Rio1 downregulates centromeric RNA levels to promote the timely assembly of structurally fit kinetochores.

Peter De Wulf (peter.dewulf@unitn.it)
University of Trento
Ksenia Smurova
University of Trento
Stefania Stancari
University of Trento
Carmela Irene
University of Trento
Giovanna Berto
University of Trento
Maria Iacovella
Eureopean Institute of Oncology
Giulia Perticari
University of Trento
Ilaria D'Ambrosio
University of Trento
Michela Damizia
University of Trento
Maria Giubettini
University of Trento
Réginald Philippe
University of Trento
Chiara Baggio
University of Trento
Elisabetta Callegaro
University of Trento
Andrea Casagranda
University of Trento
Alessandro Corsini
University of Trento
Vincenzo Polese

University of Trento
Anna Ricci
University of Trento
Erik Dassi
University of Trento https://orcid.org/0000-0003-4487-0449

Article

Keywords:

Posted Date: December 29th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1132116/v1

Rio1 downregulates centromeric RNA levels to promote the timely assembly of structurally fit kinetochores

Ksenia Smurova[1], Stefania Stancari[1,#], Carmela Irene[1,#], Giovanna Berto[1,#], Maria Giuseppina Iacovella[2], Giulia Perticari[1], Ilaria D'Ambrosio[1], Michela Damizia[1], Maria Giubettini[1], Réginald Philippe[1], Chiara Baggio[1], Elisabetta Callegaro[1], Andrea Casagranda[1], Alessandro Corsini[1], Vincenzo Gentile Polese[1], Anna Ricci[1], Erik Dassi[1], Peter De Wulf[1,*]

[1]Department of Cellular, Computational and Integrative Biology (CIBIO), University of Trento, Via Sommarive 9, 38123 Trento (TN), Italy

[2]Department of Experimental Oncology, European Institute of Oncology, Via Adamello 16, 20139 Milano (MI), Italy

#Contributed equally

*Corresponding author: peter.dewulf@unitn.it

Running title: Kinase Rio1 curtails centromeric RNA levels

Key words: Rio1, RioK1, kinetochore, centromere, pericentromere, Rat1

Kinetochores assemble on centromeres (CENs) via histone H3 variant CENP-A and low levels of CEN transcripts. RNA polymerase II (RNAPII) activity is restrained by the CEN histone code, while CEN RNA concentrations are reduced by the nuclear exosome. Using *S. cerevisiae*, we add kinase Rio1 to this scheme as it downregulates RNAPII, and promotes CEN RNA turnover via exoribonuclease Rat1. Transcription factor Cbf1, and the assembled kinetochore further restrain CEN transcription. CEN transcripts exist as long (up to 11,000nt) and short RNAs (119±40nt), which may underlie CEN identity and kinetochore recruitment. While also curtailed by Rio1, Rat1, and the exosome, periCEN RNAs (<200nt) accumulate at levels that are one order of magnitude higher than the CEN transcripts. Depleting Rio1 causes CEN and periCEN RNA buildup, kinetochore malformation, and chromosome loss. Depleting human orthologue RioK1 leads to CEN RNA accumulation and micronuclei formation, suggesting that Rio1/RioK1 activity at centromeres is conserved.

Low levels of CEN transcripts embed in centromeric chromatin to promote kinetochore recruitment. Indeed, kinetochore proteins including histone H3 variant CENP-A (Cse4 in *S. cerevisiae*), CENP-C (Mif2), and Aurora B (Ipl1) bind to both centromeric DNA and RNA[1-4]. As demonstrated with yeast and human cells, excessive or insufficient levels of CEN transcripts interfere with kinetochore recruitment[5-11]. Inflated CEN RNA concentrations characterize various malignancies[12-15] suggesting that centromere dysregulation may contribute to kinetochore malformation, resulting in chromosome mis-segregation and aneuploidy; a hallmark of cancer.

Low RNA polymerase II (RNAPII) activity at metazoan centromeres adheres to the histone code[10,16-18], which remains poorly understood in *S. cerevisiae*. Inclusion of histone H2A variant H2A.Z, normally associated with gene activation, represses CEN transcription in budding yeast[19,20]. CEN RNA levels are further kept low by the RNA-degrading nuclear exosome[21]. Regulatory factors that impinge on RNAPII to modulate its activity at centromeres remain unidentified in all eukaryotes.

To ensure kinetochore assembly on the core centromere only, fission yeast[22,23] and likely metazoans as well[9,24-26], intensively transcribe its flanking sequences. The generated transcripts are then processed into short interfering siRNAs that recruit histone-modifying enzymes and heterochromatin-binding proteins, which compact the local chromatin to insulate the centromere core. However, neither the transcription from the non-coding CEN-flanking pericentromeric domains, nor the RNAs produced from them have yet been studied in *S. cerevisiae*. As budding yeast did not evolve the siRNA pathway, it may use alternative routes to modulate centromere and pericentromere activity, which could well be conserved to higher species.

We previously reported that *S. cerevisiae* kinase Rio1 downregulates RNA polymerase I (RNAPI) at the rDNA (nucleolus), while also promoting *35S* pre-rRNA

processing via the SSU processome. Both activities clear the rDNA, allowing for its condensation and segregation[27]. Imaging and genetic interaction analyses placed Rio1 also at kinetochores/centromeres[28], suggesting it might also regulate CEN RNA production (by RNAPII) and processing.

At S-phase entry, yeast centromeric nucleosomes and kinetochores disassemble, allowing (i) the replisome to move through the centromere, (ii) RNAPII to transcribe the CEN DNA, (iii) centromeric R-loops to be resolved, and (iv) excess CEN RNAs to be degraded. Within five minutes, centromeric nucleosomes reform, CEN transcripts become embedded, and kinetochores reassemble[29]. Since Rio1 levels peak at centromeres in early-S-phase, it might participation in one or more of the above activities. Indeed, treating Rio1-depleted yeast with an RNAPII inhibitor established an involvement of the kinase in both the synthesis and turnover of CEN transcripts. Rio1 contributes to phosphorylation of the C-terminal domain of RNAPII subunit Rpb1, which dictates RNAPII activity. RNAPII downregulation is reinforced by repressive transcription factor Cbf1 and the assembled kinetochore itself. Rio1 lowers CEN RNA levels via 5'-3' exoribonuclease Rat1, which acts in parallel to the 3'-5' nuclear exosome. We further identify very long CEN transcripts of up to 11,000nts in length, a large pool of CEN RNAs with lengths between 200-5000nt, next to short (<200nt) CEN transcripts (119±40nt). Most of the latter (85%) overlap the centromeric CDEIII subsequence and its right-flanking domain onto which the kinetochore CBF3 core complexes assemble and via which the kinetochore (starting with the CCAN kinetochore protein network) is recruited. Ten percent of the CEN transcripts overlapping CDEIII also cover CDEII, onto which the CEN nucleosome binds. Five percent of the CEN transcripts overlap with CDEI and its left-flanking region, onto which CEN transcription factor Cbf1 and a second network of CCAN kinetochore proteins binds[30].

While Rio1, Rat1, and the nuclear exosome also downregulate pericentromere transcript levels, the concentrations of these RNAs exceed those deriving from the centromeres with at least one order of magnitude, depending on the chromosome. periCEN transcripts are similarly long (11,000nt) but mostly exist as short <200nt oligomers. Upon Rio1 depletion, RNAPII activity at centromeres and pericentromeres becomes derepressed, CEN and periCEN RNA levels rise, also due to reduced degradation, resulting in an untimely assembly of kinetochores that are also structurally faulty, which culminates in chromosome mis-segregation. The here-identified roles of Rio1 appear to be conserved as human cells depleted of orthologue RioK1 similarly suffered from elevated CEN RNA levels and micronuclei formation (representing mis-segregated and damaged chromosomes).

RESULTS

Characterization of budding yeast CEN and periCEN transcripts

In *S. cerevisiae*, CEN transcription and the transcripts themselves have been studied only minimally due to their extremely low concentrations[21], while the non-coding periCEN transcripts have not even been characterized yet. To examine both, we subjected two mid-exponential cultures of both wild-type yeast and a *trf4*Δ mutant, which is defective in nuclear exosome activity, to HiSeq RNA-Sequencing analysis. CEN and periCEN transcripts were distinguished, and those present in at least two of the four samples were mapped onto the yeast genome (Fig. 1a, Extended Data Fig. 1a). In wild-type yeast, CEN RNAs measured up to 11,000nt (Fig. 1b), much longer than hitherto reported (400-1,700nt[20,31]), and originated far upstream from the CEN sequence (confirmed by PCR analysis; transcript start and termination sites in Fig. 1a, Extended Data Fig. 1a). While CEN transcripts with lengths between 200-5,000nt were most abundant, a highly reproducible pool of short CEN RNAs (119±40nt) covering the centromere sequences (117±2nt) were identified (Fig. 1b, Extended

Data Fig. 1b), suggesting pre-CEN transcript processing and/or short CEN transcript production. Most (85%) of the short CEN RNAs overlapped the centromeric CDEIII sequence and its right-flanking region, onto which the most upstream kinetochore CBF3 complexes and the CCAN kinetochore network bind to recruit the additional downstream kinetochore subunits. Ten percent of the transcripts additionally harbor the centromeric CDEII sequence, which recruits and bends around the CEN nucleosome. Five percent of the CEN transcripts comprised CDEI and its left-flanking sequence, onto which CEN transcription factor Cbf1 and a second network of CCAN proteins binds[30] (Extended Data Fig. 1b). In general, the CEN transcripts occurred at a few molecules per cell, corresponding to 10^{-6} of *ACT1* mRNA levels measured (Extended Data Figs. 2a,3a). The periCEN transcripts were also very long (11,000nt, Fig. 1c) but existed mostly as heterogeneous <200nt transcripts (Fig. 1c), suggesting processing. Noteworthy, the periCEN transcript levels exceeded those of the CEN RNAs with at least one order of magnitude, depending on the chromosome (Figs. 1b,c; Extended Data Fig. 2).

