bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Functional profiling of long intergenic non-coding RNAs in fission yeast

María Rodríguez-López[1], Shajahan Anver[1*], Cristina Cotobal[1*], Stephan Kamrad[1,2,3*], Michal Malecki[1,5*], Clara Correia-Melo[2], Mimoza Hoti[1], StJohn Townsend[1,2], Samuel Marguerat[1,6], Sheng Kai Pong[1], Mary Y. Wu[4], Luis F. Montemayor[1], Michael Howell[4], Markus Ralser[2,3], Jürg Bähler[1+]

[1] University College London, Institute of Healthy Ageing and Department of Genetics, Evolution & Environment, London WC1E 6BT, United Kingdom

[2] The Francis Crick Institute, Molecular Biology of Metabolism Laboratory, London, United Kingdom

[3] Charité Universitätsmedizin Berlin, Institute of Biochemistry, Berlin, Germany

[4] The Francis Crick Institute, High Throughput Screening, London, United Kingdom

[5] Current address: Institute of Genetics and Biotechnology, Faculty of Biology, University of Warsaw, Warsaw, Poland

[6] Current address: MRC London Institute of Medical Sciences (LMS), London WC12 0NN, United Kingdom

* Equal contribution.

[+] Correspondence: j.bahler@ucl.ac.uk

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

Eukaryotic genomes express numerous long intergenic non-coding RNAs (lincRNAs) that do not overlap any coding genes. Some lincRNAs function in various aspects of gene regulation, but it is not clear in general to what extent lincRNAs contribute to the information flow from genotype to phenotype. To explore this question, we systematically analyzed cellular roles of lincRNAs in *Schizosaccharomyces pombe.* Using seamless CRISPR/Cas9-based genome editing, we deleted 141 lincRNA genes to broadly phenotype these mutants, together with 238 diverse coding-gene mutants for functional context. We applied high-throughput colony-based assays to determine mutant growth and viability in benign conditions and in response to 145 different nutrient, drug and stress conditions. These analyses uncovered phenotypes for 47.5% of the lincRNAs and 96% of the protein-coding genes. For 110 lincRNA mutants, we also performed high-throughput microscopy and flow-cytometry assays, linking 37% of these lincRNAs with cell-size and/or cell-cycle control. With all assays combined, we detected phenotypes for 84 (59.6%) of all lincRNA deletion mutants tested. For complementary functional inference, we analyzed colony growth of strains ectopically overexpressing 113 lincRNA genes under 47 different conditions. Of these overexpression strains, 102 (90.3%) showed altered growth under certain conditions. Clustering analyses provided further functional clues and relationships for some of the lincRNAs. These rich phenomics datasets associate lincRNA mutants with hundreds of phenotypes, indicating that most of the lincRNAs analyzed exert cellular functions in specific environmental or physiological contexts. This study provides groundwork to further dissect the roles of these lincRNAs in the relevant conditions.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Introduction

Genomes produce pervasive and diverse non-coding RNAs. How much genetic information is transacted by this non-coding ‘dark matter’ remains a matter of debate. A substantial, but poorly understood portion of transcriptomes consists of long intergenic non-coding RNAs (lincRNAs). LincRNAs are longer than 200 nucleotides, lack long open reading frames, and do not overlap any neighbouring coding regions. While not all lincRNAs may be functional, several have well-defined roles in gene regulation and some other cellular processes. Different lincRNAs can control gene expression at different levels, from transcription to translation, and either in *cis* (acting on neighbouring genes) or in *trans* (acting on distant genes) (Rinn and Chang, 2012; Popadin *et al.*, 2013; Ulitsky and Bartel, 2013; Yamashita, Shichino and Yamamoto, 2016; Schlackow *et al.*, 2017; Fauquenoy *et al.*, 2018). Although lincRNAs show little sequence conservation between species, functional principles seem to be conserved which can help us to understand their biology (Ulitsky, 2016). Specific lincRNAs have been implicated in complex human diseases (Batista and Chang, 2013)(Kumar *et al.*, 2013). For example, Xist exerts a tumour suppressive function (Yildirim *et al.*, 2013), TUNA is associated with neurological function and Huntington’s disease (Lin *et al.*, 2014), and lincRNA1 delays senescence (Abdelmohsen *et al.*, 2013). Moreover, lincRNAs are emerging as diagnostic molecular markers, as they can be easily detected in blood and could provide more readily accessible drug targets than proteins (Kim *et al.*, 2016; Bester *et al.*, 2018; DeWeerdt, 2019).

