G*G*G*T*A*C*T*C*T*C*C*A*T*T-3′; sincRNA-a7-1: 5′-G*T*T*C*C*C*T*T*C*T*G*A*C*G*C*C*T*G*G*T*A-3′; sincRNA-a7-2: 5′-C*A*T*C*T*T*C*T*C*C*G*T*C*T*C*T*G*C*C*A*C-3′. Asterisks indicate nucleoside phosphorothioates.

shRNA infections. pLKO.1 lentiviral vector containing short hairpin RNAs (shRNAs) or shRNA targeting GFP control was obtained from Open Biosystems. The TRC shRNA constructs were purchased from Open Biosystems using Lentiviral packaging system. Twenty-four hours after infection, selection with 2.5 μg ml^{-1} puromycin was performed for 48 h.

shRNA list. shGFP: RHS4459; MED12 (TRCN0000018576): 5′-CGGGTACTTCATACTTTGG-3′; MED17 (TRCN0000019245): 5′-CCGAGCTTGCAGTTATCTATT-3′; MED18 (TRCN0000053164): 5′-GCCAGAAATGGGAGACAAGAA-3′; MED1 (TRCN0000019800): 5′-GCCGAGTTCCTCTTATCCTAA-3′; MED21 (TRCN0000013429): 5′-CCATTGGAGTATTGCAGCAAT-3′; MED15 (TRCN0000018969): 5′-GCTCAGAACCAACCATCAC-3′; MED16 (TRCN0000022139): 5′-CGACATTGACAAGGTCATGAT-3′; MED23 (TRCN0000019196): 5′-CGCAGTTTACACGCTTCCT-3′; WDR5 (TRCN0000118050): 5′-CCAACCTTATTGTCTCAGGAT-3′.

Antibodies and recombinant proteins. Antibodies and recombinant proteins used were: Pol II (SC-899, Santa Cruz), MED1 (A300-793A, Bethyl labs), MED12 (A300-774A, Bethyl labs), histone H3Ser10p (05-598, Millipore), CDK8 (SC-1521, Santa Cruz), CCNC (cyclin C, SC-1061, Santa Cruz), human rHistone H3.1 (M2503S, NEB), and human rCCNH (pka-364, Prospec protein specialists).