Rio1 downregulates CEN and periCEN RNA levels peaking in early S-phase

We previously localized Rio1 to the rDNA (nucleolus) and centromeres[27,28]. By phosphorylating subunit Rpa43, Rio1 removes RNAPI from the rDNA. The kinase also promotes *35S* pre-rRNA cleavage by the SSU processome. Both interventions allow for rRNA condensation and segregation in late anaphase[27]. To now probe its activity at centromeres, we depleted Rio1 -carrying an auxin-inducible degron (Rio1-AID)- with 500μM auxin in 45 min (Extended Data Fig. 3b). Compared to mock treatment, a 5-, 9-and 22-fold increase in CEN8, CEN5 and CEN3 RNA levels was measured, respectively, in the auxin-treated cells. In parallel, periCEN8, periCEN5 and periCEN3 RNA concentrations rose 8-, 25-, and 8- to 50-fold, respectively (Fig. 2a, Extended Data Fig. 3c). Despite periCEN RNA levels being much

higher than the CEN transcript concentrations, their stronger relative increase upon Rio1 depletion implies that Rio1 more potently restrains the accumulation of periCEN RNAs.

Next, RT-qPCR analysis of yeast released from late G1 and tracked trough the cell cycle via budding state and spindle pole (Spc110-mCherry) distribution, revealed that CEN and periCEN RNA levels peak in early S-phase, when the CEN and periCEN regions replicate. The periCEN transcripts prevailed at much higher levels, as expected (Fig. 2b, Extended Data Fig. 3d). CEN and periCEN transcript concentrations then decreased through S-phase. In metaphase, CEN RNAs became undetectable while periCEN transcripts remained at low levels. At anaphase onset, the numbers for both started to rise, to then peak in S-phase (Fig. 2b). In yeast depleted of Rio1, which is characterized by a 15-minute S-phase delay (due to impeded rDNA replication[27]), the cell cycle profiles of the CEN and periCEN RNAs were maintained, albeit at levels that were 10-20–fold higher than those measured in the mock-treated cells (Fig. 2b). To probe whether the localization of Rio1 correlates with the above RNA profiles, we tracked 6Myc-Rio1 (anti-Myc immunofluorescence imaging) from late G1 through the cell cycle. Kinetochore protein Ndc80-3GFP (anti-GFP) acted as the reference (Fig. 2c). A dynamic Rio1 localization pattern emerged, matching those of the CEN and periCEN RNAs, suggesting a functional relationship with them. Since intracellular Rio1 protein levels remained stable (Fig. 2c), its dynamics likely reflected nuclear import and export, given that Rio1 accumulated in the nucleus of nuclear export mutant *crm1-1*[32].

To determine whether kinetochore occupancy restrains local RNAPII activity, we released wild-type yeast and the temperature-sensitive *ndc10-1* mutant from late G1 at 37°C, at which Ndc10-1 becomes inactive and unable to initiate kinetochore recruitment as a member of the CBF3 complex[33]. We observed the typical cell cycle profile for CEN RNAs albeit at concentrations that were 20-175–fold higher (depending on the centromere) than those measured in wild-type yeast (Fig. 2d). This finding suggests that bound kinetochores

curb CEN transcription, contributing to the decrease in CEN RNA levels after early S-phase (Fig. 2b).

Rio1, Cbf1, and RNase H regulate RNAPII activity at centromeres and pericentromeres

Centromere transcription factor Cbf1 was reported to activate[34] and repress[20,31] CEN transcription. Compared to wild-type yeast, we measured a 2.5-, 3- and 7.5-fold increase in CEN8, CEN3, and CEN5 RNA levels, respectively (Fig. 3a), in the *cbf1Δ* mutant, indicating that Cbf1 downregulates RNAPII activity. However, its impact was less than that of Rio1 since CEN RNA numbers increased 4-, 25-, and 10-fold, respectively, in Rio1-depleted yeast (Fig. 3a). Eliminating both (*cbf1Δ RIO1-AID* strain + 500μM auxin) triggered an additive increase in CEN RNA levels (Fig. 3a), suggesting that Cbf1 and Rio1 independently downregulate centromere transcript levels.

To determine whether Rio1 curbs RNAPII activity at centromeres and pericentromeres, we measured local RNAPII occupancy (anti-RNAPII ChIP followed by qPCR analysis) in Rio1-AID cells released from late G1 in the presence or absence of auxin. While RNAPII levels peaked in early S-phase, it was also present in G1 and metaphase (Fig. 3b), albeit at significantly lower numbers, suggesting reduced transcription activity at centromeres and pericentromeres in both cell cycle stages (Fig. 2b). Since local RNAPII occupancy did not significantly change in cells depleted of Rio1 (Fig. 3b), the latter seems involved in regulating RNAPII transcription activity. To examine this hypothesis, we probed the phosphorylation state of the C-terminal domain (CTD) of RNAPII subunit Rpb1, which dictates the RNAPII transcription cycle. The CTD comprises 26 repeats of heptamer Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7. Phosphorylated Ser5 (P-Ser5) and P-Ser7 underlie transcription initiation; P-Tyr1, P-Ser2, P-Ser5, and P-Ser7 promote elongation; whereas P-Ser2 and P-Thr4 mediate termination[35]. S-phase Rio1-AID cells containing or depleted of Rio1 were submitted to ChIP

using monoclonal antibodies recognizing each of the five phosphorylatable residues. qPCR analyses of the immunoprecipitated CEN/periCEN chromatin revealed a highly reproducible reduction (40%) in Thr4 and Ser7 phosphorylation in the auxin-treated cells (Fig. 3c). The drop in P-Thr4 and P-Ser7 levels associates local Rio1 activity with a downregulation of RNAPII initiation, elongation, and termination.

To further examine this hypothesis, we released Rio1-AID cells from late G1 in the presence or absence of auxin and RNAPII inhibitor thiolutin[36]. While CEN and periCEN transcript levels increased upon Rio1 depletion, treating the Rio1-depleted cells with 3μM thiolutin limited CEN and periCEN RNA accumulation (Fig. 4a), indicating that the elevated levels of CEN and periCEN transcripts in the absence of Rio1 were in part due to de-repressed RNAPII activity.

During transcription, R-loops are resolved by RNases H1 and H2 to prevent RNAPII from stalling. When both RNases were removed and Rio1 depleted (*rnh1*Δ *rnh2*Δ *RIO1-AID* strain + 500μM auxin), CEN and periCEN RNA levels dropped with 50-80% (Fig. 4b), attesting that R-loops curtail RNAPII activity at centromeres and pericentromeres[20,31,37].

Rio1 promotes CEN and periCEN RNA turnover

Since thiolutin treatment did not annihilate the accumulation of CEN and periCEN transcripts in the Rio1-depleted cells (Fig. 4a), Rio1 must also contribute to their turnover. The TRAMP (Trf4/Air2/Mtr4 Polyadenylation) complex together with the nuclear exosome degrade CEN RNAs in *S. cerevisiae*[21]. Indeed, depleting TRAMP activity (*TRF4-AID* strain + 500μM auxin) led to a significant increase in CEN and periCEN transcript levels (Fig. 4c, Extended Data Fig. 4). Depleting both Rio1 and Trf4 (*RIO1-AID TRF4-AID* strain + 500μM auxin) led to an additive accumulation of CEN RNAs, demonstrating that Rio1 and the exosome promote CEN transcript turnover in parallel (Fig. 4c, Extended Data Fig. 4). The periCEN transcripts

also increased when Rio1 was removed in Trf4-depleted cells, albeit to a lesser extent, indicating that while both act in parallel, additional turnover activities might intervene to prevent an excessive buildup of periCEN transcripts.

Exoribonuclease Rat1 trims cleaved pre-rRNA fragments[38] and degrades telomeric repeat-containing RNAs (TERRA)[39]. It also promotes transcription termination of RNAPI[40] and RNAPII[41,42], contributes to the initiation-elongation transition[43], mRNA splicing[44], and the decay of uncapped or splicing-defective mRNAs[44,45]. Strong genetic interactions between mutants in Rat1 and myriad kinetochore proteins[46] suggested an involvement of this RNase in CEN transcript turnover. Indeed, depleting Rat1 (*RAT1-AID* strain + 500µM auxin) triggered a strong increase in CEN RNA levels, similar to those measured in Rio1-depleted cells (Fig 4c, Extended Data Fig. 4). Removing both Rio1 and Rat1 (*RIO1-AID RAT1-AID* strain + 500µM auxin) did not reveal and additive effect on CEN transcript concentrations, indicating that both act in one pathway. Depleting all three proteins (*RIO1-AID TRF4-AID RAT1-AID* strain + 500µM auxin) confirmed CEN RNA turnover both by the Rio1-Rat1 axis, and the nuclear exosome (Fig. 4c, Extended Data Fig. 4).