Despite these efforts and insights based on studying selected lincRNAs, the systematic picture remains incomplete as the importance of most lincRNAs is unknown. Functional analyses of lincRNAs are challenging given their profusion, poor annotation, low expression, and limited methodology (Bassett *et al.*, 2014; Cao, Wahlestedt and Kapranov, 2018; Kopp and Mendell, 2018). Knowledge of lincRNA function is therefore scarce even in well-studied organisms, highlighting the need for more systematic approaches. Large-scale genetic studies of lincRNAs have emerged, starting to provide a more global picture on their functions and contributions to phenotypes (Joung *et al.*, 2017; Liu *et al.*, 2017; Bester *et al.*, 2018; Parker *et al.*, 2018; Tuck *et al.*, 2018; Wei *et al.*, 2019; Balarezo-Cisneros *et al.*, 2021). These findings suggest that many lincRNAs play specialized roles in specific conditions and, therefore, need to be analysed in the relevant conditions.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The fission yeast, *Schizosaccharomyces pombe*, is a potent genetic model system to study gene regulation and lincRNA function *in vivo* (Marguerat *et al.*, 2012; Yamashita, Shichino and Yamamoto, 2016; Atkinson *et al.*, 2018; Fauquenoy *et al.*, 2018). Although only the most highly expressed lincRNAs show purifying selection (Jeffares *et al.*, 2015), their regulation is often affected by expression quantitative trait loci (Clément-Ziza *et al.*, 2014). Notably, transposon insertions in up to 80% of non-coding regions of the *S. pombe* genome can affect fitness (Grech *et al.*, 2019). RNA metabolism of fission yeast is similar to metazoan cells. For example, RNA interference (RNAi), RNA uridylation, and PABPN1-dependent RNA degradation are conserved from fission yeast to humans, but absent in budding yeast. Genome-wide approaches by us and others have uncovered widespread lincRNAs in fission yeast (Wilhelm *et al.*, 2008; Rhind *et al.*, 2011; Eser *et al.*, 2016; Atkinson *et al.*, 2018). Nearly all *S. pombe* lincRNAs are polyadenylated and transcribed by RNA polymerase II (Marguerat *et al.*, 2012). Transcription of lincRNAs starts from nucleosome-depleted regions upstream of positioned nucleosomes (Marguerat *et al.*, 2012; Atkinson *et al.*, 2018), and the regulation of some lincRNAs involves specific transcription factors such as Gaf1 (Rodríguez-López *et al.*, 2020). Most *S. pombe* lincRNAs are cryptic in cells growing under standard laboratory conditions, being suppressed by RNA-processing pathways such as the nuclear exosome, cytoplasmic exonuclease and/or RNAi (Zhou *et al.*, 2015; Atkinson *et al.*, 2018), but they become induced during starvation or sexual differentiation (Atkinson *et al.*, 2018). A substantial portion of lincRNAs are actively translated (Duncan and Mata, 2014), raising the possibility that some of them act as small proteins. A few *S. pombe* lincRNAs have been functionally analyzed: *meiRNA* and *rse1* control meiotic differentiation (Ding *et al.*, 2012; Yamashita, Shichino and Yamamoto, 2016; Fauquenoy *et al.*, 2018), *SPNCRNA.1164* regulates the *atf1* transcription-factor gene in *trans* during oxidative stress (Leong *et al.*, 2014), several lincRNAs activate the downstream *fbp1* gene during glucose starvation (Oda *et al.*, 2015), *prt* controls *pho1* expression (Ard, Tong and Allshire, 2014; Shah *et al.*, 2014), *nc-tgp1* inhibits the *tgp1* gene by transcriptional interference (Ard, Tong and Allshire, 2014), and *nam1* regulates meiotic differentiation (Touat-Todeschini *et al.*, 2017).

Most *S. pombe* lincRNAs may not function under benign laboratory conditions, when they are typically very lowly expressed (Marguerat *et al.*, 2012; Atkinson *et al.*, 2018). Phenomics approaches seek to rigorously characterize phenotypes associated with many gene variants under diverse conditions (Brochado and Typas, 2013; Rallis and Bähler, 2016). Such broad, high-throughput phenotyping is an effective approach to uncover functional clues for unknown genes. For example, while only 34% of all budding yeast gene-deletion mutants display a

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

growth phenotype under the standard condition, 97% of these mutants show sub-optimal growth in at least one condition when assayed under a large number of chemical or environmental perturbations (Hillenmeyer *et al.*, 2008). We have established a sensitive, reproducible platform for high-throughput colony-based assays to determine cellular fitness under diverse conditions (Kamrad, Rodríguez-López, *et al.*, 2020). Here we take advantage of this potent approach to broadly investigate phenotypes of 150 lincRNAs, using deletion and/or overexpression mutants, supplemented with high-throughput microscopy and flow-cytometry assays of deletion mutants. Colonies of a representative set of 238 coding-gene mutants were phenotyped in parallel for functional comparison. Using these different assays, we collected quantitative data for over 1.1 million unique colonies and over 5.7 million cells in a wide range of conditions. This study reveals hundreds of novel lincRNA-associated phenotypes and provides a framework for follow-on studies.

Results and Discussion

Experimental strategy for functional profiling of lincRNAs

We focused on lincRNAs, rather than other types of non-coding RNAs, because 1) they are poorly characterized in general but emerge as varied regulatory factors; 2) they can be deleted without directly interfering with coding gene function; and 3) they are more likely to function in *trans* as RNAs than antisense or promoter-associated ncRNAs which can affect neighbouring or overlapping genes via their transcription (Ard, Allshire and Marquardt, 2017). For functional profiling, we selected 150 *S. pombe* lincRNA genes that produce well-defined transcripts and are well-separated from neighboring coding regions (over ~200 bp), based on genome-browser views of RNA-seq data. We established efficient high-throughput methods to genetically manipulate these lincRNAs. For deletions, we applied a CRISPR/Cas9-based method (Rodríguez-López *et al.*, 2016); this approach allowed us to knock-out the precise regions transcribed into lincRNAs without inserting any markers or other alterations, thus avoiding indirect physiological effects. For overexpression, we applied restriction-free cloning to express the lincRNAs from a plasmid under the control of the strong, inducible *nmt1* promoter (Maundrell, 1993). Gene overexpression (‘gain-of-function’) provides complementary phenotype information to gene deletion (Prelich, 2012); moreover, any phenotype caused by a lincRNA that is ectopically expressed from a plasmid points to a function that is exerted over a distance (*in*

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

trans) via the lincRNA itself, rather than via its transcription or other local effects. We managed to delete 141 different lincRNAs (111 of which with at least two independent guide RNAs) and to overexpress 113 lincRNAs, with 104 lincRNAs being both deleted and overexpressed. These lincRNAs ranged in length from ~90 to 5100 nucleotides and in GC content from 25 to 46%, with means of 820 nucleotides and 34% GC, respectively. These lincRNAs are distributed across the entire nuclear genome (Figure 1A). Information for all deletion and overexpression strains analyzed is available in Supplemental Dataset 1.