RNA purification, cDNA synthesis and quantitative PCR. Cells were collected and re-suspended in TRIzol (Invitrogen) and RNA extracted according to the manufacturer's protocol. cDNA synthesis was done using random primers with Fermentas RevertAid first-strand cDNA synthesis kit. Quantitative PCR was done using IQ SYBR green super-mixes and CFX96 Touch Real-Time PCR detection system (Bio-rad). For all quantitative PCR reactions *GAPDH* was measured for an internal control and used to normalize the data. qPCR primer sequences are listed here: ncRNA-a1-qfor: 5′-GCAAGCGGAGACTTGTCTTT-3′; ncRNA-a1-qrev: 5′-GGCTGGTCTTGAACTCCTGA-3′; ncRNA-a3-qfor: 5′-TTAAGCCCAAGGAATGGAGA-3′; ncRNA-a3-qrev: 5′-AGCGGTGTGGAATAAACTGG-3′; ncRNA-a7-q1for: 5′-ATCCTGGTGGAAAAGGCATC-3′; ncRNA-a7-q1rev: 5′-GCCTGGGAAAAGCTACTTCA-3′; ncRNA-a7-q2for: 5′-CCGTTGGCTCCACAAACCT-3′; ncRNA-a7-q2rev: 5′-CAGTGACAGTAGCAGGCATCCT-3′; MED17-qfor: 5′-CCCAGCGGATAGACTTCAGC-3′; MED17-qrev: 5′-ATTGTTCCTCACTGAGTCCCA-3′; MED18-qfor: 5′-GGGCACCATTAACATGATGGA-3′; MED18-qrev: 5′-TGGTCAAGGAAAGTCTCAGGT-3′; MED1-qfor: 5′-CTGGAACGGCTCCATGCAA-3′; MED1-qrev: 5′-CTTCTCCATGACTTGACGCAC-3′; MED21-qfor: 5′-TTGCACGAACAGCAAAAGACA-3′; MED21-qrev: 5′-TGCGCTTTGTATCTTCTCCAG-3′; MED15-qfor: 5′-ATGGACGTTTCCGGGCAAG-3′; MED15-qrev: 5′-GCATCCTCGATTTGACTGACCA-3′; MED16-qfor: 5′-TGCTGGACATGAACACACTG-3′; MED16-qrev: 5′-AGGGAACCCTGGTTGGGTA-3′; MED23-qfor: 5′-GGGCCTTCAGACAGTTTTGG-3′; MED23-qrev: 5′-AGCTGTGTTCTTTCCCACTCA-3′; MED12-qfor: 5′-CTCCCGATGTTTACCCTCAG-3′; MED12-qrev: 5′-ATGCTCATCCCCAGAGACAG-3′; NIPBL-qfor: 5′-CGGCCCGGATTATAGTCTCT-3′; NIPBL-qrev: 5′-TCAGGTGCCAGCTGTCATTA-3′; CCNT1-qfor: 5′-GCCAATCTGCTTCAGGACA-3′; CCNT1-qrev: 5′-GGGAACTGTGTGAAGGACTGA-3′; CDK9-qfor: 5′-ATACGAGAAGCTCGCCAAGA-3′; CDK9-qrev: 5′-CCTTCTCGTTTTCCATCAGC-3′; CEBPA-qfor: 5′-CAGAGGGACCGGAGTTATGA-3′; CEBPA-qrev: 5′-TTCACATTGCACAAGGCACT-3′; EP300-qfor: 5′-GGGAGGACAAACAGGATTGA-3′; EP300-qrev: 5′-CCAATCTGCTGTCCAGGATT-3′; GTF2B-qfor: 5′-TTCCTGCTTTCGGTGTGTCT-3′; GTF2B-qrev: 5′-TGGATGGTTTGGACATGTGA-3′; SMC1L1-qfor: 5′-TCGGACCATTTCAGAGGTTC-3′; SMC1L1-qrev: 5′-GGTCTTTACCCGCAGGTTG-3′; WDR5-qfor: 5′-AATTCAGCCCGAATGGAGAGT-3′; WDR5-qrev: 5′-GGATATTCCCAGCTTGTGACC-3′; CSTF1-qfor: 5′-ACAGAACCAAAGTGGGCTTGA-3′; CSTF1-qrev: 5′-ACACAGACTGAGGCTTGATTTC-3′; AURKA-qfor: 5′-TTCAGGACCTGTTAAGGCTACA-3′; AURKA-qrev: 5′-ATTTGAAGGACACAAGACCCG-3′; TAL1-qfor: 5′-CAGCCTAGTGGCTTGTCCTC-3′; TAL1-qrev: 5′-GGAGCCTGAAATTGAATGGA-3′; HOTAIR-qfor: 5′-GGTAGAAAAAGCAACCACGAAGC-3′; HOTAIR-qrev: 5′-ACATAAACCTCTGTCTGTGAGTGCC-3′; HOTTIP-qfor: 5′-CCTAAAGCCACGCTTCTTTG-3′; HOTTIP-qrev: TGCAGGCTGGAGATCCTACT-3′; STIL-qfor: 5′-CCCAACGCCAACTGGAGATTT-3′; STIL-qrev: 5′-AGTCGGATGGTCTTCTCAGTC-3′; SNAI1-qfor: 5′-TCGGAAGCCTAACTACAGCGA-3′; SNAI1-qrev: 5′-AGATGAGCATTGGCAGCGAG-3′; UBE2V1-qfor: 5′-AAGCAAGAGCGACGCAAGAT-3′; UBE2V1-qrev: 5′-CGGAAATTGCGAGGGACTT-3′.