As for the periCEN transcripts, they accumulated when Rat1 was removed, similar to depleting Rio1 or Trf4 (Fig. 4c, Extended Data Fig. 4). When these proteins were removed in double and triple combinations, we observed complex outcomes. Indeed, while the periCEN RNA levels increased, they never exceeded those in the Rio1-depleted cells (Fig. 4c). Possibly, the already high periCEN levels resulting from single-protein depletion could become pernicious when combined, triggering activities that limit periCEN RNA production and/or buildup.

Purifications of Rio1 (ProteinA-Rio1) from two independent yeast cultures followed by mass spectrometry analyses (Extended Data Table 3) revealed interacting proteins that are involved in DNA segregation (including CEN-binding kinetochore proteins Cep3, Ndc10,

and Ybp2), that regulate RNAPII activity, as well as enzymes that act on RNA (Fig. 5). Of note, Rio1 and Rat1 did not co-purify each other (Fig. 5, Extended Data Fig. 5, Extended Data Tables 3,4), suggesting that while they lie in the same pathway they could functionally interact through a(n) intermediate(s). Mass spectrometry identified 23 proteins that co-purified with both Rio1 and Rat1 (Extended Data Tables 3,4). Noteworthy members include Not3 (regulates gene transcription and transcript turnover), and Fun30 (chromatin remodeler involved in CEN silencing)[47].

Rio1 promotes timely and accurate kinetochore formation

To determine whether Rio1 depletion and consequent CEN RNA accumulation affect kinetochore assembly, we tracked kinetochore protein Ndc80-3GFP in Rio1-AID cells treated with 500μM auxin or a mock from late G1 to metaphase (Fig. 6a). In the absence of auxin, Ndc80-3GFP levels simultaneously dropped across all centromeres in early S-phase, representing timely kinetochore disassembly. Ndc80 signals next rose synchronously within five minutes, reflecting kinetochore re-formation[29]. In the Rio1-depleted cells, both Ndc80 dis- and re-assembly were untimely (asynchronous) and less explicit, suggesting aberrant kinetochore dynamics. Since Ndc80-3GFP levels were also lower in the Rio1-depleted cells (Fig. 6a) we examined kinetochore composition by quantitating the levels of Cse4-GFP-Cse4, CEN-binding proteins Mif2-GFP and Cnn1-3GFP, middle components Ame1-GFP and Mtw1-GFP, and outer protein Ndc80-3GFP in all cell cycle stages of Rio1-AID yeast treated with 500μM auxin or a mock (Fig. 6b, Extended Data Fig. 6a). We measured increased levels of Cse4 and Mif2, decreased levels of Cnn1 and Ndc80, while Ame1 and Mtw1 were not affected (Fig. 6b). We also observed a singular GFP dot, distant from the spindle-bound kinetochores, at which only Cse4 and Ame1 localized (Fig. 7a). The dot was present in all cell cycle stages, lied inside the nucleus, away from the nucleolus (Fig. 7b). Since the other

kinetochore reporter proteins did not localize at the dot, it likely represented a premature kinetochore formed on unbound chromosomes. Indeed, anti-GFP ChIP-qPCR analysis of Cse4 at CEN5 and CEN8, at *ACT1* (negative control), and at six top-scoring loci shown to recruit Cse4 when overexpressed[19,48] identified Cse4 only at centromeres in Rio1-AID cells treated with auxin or a mock (Extended Data Fig. 6b), indicating that the GFP dot represented CEN-bound Cse4, and Cse4-recruited Ame1. Of note, the cells did not arrest at the metaphase-anaphase transition, possibly because the spindle checkpoint requires structurally intact kinetochores to act on.

Given the aberrant kinetochore composition and unaligned Cse4-Ame1-CEN signal in Rio1-depleted yeast, we next examined chromosome segregation by tracking a centromere-harboring chromosome fragment (colony-sectoring assay[49], Extended Data Fig. 7). When *RIO1* and *RIO1-AID* cells were grown for 2 divisions (6h) under Rio1-lowering conditions (100μM auxin), and then transferred to agar-based medium lacking auxin, we observed an eight-fold increase in sectored colonies for the *RIO1-AID* strain (Fig. 7c), corroborating that Rio1 activity prevents chromosome loss.

RioK1 downregulates centromeric RNA levels in human cells

Rio1 promotes pre-rRNA processing[50] and small ribosomal subunit maturation[32,51-54] in yeast and human cells[51,55]. To examine whether its regulation of CEN RNA is similarly conserved, we inducibly removed RioK1 from human RPE-1 cells both by a quick (<3h) auxin-induced degradation of mAID-RioK1[56] (Extended Data Fig. 8) and a gradual knock-down (72h) using RNA interference (Extended Data Fig. 9). RT-qPCR analysis of the alphoid CEN sequence using two independent primer sets (one canonical, one CEN11-specific), consistently revealed 2.5-3-fold increase in CEN transcript levels using both depletion methods (Fig. 8a). While we also probed the pericentromeric transcripts, the redundant composition of the

periCEN regions prevented an unbiased identification. We further measured a significant increase in micronuclei in the siRNA_RioK1-depleted cells after three cell divisions (72h), as compared to the cells treated with negative-control siRNAs (Fig. 8b). Since micronuclei form following chromosome mis-segregation, our combined findings suggest that RioK1 also downregulates CEN activity in humans to prevent chromosomal instability.

Discussion

In budding yeast, dual-specificity protein kinase Rio1 (phosphorylates serine, tyrosine, and threonine *in vivo* and/or *in vitro*)[57] orchestrates a signaling network at the DNA, RNA, and protein levels to regulate growth and proliferation in response to nutrient availability, similar to kinases Tor1/mTor[28]. Within this context, we now identify Rio1 as a determinant of CEN and periCEN activity, kinetochore formation, and chromosome segregation, therewith adding a new layer to cell division control. The demand by kinetochores for extremely low levels of CEN transcripts establishes centromeres as highly negative environments for CEN RNA production. While Cbf1 represses RNAPII activity, and RNase H paces CEN transcription elongation by resolving R-loops, Rio1[27,57] reduces local RNAPII activity by phospho-regulating its C-terminal domain. The purification with Rio1 of general transcription factors, mediator subunits, and RNAPII elongation and termination factors suggests that its control of CEN RNA production may be extensive. The Rio1-RNAPII relationship might also act beyond centromeres and pericentromeres as depleting Rio1 alters the levels of various transcripts, none of which encode kinetochore subunits[28], underlining that the here-observed kinetochore phenotypes derive from aberrant kinetochore recruitment. Rio1 co-purifying Ndc10 and Cep3 (subunits of kinetochore's most upstream CBF3 complex), and CBF3 component Skp1, which regulates CBF3 formation and degradation[58,59], places Rio1 directly

at the CEN chromatin where it can monitor CEN RNA levels and orchestrate kinetochore (dis)assembly (Fig. 8c).

Rio1 further ensures low CEN RNA levels by promoting their turnover, independently from TRAMP/exosome activity[21]. Our epistatic analyses place Rio1 in one pathway with the 5'-3' exoribonuclease Rat1. Since Rat1 not only disposes of transcripts, but also incites RNAPII transcription termination[60], Rio1 might also exploit that ability of Rat1 to modulate RNAPII termination activity. Since Rio1 and Rat1 did not co-purify each other, they may collaborate indirectly. Indeed, we identified a number of proteins present in both the Rio1 and Rat1 interactomes (Extended Data Tables 3,4), some of which could functionally connect Rio1 and Rat1.

Though we identified CEN RNAs extending beyond the CEN core sequence, with lengths of up to 11,000nt; we also discriminated pools of shorter fragments (<200nt to 5,000nt). Especially intriguing are the 119±40nt transcripts that overlap the centromere, which we consistently identified in both wild-type yeast and, at higher concentrations, in yeast lacking TRAMP/nuclear exosome activity (*trf4Δ*) (Extended Data Fig. 1b). These transcripts are highly specific, and could be the result of explicit CEN processing or represent specific transcription products. Currently, no information exists as to transcription start sites upstream or within centromeres (there may be numerous, as in *S. pombe*[7]). Nevertheless, since most short CEN transcripts initiate downstream of CDEI; the binding site for transcription factor Cbf1, transcript-initiation positions could well reflect transcription start sites. Given their overlap with the kinetochore recruiting CEN elements, it is tempting to speculate that these short transcripts promote CEN identity, kinetochore assembly and stability.

We also found that the noncoding pericentromeres are actively transcribed. Unexpectedly, the levels of the produced transcripts exceeded those deriving from the enclosed centromeres with one or more orders of magnitude. Rio1 also downregulates their

production and turnover, while Rat1, TRAMP/exosome, and RNase H are similarly involved in regulating periCEN RNA levels. A common regulatory network thus acts at the CEN and periCEN domains, albeit with different intensities. Though present as long transcripts, pericentromeric RNAs accumulate largely as heterogeneous <200nt fragments. Do their exuberant levels help to insulate the centromere domain? While *S. cerevisiae*, in contrast to fission yeast, did not evolve the siRNA interference pathway, it could employ its host of RNA-processing enzymes to cleave, trim, and mature these transcripts. Similarly to *S. pombe*, the periCEN RNAs might recruit epigenetic enzymes to compact the CEN core flanking domains, allowing centromeres to stand out for kinetochore recruitment.