To provide functional context for the lincRNA deletion-mutant phenotypes, we also assayed 238 coding-gene mutants from the *S. pombe* gene-deletion library, using prototrophic mutants after crossing out the auxotrophic mutants (Malecki and Bähler, 2016). These mutants broadly cover the Gene Ontology (GO) slim Biological Process categories (Lock *et al.*, 2019), ageing-related genes (Rallis *et al.*, 2014; Sideri *et al.*, 2014), as well as 104 'priority unstudied genes' (Wood *et al.*, 2019) (Supplemental Dataset 1).

Figure 1B provides an overview of the colony- and microscopy-based phenomics assays for the deletion and overexpression mutants. To determine fitness-related traits from colony-based assays, we applied *pyphe,* our python package for phenomics analyses (Kamrad, Rodríguez-López, *et al.*, 2020). Strains were arrayed randomly around a control grid at a density of 384 colonies per plate. We assayed the deletion or overexpression strains in response to diverse environmental factors such as different nutrients and drugs as well as oxidative, osmotic, heavy-metal, protein-homeostasis and DNA-metabolism stresses (Supplemental Figure 1), including some combined factors which can reveal additional phenotypes by non-additive effects that are not evident from single conditions (Rallis, Codlin and Bähler, 2013). For drugs and stressors, we applied low and high doses, where wild-type cell growth is normal or inhibited, respectively, to uncover both sensitive or resistant mutants. For the deletion mutants, we measured colony size to determine cell growth across 149 different nutrient, drug and stress conditions (Supplemental Dataset 1). For 68 of these conditions, we also measured colony redness using the phloxine B dye to determine cell viability (Lie *et al.*, 2018; Kamrad, Rodríguez-López, *et al.*, 2020). Cell growth and viability provided complementary functional information and produced strong biological signals (Figure 1C). For the overexpression mutants, we assayed cell growth across 47 conditions (Supplemental Dataset 1). All colony-based phenotyping was performed in at least three independent biological repeats per condition, with a median number of nine repeats per lincRNA.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Overall, we collected >1,100,000 phenotype data points for cell growth and >350,000 data points for cell viability. We established a normalisation procedure based on control grids to correct for known variations between and within plates which effectively reduces noise in the data (Figure 1C; Supplemental Figure 2A) (Kamrad, Rodríguez-López, *et al.*, 2020). Together with the high number of replicates, this normalisation provided the statistical power to confidently measure growth differences as small as 5%, thus supporting the detection of subtle lincRNA mutant phenotypes (Supplemental Figure 2B). Although control conditions measured in the same batch tended to be more similar, the batch effects remained much smaller than the biological signals (Supplemental Figure 2C). Thus, our colony-based phenotyping assays produce robust and reproducible results with high sensitivity. For the lincRNA deletion mutants, we also screened for cell size and cell-cycle traits using high-throughput microscopy and flow-cytometry analyses (Figure 1B). These assays added >20,000 phenotype datasets (microscopic fields analysed), with over 5.7 million cells analysed across 338 samples. Information for all phenotyping conditions is provided in Supplemental Dataset 1.

Phenotyping of deletion mutants in benign conditions

We screened for phenotypes of the lincRNA and coding-gene deletion mutants under benign, standard-laboratory conditions using rich and minimal growth media. We looked for mutants showing a significant difference in colony growth and/or colony viability compared to wild-type cells. Among the 141 lincRNA mutants tested, five and ten mutants grew slower than wild-type cells in rich and minimal media, respectively, while one mutant grew faster in minimal medium (Figure 2A; Supplemental Dataset 2). Among the 238 coding-gene mutants tested, 26 and 48 mutants grew slower in rich and minimal media, respectively, while four mutants each grew faster in rich and minimal media, three of which in both media (Figure 2A; Supplemental Datasets 3 and 4). Among the total of 51 coding-gene mutants growing slower in our assays, 49 have previously been associated with the phenotype ontology ‘abnormal vegetative cell-population growth’ (Harris *et al.*, 2013), thus validating our assay for this phenotype. With respect to colony viability, three lincRNA mutants showed lower viability than wild-type cells, two in rich medium and one in minimal medium (Figure 2B; Supplemental Datasets 2 and 3). Among the coding-gene mutants, 103 and 42 mutants showed higher or lower viability, respectively, in either or both benign conditions (Figure 2B; Supplemental Datasets 2 and 3). In conclusion, ~2-7% of the lincRNA mutants showed growth or viability phenotypes, compared to ~11-43% of the

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

coding-gene mutants, respectively. These results suggest that coding-gene mutants are more likely to have phenotypes in standard growth conditions. The results also illustrate that colony-viability assays can uncover phenotypes for many additional mutants not evident from colony-growth assays (Lie *et al.*, 2018; Kamrad, Rodríguez-López, *et al.*, 2020).