Chromatin immunoprecipitation. Briefly, 1×10^7 to 3×10^7 A549 or MCF7 cells were collected for each independent immunoprecipitation. After chemically crosslinking, cells were fixed in 1% formaldehyde, 1× PBS with 10% serum for 15 min at room temperature. After quenching with 0.25 M glycine, rinsing twice with cold 1× PBS, cells were lysed for 15 min in cold Chip buffer (150 mM NaCl, 1% Triton X-100, 5 mM EDTA; 10 mM Tris-HCl, pH 7.5, 0.5 mM DTT) with protease inhibitors on ice for 10 min. DNA was sonicated to between 200- and 350-bp DNA fragments on a Diagenode Bioruptor according to manufacturer's protocol. The solubilized chromatin was cleared and diluted with lysis buffer (without SDS to bring down SDS concentration to 0.1–0.15%). 1 ml of the chromatin solution was incubated overnight with 2–6 μg of antibody using protein A Dyna Beads at 4 °C. The beads were washed twice with Mixed Micelle wash buffer (0.2% SDS, 1% Triton X-100, 5.2% w/v sucrose, 5 mM EDTA, 20 mM Tris-HCl, pH 8, 150 mM NaCl), 2 times with high-salt buffer (1% Triton X-100, 1 mM EDTA, 50 mM HEPES, pH 8, 0.1% w/v deoxycholate, 500 mM NaCl), 2 times with LiCl buffer (250 mM LiCl, 0.5% NP-40, 0.5% deoxycholate, 1 mM EDTA, 10 mM Tris-HCl, pH 8), and one time with cold TE buffer (10 mM Tris-HCl, pH 7.6, 1 mM EDTA). Bound chromatin was eluted at 65 °C for 10 min in 60 μl fresh elution buffer twice (TE, 1% SDS) and brought to 200 μl. Crosslinking samples were reversed overnight at 65 °C. After proteinase K treatment (0.5 mg ml^{-1}), DNA was purified using a QIAquick PCR purification kit (Qiagen). ChIP primers sequences are listed here: AURKAChip-for: 5′-CGAATCCTGCCCAATCTACC-3′; AURKAChip-rev: 5′-GATGGCGAGAAAAGCAAGAG-3′; ncRNAa7Chip-for: 5′-GGATAACCAAAAGCGTAGGAAA-3′; ncRNAa7Chip-rev: 5′-CGGTCTCTAACAACAACAGCTC-3′; CSTF1-Chip-for: 5′-TCCATTTTTCCAGGAGAGAGC-3′; CSTF1-Chip-rev: 5′-CCAGTCGTTTCTGTGGTTTTC-3′; TAL1-Chip-for: 5′-TCCTTCCCCCTTTTCCTTAC-3′; TAL1-Chip-rev: 5′-TGAACGCACTCTCACAATCC-3′; ncRNA-a3-Cfor: 5′-ATGAGAGCCTCGGAAGTTGA-3′; ncRNA-a3-Crev: 5′-CCTTGCATTCCACAGGAAGT-3′; STIL-Chip-for: 5′-AATGTTACCCACCAACCTTCC-3′; STIL-Chip-rev: 5′-CTACCCTGCAAACAGACCTCA-3′; UBE2V1-Chip-for: 5′-TGCTTTTGGGGAGAATGAAG-3′; UBE2V1-Chip-rev: 5′-TCCTCGAACAGGAGACTGGT-3′.

Protein affinity purification and RNA immunoprecipitation. The affinity-purified Mediator complex and the GFP protein control were generated from the Flag-tagged MED12 or GFP HEK293T stable cell lines.

RNA probes, internally labelled with [α-^{32}P]CTP, were transcribed using T7 or Sp6 RNA polymerase (Riboprobe System, Promega).

RNA immunoprecipitations were performed using Dynabeads Protein A (Invitrogen) according to manufacturer's recommendations. 2 μg antibody was incubated with each sample, and the purified protein complex or the control

protein together with labelled RNA probes were incubated with the beads at 4 °C. The immunoprecipitated RNAs were extracted by TRIzol (Invitrogen) and resolved on the 6% Novex TBE-Urea gels (Invitrogen).

For ultraviolet crosslinking experiments, the method is modified from that of ref. 28. Transient transfected HEK293T cells were UV-crosslinked at 254 nm (200 mJ cm^{-2}) in 10 ml ice-cold PBS and collected by scraping. Cells were incubated in lysis solution (0.1% SDS, 0.5% NP40, 0.5% sodium deoxycholate, 400 U ml^{-1} RNase Inhibitor (Roche)) and protease inhibitor at 4 °C for 25 min with rotation, followed by DNase treatment (30 U of DNase, 15 min at 37 °C).