Analysis of our previous RNA MiSeq data obtained from three independent Rio1-AID cultures treated with 500µm auxin or a mock[28] did not allow us to identify CEN transcripts (due to lower sequencing depth) but confirmed a significant increase in periCEN RNA concentrations in the Rio1-depleted cells. Importantly, we did not detect longer lengths for the periCEN transcripts, indicating that Rio1 is not involved in periCEN RNA maturation.

In human cells, the CTD of RNAPII acting at centromeres is phosphorylated at Ser2 and Ser5 by still-unidentified kinases[5,61]. Since orthologue RioK1 downregulates CEN transcript levels, its activity may add to the above CTD-based modulation of RNAPII to complement the CEN histone code[62]. Mis-regulated CEN RNA levels interfere with CENP-A recruitment[6,7,10], and cause chromosomal instability[63]. The accumulation of micronuclei in RioK1-depleted cells are in line with this model. Anomalous CEN transcript levels characterize various malignancies[12-15], establishing them as potential biomarkers for breast cancer[64]. Recent work with human cells and mice provided evidence that a dysregulation of centromeres[12-14,65] and pericentromeres[66,67] contribute to tumorigenesis. While centromeric mis-expression predicts cancer patient survival and response to radiotherapy and chemotherapy[13], increased levels of pericentromeric satellite III transcripts in mice provoked

resistance towards chemotherapeutic agent etoposide, a phenotype that was rescued by lowering satIII expression[68].

Given that *RIOK1* expression is driven by oncogenic transcription factor c-Myc[57] it is not surprising that RioK1 levels are elevated in myriad malignancies. In fact, overexpressing only *RIOK1* drives lung and breast cancer formation and metastasis *in vivo*[69]. Enhancing RioK1 instability by altering its post-translational modifications impeded the progression of gastric and colorectal cancers[70]. While RioK1 is valued especially for its role in small ribosomal subunit maturation, we hope that its present identification as a guardian of CEN activity and chromosome segregation will help to accelerate research into this understudied oncogenic kinase.

Acknowledgements

We thank Dr. Carl Wu for the Cse4-GFP-Cse4 strain, Dr. Daniele Fachinetti for the hTERT-immortalized, OsTIR1-expressing RPE-1 cell line, Dr. Anna Cereseto for recombinant Cas9, and Dr. Luca Fava for his Cas9-based genome editing protocol. We thank Romina Belli and Daniele Peroni from CIBIO's Mass Spectrometry and Proteomics facility for performing the mass spectrometry analyses, Giorgina Scarduelli and Michela Roccuzzo from CIBIO's Imaging facility for assistance with microscopy, and Veronica De Sanctis and Roberto Bertorelli from CIBIO's Next Generation Sequencing (NGS) Core facility for performing the RNA-Sequencing analyses. P.D.W. acknowledges financial support from the Italian Association for Cancer Research Investigator Grant 13243, and Euregio Science Fund Grant IPN124. The European Regional Development Fund (ERDF) 2014-2020 POR P.A.TRENTO is acknowledged for supporting CIBIO's Mass Spectrometry and Proteomics facility, and its NGS Core facility.

Author contributions

K.S., S.S., and C.I., designed, performed and analyzed the yeast genetic, biochemical, molecular biological, cell cycle, CEN and periCEN expression, and imaging experiments, and were supported in these efforts by M.G.I, C.B., E.C., A.C., V.G.P and A.R. G.B., M.G., G.P., R.P., and I.D'A. designed, performed, and analyzed the human genetic, cell biological, molecular biological, and imaging experiments, and were assisted by A.C., and M.D. E.D. analyzed the RNA-Seq data. P.D.W. conceived, supervised, and analyzed the study. P.D.W. and K.S. wrote the paper. All of the authors approved of the manuscript.

Competing financial interests

The authors declare no competing financial interests.

Data availability

The RNA-Seq data sets are available at GSE189278 (Gene Expression Omnibus, NCBI).

FIGURE LEGENDS

Fig. 1 | Characterization of centromeric and pericentromeric transcripts. a, The CEN and periCEN RNAs derived from chromosome 5 are displayed as arrows (indicating transcript orientation, start and end points) that were mapped onto the *S. cerevisiae* genome. Total RNA (including the chromatin-embedded transcripts) was isolated from two mid-exponential cultures of both wild-type yeast and a *trf4*Δ mutant (deficient in nuclear exosome activity), converted into cDNA, and deep-sequenced (HiSeq RNA-Seq). The displayed CEN and periCEN RNAs were identified in at least two of the four samples. Transcript color intensity (grey scale) correlates with their abundance (number of reads). The yellow box details the CEN5 sequence and position of the oligomers used to quantitate the CEN5 and periCEN5 RNAs by RT-qPCR analysis. **b**, The centromeric transcripts (length and numbers of read pairs) identified in two independent, mid-exponential cultures of wild-type yeast are shown and color-coded for their chromosome of origin. The inset shows the transcript read pairs with a length <200nt. **c**, The pericentromeric transcripts (length and numbers of read pairs) identified in two independent, mid-exponential cultures of wild-type yeast are shown, and color-coded for their chromosome of origin. The periCEN transcript origins lie within 500bp left and right of the CEN sequence, and not overlap with it. The inset shows the periCEN transcript read pairs with a length <200nt. Of note, the scale for the periCEN RNA read numbers is in Log_2, compared to the linear scale for the number of CEN RNA read pairs shown in panel b.

Fig. 2 | Cell cycle stage-dependent regulation of CEN and periCEN RNA levels. a, The CEN and periCEN transcripts derived from chromosomes 5 and 8 measured in wild-type yeast (*RIO1*) and in a *RIO1-AID* strain treated with 500µM auxin or a mock. Their concentrations were normalized to those of *ACT1* mRNA (both measured by RT-qPCR), and then referenced

to the corresponding CEN and periCEN RNA levels in the wild-type strain (value = 1). The data were derived from three independent biological experiments. **b**, CEN5 and periCEN5 (downstream of CEN5), CEN8 and periCEN8 (upstream of CEN8) transcript numbers through a synchronous cell cycle (α-factor arrest-and-release from late G1; T = 0 min) in the *RIO1-AID* strain treated with 500μM auxin or a mock. Transcript levels were normalized to those of *ACT1* (all measured by RT-qPCR). The number of transcript molecules were derived from the standard curves established for each primer pair (Extended Data Fig. 3a). **c**, Upper panel: quantitative immunofluorescence analysis of 6Myc-Rio1 (anti-Myc) and Ndc80-3GFP (anti-GFP) levels in spread nuclei isolated from cells synchronously released into the cell cycle from late G1 (α-factor arrest-and-release). Lower panel: western hybridization blot showing 6Myc-Rio1 protein levels (anti-Myc) through the cell cycle (Pgk1 acts as the loading control). **d**, CEN5 and CEN8 transcript levels in wild-type yeast (*NDC10*) and in a temperature-sensitive *ndc10-1* mutant synchronously released into the cell cycle at 37°C from late G1 (α-factor arrest-and-release). The transcript levels were normalized to those of *ACT1* mRNA (both quantitated by RT-qPCR). The numbers of CEN RNA molecules were calculated from the standard curves established for each primer pair (Extended Data Fig. 3a). The error bars in panels a-d represent standard deviations. Confidence levels (*p*-values) were calculated with the unpaired, two-tailed student t-test.

Fig. 3 | Rio1-mediated downregulation of CEN transcript levels, and negative control - in parallel to Cbf1- of RNA polymerase II activity. a, CEN3, CEN5, and CEN8 transcript levels were measured in wild-type yeast (*RIO1*), in the *RIO1-AID* strain treated with 500μM auxin or a mock, in a *cbf1*Δ mutant, and in a *cbf1*Δ *RIO1-AID* strain treated with 500μM auxin or a mock. Transcript concentrations were normalized to those of *ACT1* (both quantitated by RT-qPCR) and then referenced to the corresponding CEN transcript levels measured in the

wild-type strain (value = 1). The data were gathered from three independent biological experiments. **b**, RNA polymerase II (RNAPII) occupancy of CEN5/periCEN5 and CEN8/periCEN8, measured by chromatin immunoprecipitation (ChIP, monoclonal anti-RNAPII) and qPCR analysis. The *RIO1-AID* strain was synchronously released from late G1 (α-factor arrest-and-release) in the presence of 500μM auxin or a mock. Cells were analyzed in G1, S-phase, and metaphase (M). The immunoprecipitated chromatin was quantitated as % of input, and normalized to the value measured in G1 (value = 1) under each condition (auxin, mock). The data were derived from three independent biological experiments. **c**, Phosphorylation-state analysis of the C-terminal domain (comprising 26 tandem repeats of Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7) of RNAPII subunit Rpb1 at CEN5/periCEN5 and CEN8/periCEN8. The *RIO1-AID* strain was released from late G1 in the presence of 500μM auxin or a mock, and sampled in S-phase. Monoclonal antibodies singularly recognizing the phosphorylated state of each phosphorylatable residue (P-Tyr1, P-Ser2, P-Thr4, P-Ser5, P-Ser7) were used to ChIP RNAPII. Its presence at CEN5/periCEN5 and CEN8/periCEN8 was measured by qPCR analysis, expressed as % of input, and normalized to the value measured in G1 (value = 1) under both experimental conditions (auxin, mock). Anti-RNAPII ChIP-qPCR analysis acted as the positive control for the presence of RNAPII. The data were gathered from minimally three (up to five) biological experiments. The error bars in panels a-c represent standard deviations. Confidence levels (*p*-values) were calculated with the unpaired, two-tailed student t-test.