We examined additional, cellular phenotypes in rich medium for 110 lincRNA deletion mutants. Abnormal cell length or altered duration of cell-cycle stages point to defects in the cell-division cycle. Using high-throughput microscopy, we determined the length and proportion of bi-nucleated cells; these cells are fully grown and in G1/S phases of the cell cycle. In addition to wild-type cells, we used small *wee1-50* and large *cdc10-129* cell-cycle mutants as controls (Nurse and Hayles, 2019). Bi-nucleated wild-type cells showed a median length of 9.7µm, consistent with published data for ethanol-fixed cells (Heisler *et al.*, 2014). Two lincRNA mutants were significantly shorter than wild-type cells and four were longer (Figure 3A,B; Supplemental Figure 3A; Supplemental Dataset 2). Thus, these lincRNAs may be involved in the coordination of cell growth and division. Two of the size mutants, *SPNCRNA.989Δ* and *SPNCRNA.236Δ*, also showed strong slow-growth phenotypes (Figure 3C), but no anomalies in cell-cycle phases (Figure 3D; Supplemental Figure 4A,B). We independently validated the cell-length phenotypes of these two mutants by measuring calcofluor-stained cells growing in rich liquid medium fixed with formaldehyde. This analysis confirmed the shortened average length of *SPNCRNA.989Δ* cells (11.7±0.89µm; n=114) and extended median length of *SPNCRNA.236Δ* cells (12.7±0.92µm; n=155) compared to wild-type cells (12.1±0.75µm; n=129). These two mutants showed a range of other phenotypes and are further discussed below. We also detected phenotypes pointing to defects in transitions between cell-cycle phases: 22 and five lincRNA mutants showed significantly reduced and increased proportions of bi-nucleated cells, respectively, compared to the 13.2% bi-nucleated wild-type cells (Figure 3D; Supplemental Figure 3B). Four mutants showed both aberrant cell lengths and proportions of bi-nucleated cells: *SPNCRNA.819Δ* cells were shorter and had fewer bi-nucleates, while *SPNCRNA.323Δ, SPNCRNA.412Δ* and *SPNCRNA.934Δ* cells were longer and had more bi-nucleates (Figure 3A,C). We validated these microscopy data with high-throughput flow cytometry, with the results showing a good correlation (Supplemental Figure 4C). We conclude that several lincRNAs are involved in regulating cell size and/or cell-cycle progression.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Phenotyping of deletion mutants in multiple nutrient, drug and stress conditions

We assayed for colony-size (growth) phenotypes of the lincRNA and coding-gene mutants in the presence of various stresses or other treatments, relative to the same mutants growing in benign conditions and normalized for wild-type growth (Methods). To this end, we applied the same significance thresholds as for benign conditions. Among the 141 lincRNA mutants tested, 60 (43%) showed growth phenotypes in at least one condition (Supplemental Datasets 2 and 3). Together, these 60 mutants showed 211 growth phenotypes across conditions, with 69 of the 145 conditions producing phenotypes in at least one mutant (Figure 4A). The 211 hits included 150 resistant and 61 sensitive phenotypes (*i.e.,* mutants showing larger or smaller colonies, respectively, in assay conditions compared to the control condition, each relative to wild-type). Seven lincRNA mutants showed growth phenotypes in at least five conditions, with *SPNCRNA.236Δ* showing the most phenotypes, being resistant in 26 and sensitive in 2 conditions (Supplemental Datasets 2 and 3). Among all conditions, the most phenotypes were triggered by 0.075% MMS (causing DNA damage; 13 hits) and Brefeldin A (inhibiting protein transport from endoplasmic reticulum to Golgi; 18 hits).

Due to possible off-target mutations introduced by CRISPR/Cas9, we generated independent deletion mutants using different guide RNAs targeting the same lincRNA gene. These independent mutants generally produced highly similar growth phenotypes (Figure 4B). Of the 161 phenotypes associated with lincRNAs represented by two or more independent mutants, 112 phenotypes agreed between the corresponding mutants (all mutants showed median effect sizes >5%), and 27 hits showed a similar trend (median effect sizes >2%). In 16 cases, at least one guide RNA showed no phenotype (median effect sizes <2%), and in only 6 cases did the guide RNAs show opposite effects (Figure 4B). These results indicate that any secondary effects from CRISPR/Cas9-based gene deletions did not affect the consistency of our phenotype results in the vast majority of cases.

Among the 238 coding-gene mutants tested, 223 (93.7%) showed growth phenotypes in at least one condition, 104 of which representing priority unstudied genes that have remained entirely uncharacterized (Wood *et al.*, 2019). Together, these 223 mutants showed 1924 growth phenotypes across conditions, with 119 of the 145 conditions tested producing phenotypes in at least one mutant (Figure 4C; Supplemental Datasets 3 and 4). The 1924 hits included 651 resistant and 1273 sensitive phenotypes.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

We also assayed for colony-viability phenotypes of the lincRNA and coding-gene mutants across stress or other treatments relative to mutant cells growing in benign control conditions, normalized for wild-type growth. To this end, we applied the same quantitative redness scores and significance thresholds as for the benign conditions. Among the 141 lincRNA mutants tested, 25 (17.7%) differed in viability in at least one condition compared to wild-type cells (Supplemental Dataset 2). Together, these 25 mutants showed 98 phenotype hits across conditions, with 45 of the 67 conditions tested producing phenotypes in at least one mutant (Figure 4A). The 98 hits included 86 resistant and 12 sensitive phenotypes (higher and lower viability than wild-type, respectively). Two lincRNA mutants, which were sensitive in the benign condition, caused ~56% of the hits, all resistant, in conditions that partially suppressed this sensitive phenotype: *SPNCRNA.989Δ* (31 hits) and *SPNCRNA.1343Δ* (24 hits) (Supplemental Dataset 2). These lincRNAs are discussed further down. Among all conditions, the highest number of hits with viability phenotypes were observed in rich medium with 0.5M KCl or with 0.005% MMS (6 hits each) and in minimal medium with canavanine (5 hits).

Among the 238 coding-gene mutants tested, 172 (72.3%) showed viability phenotypes in at least one condition. Together, these 172 mutants showed 1874 phenotype hits across conditions, with 57 of the 67 conditions tested producing phenotypes in at least one mutant (Figure 4C; Supplemental Datasets 3 and 4). The 1874 hits included 1535 resistant and 339 sensitive phenotypes.