***In vitro* transcription of ncRNA-a.** *In vitro* transcription followed the protocol provided by the Riboprobe *in vitro* transcription systems (Promega); the designed oligonucleotides are listed here: ncRNA-a1-T7FOR: 5′-TAATACGACTCACTATAGGGgcgcatctcctacggcctccag-3′; ncRNA-a1-sp6REV: 5′-ATTTAGGTGACACTATAGAAccaagtgtcctgtgtaataggc-3′; ncRNA-a3-T7FOR: 5′-TAATACGACTCACTATAGGGgaagttgagcttcaggcgcggctctt-3′; ncRNA-a3-sp6REV: 5′-ATTTAGGTGACACTATAGAAAaccagcctcagcggtgtggaata-3′; ncRNA-a7-T7FOR: 5′-TAATACGACTCACTATAGGGctgtggcagagacggagaa-3′; ncRNA-a7-SP6REV: 5′-ATTTAGGTGACACTATAGAAacagccttagactgtgaaattttattg-3′; HOTAIR-T7FOR: 5′-TAATACGACTCACTATAGGGAGGCCCCAAAGAGTCTGATGTT-3′; HOTAIR-SP6REV: 5′-ATTTAGGTGACACTATAGAACTCAGGTTTTTCCAGCGTTCTC-3′. Uppercase letters inciate either the T7 or SP6 polymerase binding site, whereas lowercase letters indicate specific sequences for each ncRNA.

Kinase assays. Kinase reactions were carried out with 10–50 ng of purified Mediator protein complex and 500 ng human histone H3.1 recombinant (New England Biolabs) substrates in kinase buffer (25 mM Tris pH 8.0, 100 mM KCl, 100 μM ATP, 10 mM $MgCl_2$, 2 mM DTT) with the addition of 2.5 μCi [γ-^{32}P]ATP at 24 °C. Samples were resolved on a 4–12% SDS–polyacrylamide gel (Invitrogen) followed by transfer to PVDF membrane (Millipore) and autoradiography.