Fig. 4 | Regulation of CEN and periCEN transcript levels by Rio1, RNase H, Rat1, and TRAMP/nuclear exosome. a, CEN8 and periCEN8 (upstream of CEN8) transcript levels measured by RT-qPCR analysis in wild-type yeast (*RIO1*) treated with 3μM thiolutin or a mock, and in the *RIO1-AID* strain treated with 500μM auxin and 3μM thiolutin. The cells

(three independent biological experiments) were synchronously released into the cell cycle from late G1 (α-factor arrest-and-release). Samples were taken in G1 (T = 0 min), S-phase, and metaphase. Transcript levels were quantitated by RT-qPCR analysis, and normalized to those measured in the untreated wild-type strain (*RIO1*) at T = 0 min. **b**, CEN5 and periCEN5 (downstream of CEN5), CEN8 and periCEN8 (upstream of CEN8) transcript levels in the *RIO1-AID* and *RIO1-AID rnh1*Δ *rnh2*Δ strains treated with 500μM auxin were quantitated and normalized to those of *ACT1* mRNA (all quantitated by RT-qPCR). The data were gathered from three independent biological experiments. **c**, CEN5, CEN8, periCEN5 (downstream of CEN5), and periCEN8 (upstream of CEN8) transcript levels were quantitated by RT-qPCR analysis in the indicated strains, which were treated with 500μM auxin. All measures were normalized to those of *ACT1* mRNA (quantitated by RT-qPCR analysis) and referenced to the values for the *RIO1* (wild-type) strain (value = 1). The data were gathered from three independent biological experiments. The striped-line bar color patterns represent the double and triple protein depletions, with each color in it representing one depleted protein. The error bars in the above panels represent standard deviations. Confidence levels (*p*-values) were calculated with the unpaired, two-tailed student t-test.

Fig. 5 | The Rio1 interactome. Rio1-protein interaction map derived from two independent ProteinA-Rio1 purifications followed by mass spectrometry analysis. Two negative control purifications were performed in parallel, and the proteins specifically co-purifying with Rio1 at a confidence level of $p \leq 0.05$ were withheld. The confidence level for each protein is indicated by color intensity. High-scoring co-purifying proteins are involved in various aspects of cell division (chromosome segregation, regulation of cell cycle progression, DNA replication and repair) and RNA biology (transcription and its regulation, chromatin remodeling, RNA modification and degradation) are shown. Casein kinase 2 is an established

activator of Rio1. The proteins were functionally clustered based on their Gene Ontology function[47]. All co-purifying proteins and the corresponding volcano enrichment plots are presented in Extended Data Table 3.

Fig. 6 | Rio1 promotes the timely formation of structurally accurate kinetochores. a, Real-time imaging-based quantification of Ndc80-3GFP fluorescence levels in a *RIO1-AID* strain that was released from late G1 (α-factor arrest-and-release) in the presence of 500μM auxin or a mock, and tracked through S-phase. The number of cells analyzed (n) per microscopy field at each time-point is indicated. The cells were imaged across 25 Z-planes, which were then vertically projected at maximum intensity, and the fluorescence signals measured. The plotted data were gathered from three independent biological experiments. A.U. = arbitrary units. **b**, Quantification of fluorescence levels of GFP-labeled kinetochore reporter proteins at spindle-bound kinetochores in *RIO1-AID* strains treated for 45 min with 500μM auxin or a mock. The number of cells analyzed (n) per cell cycle stage is indicated (G1 = G1, S = S-phase, M = metaphase, A = anaphase). The cells were imaged across different microscopy fields through 25 Z-planes, which were then vertically projected at maximum fluorescence intensity for signal measurement. A.U. = arbitrary units. The data were gathered from three independent biological experiments. The error bars in panels a and b represent standard deviations. Confidence levels (*p*-values) were calculated with the unpaired, two-tailed student t-test.

Fig. 7 | Unaligned ectopic localization of Cse4 and Ame1 in Rio1-depleted yeast suffering from chromosome mis-segregation a, Fluorescence levels of Cse4-GFP-Cse4 and Ame1-GFP in *RIO1-AID* strains with marked spindle poles (Spc110-mCherry) and spindle microtubules (Tub1-mCherry) treated for 45 min with 500μM auxin or a mock. The number

of cells analyzed (n) per cell cycle stage is indicated (G1 = G1 phase, S = S phase, M = metaphase, A = anaphase). The cells were imaged across different microscopy fields through 25 Z-planes, which were then vertically projected and measured for maximum fluorescence intensity. The right plots, which accompany the images, show the number of cells at different cell cycle stages that contain an unaligned ectopic GFP-Cse4-GFP or Ame1-GFP signal (indicated by the white arrow heads). The lower plot shows the ratio between the ectopic signal of Cse4-GFP-Cse4 and Ame1-GFP, and that of Cse4-GFP-Cse4 and Ame1-GFP at each set of clustered kinetochores in the Rio1-AID cells treated with 500μM auxin. **b**, Localization analysis of the unaligned ectopic Cse4-GFP-Cse4 signals (indicated by the white arrow heads) in *RIO1-AID* cells with marked spindle poles (Spc110-mCherry), genomic chromatin (Rap1-CFP), and nucleolus (rDNA array on chromosome 12; Nop1-mCherry), treated for 45 min with 500μM auxin. **c**, Measurement of chromosome fragment transmission fidelity in wild-type yeast (*RIO1*) and in a *RIO1-AID* strain treated with 100μM auxin for 6h (2 cell divisions) in 2% glucose synthetic medium. The cells were next grown on 2% glucose synthetic agar medium lacking auxin (5d). Colonies with red-colored sectors, indicative of chromosome reporter loss (Extended Data Fig. 7), were counted, and normalized to those measured in the wild-type *RIO1* strain (value = 1). The data were gathered from three independent biological experiments. The error bars in panels a and c represent standard deviations. Confidence levels (*p*-values) were calculated with the unpaired, two-tailed student t-test.

Fig. 8 | Accumulation of CEN RNA and micronuclei in RioK1-depleted human cells, and working model of Rio1 activity at yeast centromeres and pericentromeres. a, Left: CEN11 (D11Z1) transcript levels measured by RT-qPCR analysis in *mAID-RIOK1* RPE-1 cells treated for 3h with 500μM auxin or a mock. The measurements, gathered from five

independent biological experiments, were normalized to *GAPDH* mRNA levels and referenced to CEN11 RNA levels measured in the mock-treated cells. mAID-RioK1 levels (rabbit anti-RioK1 western hybridization blot) before and after 3h of treatment with 500µM auxin are shown. GAPDH acted as the loading control. Right: CEN alphoid transcript levels measured by RT-qPCR analysis in RPE-1 cells treated for 72h with anti-*RIOK1* or negative-control siRNA primers. The data, gathered from three independent biological experiments, were normalized to *GAPDH* mRNA concentrations. Accompanying RioK1 levels (mouse anti-RioK1 western hybridization blot) before and after 72h of siRNA_Ctrl and siRNA_RioK1 treatment. **b**, Left: number of RPE-1 cells (n) containing micronuclei following 72h treatment with anti-RioK1 or negative-control siRNA primers. Right: Interphase RPE-1 cells with stained chromosomes (Hoechst 33342) present in an asynchronous culture that was treated (72h) with siRNA_Ctrl or siRNA_RioK1 primers. The small DNA specimens represent micronuclei (mis-segregated and damaged chromosomes). The error bars in panels a and b represent standard deviations. Confidence levels (*p*-values) were calculated with the unpaired, two-tailed student t-test. **c**, Working model of how Rio1 ensures low CEN and periCEN RNA levels following CEN/periCEN replication in early S-phase. Rio1 downregulates RNAPII activity at centromeres and pericentromeres, and promotes CEN and periCEN RNA turnover by 5'-3' exoribonuclease Rat1, acting independently from 5'-3' TRAMP/nuclear exosome. Transcription factor Cbf1 represses RNAPII, while RNase H resolves local R-loops therewith modulating RNAPII processivity. Although Rio1 strongly downregulates periCEN transcript levels, the latter remain a magnitude higher than those deriving from centromeres. CEN and periCEN RNAs are produced as long pre-transcripts which become processed (indicated by scissors), including into 119±40nt oligomers that overlap the CEN sequence. Maturated periCEN RNAs exist as

heterogeneous oligomers <200nt. Rio1 is not involved in periCEN RNA processing, and only promotes CEN and periCEN transcript turnover.