We then explored the relationships between colony-growth and -viability for the 67 conditions used to measure both phenotypes. The lincRNA mutants produced 140 growth phenotypes and 98 viability phenotypes, but in only 24 instances were both phenotypes associated with the same mutant (Figure 4D). The coding-gene mutants showed 1216 growth phenotypes and 1874 viability phenotypes, with only 310 instances where both phenotypes were associated with the same mutant (Figure 4D). A large excess of high-viability phenotypes was evident for coding-gene and, even more so, for lincRNA mutants (Figure 4D). Thus, slowly growing mutants did often show higher viability rather than lower viability, especially in coding-gene mutants. Together, these results further highlight that the colony-viability assays produce orthogonal phenotype information to the colony-growth assays and can uncover many additional phenotypes (Lie *et al.*, 2018; Kamrad, Rodríguez-López, *et al.*, 2020).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The lincRNAs that showed phenotypes were distributed across the genome (Figure 1A). They were not enriched in any particular gene-expression patterns, showing diverse responses to genetic or physiological perturbations (Supplemental Figure 5A). The lincRNAs associated with phenotypes were of similar length as those without phenotypes, but they tended to have a higher GC content (Supplemental Figure 5B). This result raises the possibility that the GC content reflects or even determines the likelihood of lncRNA function.

In conclusion, substantial proportions of the lincRNA mutants showed growth (43%) and/or viability (18%) phenotypes in some stress conditions, and the majority of coding-gene mutants showed phenotypes in these conditions (72-94%). With respect to viability phenotypes, much larger proportions of both lincRNA and coding-gene mutants were resistant (87.8% and 81.9%, respectively). This bias could partly reflect that many mutants are growing somewhat more slowly in benign conditions (Figure 2A), a trade-off that may render them more resilient to stresses (López-Maury, Marguerat and Bähler, 2008). Together, these analyses show that phenomics assays can effectively uncover functional clues not only for protein-coding genes but also for many lincRNAs.

Integrated analyses of functional signatures from deletion mutants

Using unsupervised clustering, we mined the rich deletion-mutant phenotype data to explore functional profiles for both protein-coding and lincRNA genes. For the phenotype calling described above, we wanted to identify functional clues and gene-environment interactions with high confidence (*i.e.,* low false discovery rate). Here, using a less conservative analysis, we applied a multivariate, global approach by converting effect sizes to a modified z-score to indicate the deviation from the expected phenotype value in units of standard deviations from the wild-type control in the same condition. Several conditions involved the same stressor, *e.g.* the same drug used at different doses (Supplemental Dataset 1). We aggregated such related conditions and used the strongest median response for each mutant and set of conditions (Methods). The protein-coding mutants generally showed stronger phenotypes than the lincRNA mutants as measured by the magnitude of the effect sizes (Supplemental Figure 6A). To compare phenotypes across the two types of mutants, we discretized the data, classing mutants as either sensitive (-1), resistant (+1), or similar (0) to their fitness in the corresponding control condition (Supplemental Figure 6B; Supplemental Dataset 5). Thresholds were chosen at ±1.5

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

standard deviations for both growth and viability data, which resulted in ~23% of all data points classed as non-zero in each dataset. We limited this analysis to 41 sets of aggregated ‘core’ conditions in which all mutants were phenotyped (Supplemental Dataset 5).

Applying this analysis, most lincRNA mutants showed few or no phenotypes across the 41 core conditions, while 16 lincRNA mutants showed strong phenotype profiles across many conditions. Such uneven distribution in the phenotype numbers associated with lincRNAs indicates that the data reflect biology rather than technical noise. In total, 194 mutants showed a phenotype in five or more sets of conditions, including the 16 lincRNA mutants, and these mutants were used for hierarchical clustering. Clear patterns were evident, and we divided the genes into three main clusters (Figure 5A; Supplemental Figure 6C; Supplemental Dataset 5). Clusters 1, 2 and 3 contained two, ten and four lincRNAs, respectively, providing an opportunity to infer function through ‘guilt by association’ with known protein-coding genes in the same clusters. This approach was somewhat limited because only 115 of the 178 protein-coding genes in the clusters had known or inferred biological roles. Using the AnGeLi tool (Bitton *et al.*, 2015), we identified functional enrichments for the clusters as described below.

Cluster 1 showed the most defined phenotype signature, characterized by many mutants displaying higher viability in 15 stress conditions, lower viability in the benign conditions and in canavanine A, and slow growth in benign conditions and several drugs tested, including hydrogen peroxide and antimycin A (Figure 5A; Supplemental Figure 6C; Supplemental Dataset 5). This cluster was enriched in various GO categories related to protein localization/transport, cellular respiration, phosphate metabolism, and protein translation (the latter including five cytosolic/mitochondrial ribosomal subunits, six translation factors, and three subunits of the elongator complex). The cluster also included nine genes involved in nutrient- or stress-dependent signalling (Supplemental Dataset 5). With respect to phenotype ontology (Harris *et al.*, 2013), this cluster was enriched in multiple terms related to cytoskeleton aberrations, abnormal respiration and translation as well as altered cell growth and stress sensitivity. Indeed, 80% of the mutants in this cluster have previously been associated with decreased cell population growth, and 87% are associated with increased sensitivity to chemicals. These enrichments validate our phenotype data.