Chromosome conformation capture (3C). 3C assays followed the protocol from ref. 29, with minor modifications. A549 or MCF7 cells were filtered through a 70 μm filter to obtain a single-cell preparation. 1×10^7 cells were then fixed in 1% formaldehyde for 30 min at room temperature, followed by 15 min crosslinking. The reaction was quenched with 0.25 M glycine and cells were collected by centrifugation at 240*g* at 4 °C. The pellet was re-suspended in 0.5 ml cold lysis buffer (10 mM Tris-HCl, pH 7.5; 10 mM NaCl; 5 mM $MgCl_2$; 0.1 mM EGTA) with freshly added protease inhibitors (Roche) and was lysed on ice for 15 min. Cells were Dounce homogenized on ice with pestle B using 15–20 strokes. The nuclei were collected by centrifugation at 500*g* for 10 min at 4 °C and were re-suspended in 0.5 ml of 1.2 × restriction enzyme buffer (NEB), including 0.3% SDS and were incubated for 1 h at 37 °C while shaking at 1,200 r.p.m. Two per cent (final concentration) Triton X-100 was added to the nuclei and then the samples were incubated for 1 h at 37 °C while shaking. 400 U restriction enzyme was added to the nuclei and the samples were incubated at 37 °C overnight while shaking. 10 μl of the samples were collected before and after the enzyme reaction to evaluate digestion efficiency. The reaction was stopped by addition of 1.6% SDS (final concentration) and incubation at 65 °C for 30 min while shaking at 1,200 r.p.m. The sample was then diluted 10-fold with 1.15× ligation buffer (NEB) with added 1% Triton X-100 and was incubated for 1 h at 37 °C while shaking at 900 r.p.m. 100 U T4 DNA ligase (NEB) were added to the sample and the reaction was carried at 16 °C for 4 h followed by 30 min at room temperature. 300 μg of Proteinase K were added to the sample and the reaction was carried at 65 °C overnight for de-crosslink. RNA was removed by adding 300 μg of RNase and incubating the sample for 1 h at 37 °C. DNA was purified by twice phenol-chloroform extraction and ethanol precipitation. Purified DNA was then analysed by conventional or quantitative PCR. As a control for ligation products the BAC clones were digested with 10 U of restriction enzyme overnight and then incubated with 10 U T4 DNA-ligase at 16 °C overnight. The DNA was extracted by phenol-chloroform and precipitated with ethanol. Purified DNA was then analysed by conventional or quantitative PCR. For real-time PCR, the ΔCt method was applied for analysing data, using the BAC clone Ct values as control. Ct values were normalized for each primer pair by setting the Ct value of 100 ng of BAC-clone control random ligation matrix DNA at a value of 1. Primer sequences for PCR are listed here: Backg-Control-AURKA-for: 5′-CGTATGAGCTGGCAATAGCA-3′; Backg-Control-AURKA-rev: 5′-GGAGGCATGAGAGCTTGTTC-3′; Backg-Control-A7-for: 5′-ACTGGCTGAGGTTTGTGGAG-3′; Backg-Control-A7-rev: 5′-GGCCGATTTGAGAGAGTGAG-3′; Backg-Control-SNAI1-for: 5′-CAAGAGGGGAAGGAGAGGAG-3′; Backg-Control- SNAI1-rev: 5′-TCCTAAGTCCCCAGTCTCCA-3′; ncRNA-a7 anchor primer: 5′-CTAGGTCAGGCCCAAAGAGA-3′; ncRNA-a7 anchor control: 5′-ACAGAATAACCCATCCCTTTTCC-3′; A7-anchorDigTest-for: 5′-GCTTAACCCAGTGCAAGCTC-3′; A7-anchorDigTest-rev: 5′-AGTTTGGAGGCTGGGAATCT-3′; AURKA-N5DigTest-for: 5′-CCTGATTCTCACCCTCTTGG-3′; AURKA-N5DigTest-rev: 5′-GAGGCCCCATTCCTTAACAT-3′; SNAI1-N2DigTest-for: 5′-GGGAAAATGAGAGCAAACTCC-3′; SNAI1-N2DigTest-rev: 5′-AAGGCCAATGTGACAGAGATG-3′; A3-DigEcoRI-for: 5′-CACTTACAGGAACAGAGGACTCA-3′; A3-DigEcoRI-rev: 5′-CAACTTCTAGATCAGCGCCTTC-3′; 3C AURKA-BglII-A: 5′-GGGACATAAATGGGAGACTGG-3′; 3C AURKA-BglII-B: 5′-TGGGAGGAACTCTTCACCTCT-3′; 3C AURKA-BglII-C: 5′-TGGCTCTTTGGTCTTGTCTCT-3′; 3C AURKA-BglII-D: 5′-GTTTGGGCCGTTTCTTATTG-3′; 3C AURKA-BglII-E: 5′-CAGACCTGATTCTCACCCTCTTG-3′; 3C AURKA-BglII-F: 5′-CACCAACTTTTCTTCTCCACCAA-3′; 3C AURKA-BglII-G: 5′-CCACAGGCACTACTGTTACCAA-3′; 3C BglII.E-R1: 5′-CCAGGGAGGACACAGCATAA-3′; 3C BglII.E-R2: 5′-TGGTGTCTGGGTTGAAGTAACA-3′; 3C BglII.E-R3: 5′-TCGGCACCTGGTAAACTAGC-3′; ncRNA-a3 anchor primer: 5′-CAACTTCTAGATCAGCGCCTTC-3′; ncRNA-a3 anchor control: 5′-GTCTCTTTCTTTGATTGGAGTGG-3′; TAL1-1: 5′-GGGAATGGCAGAAGACAGAA-3′; TAL1-2: 5′-GTGTCGGGTTTGTTTTTACCTC-3′; TAL1-3: 5′-CTCCATTTTTGCCTCTGTCTCT-3′; TAL1-4: 5′-ATCTTTTCCGTGTAGCTCTTCCT-3′; TAL1-5: 5′-TGGGTTTGATTGCTCAGTACC-3′; TAL1-6: 5′-TCAATTTTGCCTGCCATACA-3′; TAL1-7: 5′-TTAAAAGGGCAGGGAGTTAAAAG-3′; TAL1-8: 5′-CAGATAGCCAGGAAGCAAGG-3′; SNAI1-1: 5′-GCTTTGGTCCTGAACTCCTG-3′; SNAI1-2: 5′-AACAAACTCACATTAATAAAATATCCAGA-3′; SNAI1-3: 5′-CTACAGTCGTGTGCCACCAA-3′; SNAI1-4: 5′-ATTGGCAAGAAGGGGAAAAT-3′; SNAI1-5: 5′-TCCAACATCCCAGACCTTTC-3′; SNAI1-6: 5′-GACCATGAGAAAGCCAGAGG-3′; SNAI1-7: 5′-AAGCTCCAGTGGCATGTGAT-3′; SNAI1-8: 5′-TGGAGTCCTGGTGTGTTCTG-3′. BAC clone ID: RP-11-664B18, RP-11-1067D15, RP11-742I1 (Empire Genomics).

28. Jeon, Y. & Lee, J. T. YY1 tethers Xist RNA to the inactive X nucleation center. *Cell* **146,** 119–133 (2011).
29. Hagège, H. *et al.* Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nature Protocols* **2,** 1722–1733 (2007).