REFERENCES

1. Quénet, D. & Dalal, Y. A long non-coding RNA is required for targeting centromeric protein A to the human centromere. *eLife* **3**(2014).
2. McNulty, S.M., Sullivan, L.L. & Sullivan, B.A. Human Centromeres Produce Chromosome-Specific and Array-Specific Alpha Satellite Transcripts that Are Complexed with CENP-A and CENP-C. *Dev Cell* **42**, 226-240 e6 (2017).
3. Du, Y., Topp, C.N. & Dawe, R.K. DNA Binding of Centromere Protein C (CENPC) Is Stabilized by Single-Stranded RNA. *PLOS Genetics* **6**, e1000835 (2010).
4. Silver, D.L., Jambhekar, A., Emerman, A.B., Schweidenback, C.T.H. & Blower, M.D. RNA Stimulates Aurora B Kinase Activity during Mitosis. *PLoS ONE* **9**, e100748 (2014).
5. Chan, F.L. *et al.* Active transcription and essential role of RNA polymerase II at the centromere during mitosis. *Proc Natl Acad Sci U S A* **109**, 1979-84 (2012).
6. Grenfell, A.W., Heald, R. & Strzelecka, M. Mitotic noncoding RNA processing promotes kinetochore and spindle assembly in Xenopus. *J Cell Biol* **214**, 133-41 (2016).
7. Catania, S., Pidoux, A.L. & Allshire, R.C. Sequence features and transcriptional stalling within centromere DNA promote establishment of CENP-A chromatin. *PLoS Genet* **11**, e1004986 (2015).
8. Hill, A. & Bloom, K. Genetic manipulation of centromere function. *Mol Cell Biol* **7**, 2397-405 (1987).
9. Saffery, R. *et al.* Transcription within a functional human centromere. *Mol Cell* **12**, 509-16 (2003).
10. Bergmann, J.H. *et al.* Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. *The EMBO Journal* **30**, 328-340 (2011).
11. Bergmann, J.H., Martins, N.M.C., Larionov, V., Masumoto, H. & Earnshaw, W.C. HACking the centromere chromatin code: insights from human artificial chromosomes. *Chromosome Research* **20**, 505-519 (2012).
12. Ting, D.T. *et al.* Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. *Science* **331**, 593-6 (2011).
13. Zhang, W. *et al.* Centromere and kinetochore gene misexpression predicts cancer patient survival and response to radiotherapy and chemotherapy. *Nat Commun* **7**, 12619 (2016).
14. Zhu, Q. *et al.* Heterochromatin-Encoded Satellite RNAs Induce Breast Cancer. *Mol Cell* **70**, 842-853 e7 (2018).
15. Bersani, F. *et al.* Pericentromeric satellite repeat expansions through RNA-derived DNA intermediates in cancer. *Proc Natl Acad Sci U S A* **112**, 15148-53 (2015).
16. Molina, O. *et al.* Epigenetic engineering reveals a balance between histone modifications and transcription in kinetochore maintenance. *Nat Commun* **7**, 13334 (2016).
17. Ishikura, S., Nakabayashi, K., Nagai, M., Tsunoda, T. & Shirasawa, S. ZFAT binds to centromeres to control noncoding RNA transcription through the KAT2B-H4K8ac-BRD4 axis. *Nucleic Acids Res* **48**, 10848-10866 (2020).
18. Zhu, J., Cheng, K.C.L. & Yuen, K.W.Y. Histone H3K9 and H4 Acetylations and Transcription Facilitate the Initial CENP-A(HCP-3) Deposition and De Novo Centromere Establishment in Caenorhabditis elegans Artificial Chromosomes. *Epigenetics Chromatin* **11**, 16 (2018).

19. Hildebrand, E.M. & Biggins, S. Regulation of Budding Yeast CENP-A levels Prevents Misincorporation at Promoter Nucleosomes and Transcriptional Defects. *PLoS Genet* **12**, e1005930 (2016).
20. Ling, Y.H. & Yuen, K.W.Y. Point centromere activity requires an optimal level of centromeric noncoding RNA. *Proc Natl Acad Sci U S A* **116**, 6270-6279 (2019).
21. Houseley, J., Kotovic, K., El Hage, A. & Tollervey, D. Trf4 targets ncRNAs from telomeric and rDNA spacer regions and functions in rDNA copy number control. *The EMBO journal* **26**, 4996-5006 (2007).
22. Volpe, T.A. *et al.* Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**, 1833-7 (2002).
23. Kagansky, A. *et al.* Synthetic Heterochromatin Bypasses RNAi and Centromeric Repeats to Establish Functional Centromeres. *Science* **324**, 1716-1719 (2009).
24. Fukagawa, T. *et al.* Dicer is essential for formation of the heterochromatin structure in vertebrate cells. *Nature Cell Biology* **6**, 784-791 (2004).
25. Kanellopoulou, C. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes & Development* **19**, 489-501 (2005).
26. Hsieh, C.-L. *et al.* WDHD1 modulates the post-transcriptional step of the centromeric silencing pathway. *Nucleic Acids Research* **39**, 4048-4062 (2011).
27. Iacovella, M.G. *et al.* Rio1 promotes rDNA stability and downregulates RNA polymerase I to ensure rDNA segregation. *Nature communications* **6**, 6643 (2015).
28. Iacovella, M.G. *et al.* Integrating Rio1 activities discloses its nutrient-activated network in Saccharomyces cerevisiae. *Nucleic Acids Res* **46**, 7586-7611 (2018).
29. Kitamura, E., Tanaka, K., Kitamura, Y. & Tanaka, T.U. Kinetochore microtubule interaction during S phase in Saccharomyces cerevisiae. *Genes Dev* **21**, 3319-30 (2007).
30. Guan, R. *et al.* Structural and dynamic mechanisms of CBF3-guided centromeric nucleosome formation. *Nature Communications* **12**(2021).
31. Chen, C.F., Pohl, T.J., Chan, A., Slocum, J.S. & Zakian, V.A. Saccharomyces cerevisiae Centromere RNA Is Negatively Regulated by Cbf1 and Its Unscheduled Synthesis Impacts CenH3 Binding. *Genetics* **213**, 465-479 (2019).
32. Vanrobays, E., Gelugne, J.P., Gleizes, P.E. & Caizergues-Ferrer, M. Late cytoplasmic maturation of the small ribosomal subunit requires RIO proteins in Saccharomyces cerevisiae. *Molecular and cellular biology* **23**, 2083-95 (2003).
33. Goh, P.Y. & Kilmartin, J.V. NDC10: a gene involved in chromosome segregation in Saccharomyces cerevisiae. *J Cell Biol* **121**, 503-12 (1993).
34. Ohkuni, K. & Kitagawa, K. Endogenous transcription at the centromere facilitates centromere activity in budding yeast. *Curr Biol* **21**, 1695-703 (2011).
35. Eick, D. & Geyer, M. The RNA polymerase II carboxy-terminal domain (CTD) code. *Chem Rev* **113**, 8456-90 (2013).
36. Trcek, T., Larson, D.R., Moldon, A., Query, C.C. & Singer, R.H. Single-molecule mRNA decay measurements reveal promoter- regulated mRNA stability in yeast. *Cell* **147**, 1484-97 (2011).
37. Mishra, P.K. *et al.* R-loops at centromeric chromatin contribute to defects in kinetochore integrity and chromosomal instability in budding yeast. *Molecular Biology of the Cell* **32**, 74-89 (2021).
38. Axt, K., French, S.L., Beyer, A.L. & Tollervey, D. Kinetic analysis demonstrates a requirement for the Rat1 exonuclease in cotranscriptional pre-rRNA cleavage. *PLoS One* **9**, e85703 (2014).

39. Luke, B. *et al.* The Rat1p 5' to 3' exonuclease degrades telomeric repeat-containing RNA and promotes telomere elongation in Saccharomyces cerevisiae. *Mol Cell* **32**, 465-77 (2008).
40. El Hage, A., Koper, M., Kufel, J. & Tollervey, D. Efficient termination of transcription by RNA polymerase I requires the 5' exonuclease Rat1 in yeast. *Genes Dev* **22**, 1069-81 (2008).
41. Kim, M. *et al.* The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* **432**, 517-22 (2004).
42. West, S., Gromak, N. & Proudfoot, N.J. Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature* **432**, 522-5 (2004).
43. Harlen, K.M. & Churchman, L.S. Subgenic Pol II interactomes identify region-specific transcription elongation regulators. *Mol Syst Biol* **13**, 900 (2017).
44. Dhoondia, Z. *et al.* A termination-independent role of Rat1 in cotranscriptional splicing. *Nucleic Acids Res* **49**, 5520-5536 (2021).
45. Jiao, X. *et al.* Identification of a quality-control mechanism for mRNA 5'-end capping. *Nature* **467**, 608-11 (2010).
46. Costanzo, M. *et al.* A global genetic interaction network maps a wiring diagram of cellular function. *Science* **353**(2016).
47. Cherry, J.M. *et al.* Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic acids research* **40**, D700-5 (2012).
48. Lefrancois, P., Auerbach, R.K., Yellman, C.M., Roeder, G.S. & Snyder, M. Centromere-like regions in the budding yeast genome. *PLoS Genet* **9**, e1003209 (2013).
49. Spencer, F., Gerring, S.L., Connelly, C. & Hieter, P. Mitotic chromosome transmission fidelity mutants in Saccharomyces cerevisiae. *Genetics* **124**, 237-49 (1990).
50. Vanrobays, E. *et al.* Processing of 20S pre-rRNA to 18S ribosomal RNA in yeast requires Rrp10p, an essential non-ribosomal cytoplasmic protein. *The EMBO journal* **20**, 4204-13 (2001).
51. Ferreira-Cerca, S., Kiburu, I., Thomson, E., LaRonde, N. & Hurt, E. Dominant Rio1 kinase/ATPase catalytic mutant induces trapping of late pre-40S biogenesis factors in 80S-like ribosomes. *Nucleic acids research* **42**, 8635-47 (2014).
52. Hector, R.D. *et al.* Snapshots of pre-rRNA structural flexibility reveal eukaryotic 40S assembly dynamics at nucleotide resolution. *Nucleic acids research* **42**, 12138-54 (2014).
53. Turowski, T.W. *et al.* Rio1 mediates ATP-dependent final maturation of 40S ribosomal subunits. *Nucleic acids research* **42**, 12189-99 (2014).
54. Belhabich-Baumas, K. *et al.* The Rio1p ATPase hinders premature entry into translation of late pre-40S pre-ribosomal particles. *Nucleic acids research* **45**, 10824-10836 (2017).
55. Widmann, B. *et al.* The kinase activity of human Rio1 is required for final steps of cytoplasmic maturation of 40S subunits. *Molecular biology of the cell* **23**, 22-35 (2012).
56. Natsume, T., Kiyomitsu, T., Saga, Y. & Kanemaki, M.T. Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. *Cell Rep* **15**, 210-218 (2016).
57. Berto, G., Ferreira-Cerca, S. & De Wulf, P. The Rio1 protein kinases/ATPases: conserved regulators of growth, division, and genomic stability. *Curr Genet* **65**, 457-466 (2019).