Cluster 1 contained the two lincRNAs, *SPNCRNA.989* and *SPNCRNA.1343*, which accounted for ~56% of the colony-viability phenotypes among the lincRNA deletion mutants. When

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

overexpressed, however, they generated just 1-2 hits, much fewer than average (see below). This pattern suggests that these lincRNAs may function *in cis*, regulating nearby genes. Notably, both lincRNAs are located upstream of genes regulated by the Pho7 transcription factor (Schwer *et al.*, 2017), which functions during phosphate starvation and other stresses (Carter-O'Connell *et al.*, 2012): *SPNCRNA.989* and *SPNCRNA.1343* are divergently expressed to *atd1* and *tgp1,* respectively (Supplemental Figure 7). *SPNCRNA.1343* partially overlaps with the *nc-tgp1* RNA that regulates phosphate homeostasis by repressing the adjacent *tgp1* gene via transcriptional interference; deletion of *SPNCRNA.1343* has been shown to increase *tgp1* expression by inhibiting *nc-tgp1* expression (Ard, Tong and Allshire, 2014; Ard and Allshire, 2016; Shah *et al.*, 2014; Garg *et al.*, 2018; Yague-Sanz *et al.*, 2020). Inspection of the region upstream of *SPNCRNA.989* suggested a regulatory mechanism similar to *tgp1*, with divergent transcripts towards *atd1* likely driven by a bi-directional promoter from the nucleosome-depleted region upstream of *SPNCRNA.989* (Supplemental Figure 7). These patterns suggest that additional Pho7-regulated genes, like *atd1*, are controlled via upstream RNAs, similar to the *tgp1, pho84* and *pho1* genes that respond to phosphate limitation (Carter-O'Connell *et al.*, 2012). The similar phenotypes of *SPNCRNA.989Δ* and *SPNCRNA.1343Δ* mutants therefore suggests that these lincRNA deletions interfere with the expression of their neighboring genes and thus with processes affected by this regulon. In spotting assays, the phenotypes of *SPNCRNA.989Δ* and *SPNCRNA.1343Δ* often differed from those of *tgp1Δ* and *atd1Δ* (Supplemental Figure 8). These results are consistent with the lincRNA deletion leading to induction, rather than repression, of their coding-gene neighbors.

Cluster 2 contained a majority of genes that are not associated with any functional annotations, including ten lincRNAs genes (Supplemental Figure 6C; Supplemental Dataset 5). This cluster was enriched for long-lived mutants and for genetic interactions (based on Biogrid data; (Breitkreutz *et al.*, 2008), meaning that the protein-coding genes within this cluster are ~4 times more likely to interact with each other than expected by chance. This cluster included seven genes involved in stress and/or nutrient signalling pathways and six genes for transcription factors functioning during stress/nutrient responses or in unknown processes. The phenotype data in this cluster were sparse and lacked a convincing functional signature across the coding and lincRNA genes.

Cluster 3 was characterized by most mutants showing rapid growth in valproic acid, formamide and sodium orthovanadate, and many of these mutants also showed higher viability in benign

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

conditions but lower viability in valproic acid. This cluster was enriched for long-lived mutants and for energy metabolism, including four genes each functioning in glycolysis and the TCA cycle. Intriguingly, one of the four lincRNA genes in this cluster, *SPNCRNA.236*, is located upstream the pyruvate-kinase gene *pyk1,* which is involved in the last step of glycolysis to generate pyruvate for the TCA cycle or fermentation*.* The finding that the *SPNCRNA.236Δ* mutant leads to a similar phenotypic signature as does deletion of glycolysis or TCA-cycle genes raises the possibility that *SPNCRNA.236* acts *in cis* to control *pyk1* expression. Consistent with this idea, *SPNCRNA.236Δ* mutants grow slowly (Figure 2A) while increased activity of Pyk1 leads to faster growth (Kamrad, Grossbach, *et al.*, 2020). However, *SPNCRNA.236* also generates phenotypes in 11 conditions when overexpressed from a plasmid (Supplemental Dataset 6), including faster growth in minimal medium which is the opposite of the slower growth of the deletion mutant in the same condition. Thus, it is also possible that *SPNCRNA.236* can act *in trans*.

We validated phenotypes of five lincRNA deletions from Clusters 1 and 3 as well as deletions from neighboring coding genes, using serial-dilution spotting assays under 13 conditions. Detection of subtle phenotypes involving 5% differences in growth is difficult with such spotting assays. Nevertheless, we could confirm 11 (84%) of the phenotypes detected by the high-throughput colony-based assays (Supplemental Figure 8). We conclude that there is generally a good agreement between these different phenotyping assays.

To further explore our dataset, we discretized all deletion-mutant phenotypes, both from lincRNAs and coding genes (Supplemental Datasets 2-4). Pearson correlations of phenotype profiles were then used for constructing a network that was visualized with Cytoscape (Shannon *et al.*, 2003). The network included several distinct clusters. A large, tight cluster consisted mostly of protein-coding genes (Figure 5B, highlighted in red). This cluster, which was similar to Cluster 1 (Figure 5A), included *SPNCRNA.989* and was enriched for genes involved in phosphate metabolism and translation and 89% of the mutants in this cluster displayed slow growth phenotypes (Figure 5B). Another large cluster was enriched for lysine metabolism with 18% of the mutants showing ageing-related phenotypes such as increased lifespan during quiescence (Sideri *et al.*, 2014). This cluster also included three lincRNAs: *SPNCRNA.318, SPNCRNA.426* and *SPNCRNA.965* (Figure 5B, highlighted in blue). Network analysis of the whole phenotypic dataset revealed further connections between several lincRNAs and coding genes. For example, a negative phenotypic correlation was evident between *SPNCRNA.1460*

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

and *fbp1* (Figure 5B, upper left); *fbp1* is a key gene responding to glucose starvation that is regulated by upstream non-coding RNAs (Hoffman and Winston, 1990; Hirota *et al.*, 2008)(Oda *et al.*, 2015). *SPNCRNA.1460* is located upstream of *scr1,* encoding a transcriptional repressor that negatively regulates *fbp1* (Tanaka et al., 1998; Vassiliadis et al., 2019). This link raises the possibility that *SPNCRNA.1460* controls *scr1* expression and, therefore, *fbp1* expression*.* The same cluster also included *car2,* which is also implicated in carbon metabolism, and two priority unstudied genes, whose association suggests that they function in similar processes. Interestingly, some clusters consisted exclusively or mostly of lincRNAs (Figure 5B, upper left). Naturally, these clusters showed no functional enrichments, but they point to several lincRNAs acting in related cellular processes, possibly together.

Phenotyping of lincRNA overexpression mutants in multiple conditions

Gene overexpression provides complementary phenotype information to gene deletion (Prelich, 2012). We constructed strains that ectopically overexpressed 113 lincRNAs from a plasmid under the strong *nmt1* promoter in minimal medium (Methods). We then looked for differences in colony growth under benign conditions compared to empty-vector control cells. We also looked for growth phenotypes in the presence of various stresses or other treatments, relative to growth in benign control conditions and normalized for the growth of empty-vector control cells. In the benign condition, most lincRNA overexpression strains grew faster compared to the empty-vector control. This pattern may reflect an indirect effect of lincRNA transcription by increasing plasmid copy numbers and/or expression of the budding yeast *LEU2* marker that is limiting for growth. Therefore, we normalized the colony growth of overexpression mutants in the stress conditions by the growth in the benign condition to correct for this potential bias. We used a more stringent significance threshold for the overexpression mutants than for the deletion mutants (Figure 6A), because ectopic overexpression of genes involves cell-to-cell variation in plasmid copy numbers, leading to higher phenotypic heterogeneity (Siam, Dolan and Forsburg, 2004). Among the 113 lincRNA overexpression strains tested, 102 (90.3%) showed growth phenotypes in at least one condition (Figure 6A; Supplemental Dataset 6). Together, these 102 overexpression strains showed 565 growth phenotypes across conditions. The 565 hits included 347 resistant and 218 sensitive phenotypes (*i.e.,* mutants showing larger or smaller colonies, respectively, in the assay condition than in the control). Fourteen lincRNA overexpression strains showed more consistent phenotypes in ten or more conditions, topped

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

by *SPNCRNA.335* that showed sensitive and resistant phenotypes in twelve and three conditions, respectively (Supplemental Figure 9A).

With respect to the 47 conditions tested, 42 produced phenotypes in at least one mutant (Supplemental Dataset 6). Over 80% of the 565 phenotypes came from only 21 of the 47 conditions, and ~24% of the phenotypes came from just three conditions: proline as a nitrogen source, 5 mM valproic acid (VPA), and 10 mM hydroxyurea (HU) (Supplemental Figure 9B). Proline is a poor nitrogen source which causes nitrogen stress and slows growth (Davie, Forte and Petersen, 2015). Notably, the expression of many lincRNAs is strongly induced under nitrogen starvation (Atkinson *et al.*, 2018). Together with the results presented here, this indicates that several lincRNAs function during nitrogen stress. VPA is an inhibitor of histone deacetylases, and some lincRNAs are involved in histone modification (Rinn and Chang, 2012). Roles for lincRNAs in neurological disorders such as epilepsy are also emerging, and VPA can ameliorate epilepsy partly by repressing some of these lincRNAs (Hauser, Henshall and Lubin, 2018). HU inhibits DNA synthesis, and our findings suggest that several lincRNAs are involved in related processes. For example, some lincRNAs have been implicated in double-strand break repair (Bader *et al.*, 2020). We propose that several lincRNAs can influence cellular growth *in trans* when ectopically overexpressed under conditions that affect certain cellular processes.

We looked for functional signatures in the lincRNA overexpression phenotypes using hierarchical clustering (Figure 6B). Unlike for lincRNA deletion mutants, this data set did not provide the functional context from coding-gene mutants. We observed a conspicuous sub-cluster of four lincRNA overexpression strains that showed slower growth in many of the conditions (Figure 6B; highlighted in red). One of these four overexpressed lincRNAs was the well-characterized *meiRNA* that functions in the induction of meiosis (Watanabe and Yamamoto, 1994). In mitotically growing cells, *meiRNA* binds to Mmi1 via its DSR motif and is degraded by the nuclear exosome, while upon induction of meiosis, *meiRNA* binds to the RRM motif of Mei2 which in turn promotes meiosis (Yamashita, 2019). The other three lincRNAs in this sub-cluster contain motifs for potential Mei2 binding, including two DSR motifs. Moreover, like *meiRNA,* the other three lincRNAs are also de-repressed in nuclear-exosome mutants and during meiosis (Atkinson *et al.*, 2018). Together, these findings raise the possibility that the other three lincRNAs in the sub-cluster also function in meiosis.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

We compared the phenotype data from deletion and overexpression mutants. We obtained data for both types of mutants for 104 lincRNAs in 22 conditions, 18 of which showed phenotypes. Of these 104 lincRNAs, only seven did not produce any phenotypes in any condition tested. Under the 18 conditions, a higher proportion of lincRNA overexpression mutants (86.5%) than deletion mutants (32.7%) produced phenotype hits in at least one condition. Moreover, lincRNA overexpression generally resulted in larger effect sizes, *i.e.* stronger phenotypes, than did lincRNA deletion (Figure 6C). Similar trends have been reported for coding-gene mutants. For example, 646 and 1302 growth phenotypes are caused by deletion and overexpression mutants, respectively, of non-essential budding yeast genes (Yoshikawa *et al.*, 2011), and 64 transcription-factor genes of fission yeast show growth phenotypes when overexpressed but not when deleted (Vachon *et al.*, 2013). Only a few lincRNAs showed phenotypes as both deletion and overexpression mutants. For example, *SPNCRNA.236* showed rapid-growth (resistant) phenotypes in both overexpression and deletion mutants in five conditions, while in the benign condition, the overexpression and deletion mutant showed rapid and slow growth, respectively (Supplemental Datasets 2 and 6). In general, our phenotype data for lincRNA deletion and overexpression mutants showed little overlap and poor correlation (Figure 6C). These results illustrate the complementary information provided by these two types of mutants.