58. Leber, V., Nans, A. & Singleton, M.R. Structural basis for assembly of the CBF3 kinetochore complex. *The EMBO Journal* **37**, 269-281 (2017).
59. Rodrigo-Brenni, M.C., Thomas, S., Bouck, D.C. & Kaplan, K.B. Sgt1p and Skp1p Modulate the Assembly and Turnover of CBF3 Complexes Required for Proper Kinetochore Function. *Molecular Biology of the Cell* **15**, 3366-3378 (2004).
60. Park, J., Kang, M. & Kim, M. Unraveling the mechanistic features of RNA polymerase II termination by the 5′ -3′ exoribonuclease Rat1. *Nucleic Acids Research* **43**, 2625-2637 (2015).
61. Chan, F.L. & Wong, L.H. Transcription in the maintenance of centromere chromatin identity. *Nucleic Acids Res* **40**, 11178-88 (2012).
62. Smurova, K. & De Wulf, P. Centromere and Pericentromere Transcription: Roles and Regulation ... in Sickness and in Health. *Front Genet* **9**, 674 (2018).
63. Ichida, K. *et al.* Overexpression of satellite alpha transcripts leads to chromosomal instability via segregation errors at specific chromosomes. *Int J Oncol* (2018).
64. Kakizawa, N. *et al.* High relative levels of satellite alpha transcripts predict increased risk of bilateral breast cancer and multiple primary cancer in patients with breast cancer and lacking BRCArelated clinical features. *Oncol Rep* **42**, 857-865 (2019).
65. Kishikawa, T. *et al.* Satellite RNA Increases DNA Damage and Accelerates Tumor Formation in Mouse Models of Pancreatic Cancer. *Molecular Cancer Research* **16**, 1255-1262 (2018).
66. Miyata, K. *et al.* Pericentromeric noncoding RNA changes DNA binding of CTCF and inflammatory gene expression in senescence and cancer. *Proceedings of the National Academy of Sciences* **118**, e2025647118 (2021).
67. Nogalski, M.T. & Shenk, T. HSATII RNA is induced via a noncanonical ATM-regulated DNA damage response pathway and promotes tumor cell proliferation and movement. *Proc Natl Acad Sci U S A* **117**, 31891-31901 (2020).
68. Kanne, J. *et al.* Pericentromeric Satellite III transcripts induce etoposide resistance. *Cell Death & Disease* **12**(2021).
69. Weinberg, F. *et al.* The Atypical Kinase RIOK1 Promotes Tumor Growth and Invasive Behavior. *EBioMedicine* **20**, 79-97 (2017).
70. Hong, X. *et al.* Targeting posttranslational modifications of RIOK1 inhibits the progression of colorectal and gastric cancers. *Elife* **7**(2018).

Figures

Figure 1

Characterization of centromeric and pericentromeric transcripts. a, The CEN and periCEN RNAs derived from chromosome 5 are displayed as arrows (indicating transcript orientation, start and end points) that were mapped onto the *S. cerevisiae* genome. Total RNA (including the chromatin-embedded transcripts) was isolated from two mid-exponential cultures of both wild-type yeast and a *trf4Δ* mutant (deficient in nuclear exosome activity), converted into cDNA, and deep-sequenced (HiSeq RNA-Seq). The displayed CEN and periCEN RNAs were identified in at least two of the four samples. Transcript color intensity (grey scale) correlates with their abundance (number of reads). The yellow box details the CEN5 sequence and position of the oligomers used to quantitate the CEN5 and periCEN5 RNAs by RT-qPCR analysis. **b**, The centromeric transcripts (length and numbers of read pairs) identified in two independent, mid-exponential cultures of wild-type yeast are shown and color-coded for their chromosome of origin. The inset shows the transcript read pairs with a length <200nt. **c**, The pericentromeric transcripts (length and numbers of read pairs) identified in two independent, mid-exponential cultures of wild-type yeast are shown, and color-coded for their chromosome of origin. The periCEN transcript origins lie within 500bp left and right of the CEN sequence, and not overlap with it. The inset shows the periCEN transcript read pairs with a length <200nt. Of note, the scale for the periCEN RNA read numbers is in Log2, compared to the linear scale for the number of CEN RNA read pairs shown in panel b.

Figure 2

Cell cycle stage-dependent regulation of CEN and periCEN RNA levels. a, The CEN and periCEN transcripts derived from chromosomes 5 and 8 measured in wild-type yeast (*RIO1*) and in a *RIO1-AID* strain treated with 500μM auxin or a mock. Their concentrations were normalized to those of *ACT1* mRNA (both measured by RT-qPCR), and then referenced to the corresponding CEN and periCEN RNA levels in the wild-type strain (value = 1). The data were derived from three independent biological experiments. **b**, CEN5 and periCEN5 (downstream of CEN5), CEN8 and periCEN8 (upstream of CEN8) transcript numbers through a synchronous cell cycle (α-factor arrest-and-release from late G1; T = 0 min) in the *RIO1-AID* strain treated with 500μM auxin or a mock. Transcript levels were normalized to those of *ACT1* (all measured by RT-qPCR). The number of transcript molecules were derived from the standard curves established for each primer pair (Extended Data Fig. 3a). **c**, Upper panel: quantitative immunofluorescence analysis of 6Myc-Rio1 (anti-Myc) and Ndc80-3GFP (anti-GFP) levels in spread nuclei isolated from cells synchronously released into the cell cycle from late G1 (α-factor arrest-and-release). Lower panel: western hybridization blot showing 6Myc-Rio1 protein levels (anti-Myc) through the cell cycle (Pgk1 acts as the loading control). **d**, CEN5 and CEN8 transcript levels in wild-type yeast (*NDC10*) and in a temperature-sensitive *ndc10-1*

mutant synchronously released into the cell cycle at 37°C from late G1 (α-factor arrest-and-release). The transcript levels were normalized to those of *ACT1* mRNA (both quantitated by RT-qPCR). The numbers of CEN RNA molecules were calculated from the standard curves established for each primer pair (Extended Data Fig. 3a). The error bars in panels a-d represent standard deviations. Confidence levels (*p*-values) were calculated with the unpaired, two-tailed student t-test.

Figure 3

Rio1-mediated downregulation of CEN transcript levels, and negative control -in parallel to Cbf1- of RNA polymerase II activity. a, CEN3, CEN5, and CEN8 transcript levels were measured in wild-type yeast (*RIO1*), in the *RIO1-AID* strain treated with 500μM auxin or a mock, in a *cbf1*Δ mutant, and in a *cbf1Δ RIO1-AID* strain treated with 500μM auxin or a mock. Transcript concentrations were normalized to those of *ACT1* (both quantitated by RT-qPCR) and then referenced to the corresponding CEN transcript levels measured in the wild-type strain (value = 1). The data were gathered from three independent biological experiments. **b**, RNA polymerase II (RNAPII) occupancy of CEN5/periCEN5 and CEN8/periCEN8, measured by chromatin immunoprecipitation (ChIP, monoclonal anti-RNAPII) and qPCR analysis. The *RIO1-AID* strain was synchronously released from late G1 (α-factor arrest-and-release) in the presence of 500μM auxin or a mock. Cells were analyzed in G1, S-phase, and metaphase (M). The immunoprecipitated chromatin was quantitated as % of input, and normalized to the value measured in G1 (value = 1) under each condition (auxin, mock). The data were derived from three independent biological experiments. **c**, Phosphorylation-state analysis of the C-terminal domain (comprising 26 tandem repeats of Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7) of RNAPII subunit Rpb1 at CEN5/periCEN5 and CEN8/periCEN8. The *RIO1-AID* strain was released from late G1 in the presence of 500μM auxin or a mock, and sampled in S-phase. Monoclonal antibodies singularly recognizing the phosphorylated state of each phosphorylatable residue (P-Tyr1, P-Ser2, P-Thr4, P-Ser5, P-Ser7) were used to ChIP RNAPII. Its presence at CEN5/periCEN5 and CEN8/periCEN8 was measured by qPCR analysis, expressed as % of input, and normalized to the value measured in G1 (value = 1) under both experimental conditions (auxin, mock). Anti-RNAPII ChIP-qPCR analysis acted as the positive control for the presence of RNAPII. The data were gathered from minimally three (up to five) biological experiments. The error bars in panels a-c represent standard deviations. Confidence levels (*p*-values) were calculated with the unpaired, two-tailed student t-test.

Figure 4

Regulation of CEN and periCEN transcript levels by Rio1, RNase H, Rat1, and TRAMP/nuclear exosome. a, CEN8 and periCEN8 (upstream of CEN8) transcript levels measured by RT-qPCR analysis in wild-type yeast (*RIO1*) treated with 3μM thiolutin or a mock, and in the *RIO1-AID* strain treated with 500μM auxin and 3μM thiolutin. The cells (three independent biological experiments) were synchronously released into the cell cycle from late G1 (α-factor arrest-and-release). Samples were taken in G1 (T = 0 min), S-phase,

and metaphase. Transcript levels were quantitated by RT-qPCR analysis, and normalized to those measured in the untreated wild-type strain (*RIO1*) at T = 0 min. **b**, CEN5 and periCEN5 (downstream of CEN5), CEN8 and periCEN8 (upstream of CEN8) transcript levels in the *RIO1-AID* and *RIO1-AID rnh1Δ rnh2Δ* strains treated with 500μM auxin were quantitated and normalized to those of *ACT1* mRNA (all quantitated by RT-qPCR). The data were gathered from three independent biological experiments. **c**, CEN5, CEN8, periCEN5 (downstream of CEN5), and periCEN8 (upstream of CEN8) transcript levels were quantitated by RT-qPCR analysis in the indicated strains, which were treated with 500μM auxin. All measures were normalized to those of *ACT1* mRNA (quantitated by RT-qPCR analysis) and referenced to the values for the *RIO1* (wild-type) strain (value = 1). The data were gathered from three independent biological experiments. The striped-line bar color patterns represent the double and triple protein depletions, with each color in it representing one depleted protein. The error bars in the above panels represent standard deviations. Confidence levels (*p*-values) were calculated with the unpaired, two-tailed student t-test.

Figure 5

The Rio1 interactome. Rio1-protein interaction map derived from two independent ProteinA-Rio1 purifications followed by mass spectrometry analysis. Two negative control purifications were performed in parallel, and the proteins specifically co-purifying with Rio1 at a confidence level of $p \leq 0.05$ were withheld. The confidence level for each protein is indicated by color intensity. High-scoring co-purifying proteins are involved in various aspects of cell division (chromosome segregation, regulation of cell cycle progression, DNA replication and repair) and RNA biology (transcription and its regulation, chromatin remodeling, RNA modification and degradation) are shown. Casein kinase 2 is an established activator of Rio1. The proteins were functionally clustered based on their Gene Ontology function47. All co-purifying proteins and the corresponding volcano enrichment plots are presented in Extended Data Table 3.

Figure 6

Rio1 promotes the timely formation of structurally accurate kinetochores. a, Real-time imaging-based quantification of Ndc80-3GFP fluorescence levels in a *RIO1-AID* strain that was released from late G1 (α-factor arrest-and-release) in the presence of 500μM auxin or a mock, and tracked through S-phase. The number of cells analyzed (n) per microscopy field at each time-point is indicated. The cells were imaged across 25 Z-planes, which were then vertically projected at maximum intensity, and the fluorescence signals measured. The plotted data were gathered from three independent biological experiments. A.U. = arbitrary units. **b**, Quantification of fluorescence levels of GFP-labeled kinetochore reporter proteins at spindle-bound kinetochores in *RIO1-AID* strains treated for 45 min with 500μM auxin or a mock. The number of cells analyzed (n) per cell cycle stage is indicated (G1 = G1, S = S-phase, M = metaphase, A = anaphase). The cells were imaged across different microscopy fields through 25 Z-planes, which were

then vertically projected at maximum fluorescence intensity for signal measurement. A.U. = arbitrary units. The data were gathered from three independent biological experiments. The error bars in panels a and b represent standard deviations. Confidence levels (*p*-values) were calculated with the unpaired, two-tailed student t-test.

Figure 7

Unaligned ectopic localization of Cse4 and Ame1 in Rio1-depleted yeast suffering from chromosome mis-segregation a, Fluorescence levels of Cse4-GFP-Cse4 and Ame1-GFP in *RIO1-AID* strains with marked spindle poles (Spc110-mCherry) and spindle microtubules (Tub1-mCherry) treated for 45 min with 500μM auxin or a mock. The number of cells analyzed (n) per cell cycle stage is indicated (G1 = G1 phase, S = S phase, M = metaphase, A = anaphase). The cells were imaged across different microscopy fields through 25 Z-planes, which were then vertically projected and measured for maximum fluorescence intensity. The right plots, which accompany the images, show the number of cells at different cell cycle stages that contain an unaligned ectopic GFP-Cse4-GFP or Ame1-GFP signal (indicated by the white arrow heads). The lower plot shows the ratio between the ectopic signal of Cse4-GFP-Cse4 and Ame1-GFP, and that of Cse4-GFP-Cse4 and Ame1-GFP at each set of clustered kinetochores in the Rio1-AID cells treated with 500μM auxin. **b**, Localization analysis of the unaligned ectopic Cse4-GFP-Cse4 signals (indicated by the white arrow heads) in *RIO1-AID* cells with marked spindle poles (Spc110-mCherry), genomic chromatin (Rap1-CFP), and nucleolus (rDNA array on chromosome 12; Nop1-mCherry), treated for 45 min with 500μM auxin. **c**, Measurement of chromosome fragment transmission fidelity in wild-type yeast (*RIO1*) and in a *RIO1-AID* strain treated with 100μM auxin for 6h (2 cell divisions) in 2% glucose synthetic medium. The cells were next grown on 2% glucose synthetic agar medium lacking auxin (5d). Colonies with red-colored sectors, indicative of chromosome reporter loss (Extended Data Fig. 7), were counted, and normalized to those measured in the wild-type *RIO1* strain (value = 1). The data were gathered from three independent biological experiments. The error bars in panels a and c represent standard deviations. Confidence levels (*p*-values) were calculated with the unpaired, two-tailed student t-test.

Figure 8

Accumulation of CEN RNA and micronuclei in RioK1-depleted human cells, and working model of Rio1 activity at yeast centromeres and pericentromeres. a, Left: CEN11 (D11Z1) transcript levels measured by RT-qPCR analysis in *mAID-RIOK1* RPE-1 cells treated for 3h with 500μM auxin or a mock. The measurements, gathered from five independent biological experiments, were normalized to *GAPDH* mRNA levels and referenced to CEN11 RNA levels measured in the mock-treated cells. mAID-RioK1 levels (rabbit anti-RioK1 western hybridization blot) before and after 3h of treatment with 500μM auxin are shown. GAPDH acted as the loading control. Right: CEN alphoid transcript levels measured by RT-qPCR analysis in RPE-1 cells treated for 72h with anti-*RIOK1* or negative-control siRNA primers. The data, gathered from

three independent biological experiments, were normalized to *GAPDH* mRNA concentrations. Accompanying RioK1 levels (mouse anti-RioK1 western hybridization blot) before and after 72h of siRNA_Ctrl and siRNA_RioK1 treatment. **b**, Left: number of RPE-1 cells (n) containing micronuclei following 72h treatment with anti-RioK1 or negative-control siRNA primers. Right: Interphase RPE-1 cells with stained chromosomes (Hoechst 33342) present in an asynchronous culture that was treated (72h) with siRNA_Ctrl or siRNA_RioK1 primers. The small DNA specimens represent micronuclei (mis-segregated and damaged chromosomes). The error bars in panels a and b represent standard deviations. Confidence levels (*p*-values) were calculated with the unpaired, two-tailed student t-test. **c**, Working model of how Rio1 ensures low CEN and periCEN RNA levels following CEN/periCEN replication in early S-phase. Rio1 downregulates RNAPII activity at centromeres and pericentromeres, and promotes CEN and periCEN RNA turnover by 5'-3' exoribonuclease Rat1, acting independently from 5'-3' TRAMP/nuclear exosome. Transcription factor Cbf1 represses RNAPII, while RNase H resolves local R-loops therewith modulating RNAPII processivity. Although Rio1 strongly downregulates periCEN transcript levels, the latter remain a magnitude higher than those deriving from centromeres. CEN and periCEN RNAs are produced as long pre-transcripts which become processed (indicated by scissors), including into 119±40nt oligomers that overlap the CEN sequence. Maturated periCEN RNAs exist as heterogeneous oligomers <200nt. Rio1 is not involved in periCEN RNA processing, and only promotes CEN and periCEN transcript turnover.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ExtendedDataFiguresfcombined.pdf
- ExtendedDataFigureLegendsf.pdf
- ExtendedDataTable1f.pdf
- ExtendedDataTable2f.pdf
- ExtendedDataTable3f.xlsx
- ExtendedDataTable4f.xlsx