How might lincRNA overexpression result in more phenotypes than lincRNA deletion? Overexpression of a protein-coding gene can burden cells via resource-consuming translation or toxic protein levels (Moriya, 2015; Bolognesi and Lehner, 2018). Although overexpression of a single lincRNA should not affect resource allocation, it is possible that these AT-rich RNAs engage in non-specific molecular interactions. Recent findings indicate that RNAs are assembly prone and must be tightly regulated as they can promote paraspeckles, stress granules and phase separation (Fox *et al.*, 2018; Van Treeck *et al.*, 2018). Such processes could trigger the overexpression phenotypes in certain physiological conditions, reflecting that lincRNAs are biologically active molecules. Nevertheless, the observed 565 phenotype hits amount to only 10.6% of the potential 5311 hits if overexpression of all 113 lincRNAs caused phenotypes in the 47 conditions tested (Supplemental Dataset 6). Thus, overexpressed lincRNAs do not generally lead to any non-specific or toxic effects, and the observed phenotypes may, therefore, mostly reflect specific lincRNA functions. Given that many lincRNAs may function in specialized conditions (Cabili *et al.*, 2011; Derrien *et al.*, 2012; Pauli *et al.*, 2012; Hon *et al.*, 2017; Atkinson *et al.*, 2018), deletion mutants will only reveal phenotypes when assayed in the relevant conditions. On the other hand, the ‘gain-of-function’ overexpression mutants may also reveal

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

phenotypes in conditions where the lincRNAs do not normally function. Notably, phenotypes that are caused by lincRNAs being ectopically expressed from plasmids point to a function that is exerted *in trans*, via the lincRNA itself, rather than via its transcription or other *cis* effects. Our findings therefore raise the possibility that many of the lincRNAs tested can function over a distance.

Conclusions

We applied a phenomics approach to explore the functional importance of *S. pombe* lincRNAs, including colony-based and cellular assays of deletion mutants and colony-based assays of overexpression strains. A panel of deletion mutants of coding-genes were screened in parallel for comparison and functional context. Together, these assays revealed phenotypes for 84 of 141 deleted lincRNAs, 229 of 238 deleted coding-genes, and 102 of 113 overexpressed lincRNAs. This extensive phenotyping uncovers lincRNAs that contribute to cellular resistance or sensitivity in specific conditions, reflected by altered colony growth and/or viability, and lincRNAs that are involved in the size control and the cell-division cycle. As expected, higher proportions of coding-gene mutants showed phenotypes and these phenotypes tended to be stronger (larger effect sizes) than for lincRNA mutants. In benign conditions, the lincRNA mutants were ~3- and 30-fold less likely to show phenotypes for colony growth or viability, respectively, than coding-gene mutants. This difference was less pronounced in the nutrient, drug and stress conditions, where many more lincRNA mutants revealed phenotypes, at only ~2- to 4-fold lower proportions than coding-gene mutants for colony growth or viability, respectively. Moreover, compared to lincRNA deletion mutants, the lincRNA overexpression strains were ~2-fold more likely to show phenotypes, which also tended to be stronger. Together, these findings support the notion that most lincRNAs play specialized roles in specific conditions. The findings also indicate that lincRNAs in general have more subtle functions than proteins, *e.g.* in fine-tuning of gene expression. Accordingly, it was important that our high-throughput assays were highly sensitive to detect subtle phenotypes. We conclude that a substantial proportion of lincRNAs exert cellular functions under certain conditions, and many of which may act *in trans* as RNAs. This analysis provides a rich framework to mechanistically dissect the functions of these lincRNAs in the physiologically relevant conditions.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Methods

Deletion and overexpression strain libraries

Using a CRISPR/Cas9-based approach and primer-design tool for seamless genome editing (Rodríguez-López *et al.*, 2016), we deleted 141 different lincRNA genes located across all the *S. pombe* chromosomes (Figure 1A; Supplemental Dataset 1 for coordinates). In total, 113 lincRNA genes were deleted in the *972 h*$^-$ background, and 70 lincRNA genes were deleted in the *968 h*^{90} background, the latter including 15 newly identified lincRNAs (Atkinson *et al.*, 2018). Thirty lincRNAs were deleted with one guideRNA (gRNA), 103 were deleted using two gRNAs, and eight were deleted using three gRNAs. For the protein-coding deletion mutants, we generated a prototroph version of Bioneer V.5 deletion library (Kim *et al.*, 2010) as described (Malecki and Bähler, 2016). Strains were arranged into 384-colony format using a RoToR HDA colony-pinning robot (Singer Instruments), including a 96-colony grid of wild-type *972 h*$^-$ strains for plate normalization (Kamrad, Rodríguez-López, *et al.*, 2020). We selected a subset of genes to broadly cover all main GO categories, together with 91 uncharacterized genes. Supplemental Dataset 1 provides information on the individual strains.

We generated ectopic overexpression constructs for 113 long intergenic lincRNAs using the *nmt1* promoter (Maundrell, 1993). The predicted full-length lincRNA sequences were amplified by PCR and cloned into the pJR1-41XL vector (Moreno, Durán and Ribas, 2000) using the CloneEZ® PCR Cloning Kit (GenScript). Each plasmid was checked by PCR for correct insert size. All primer sequences are provided in Supplemental Dataset 1. Plasmids were then transformed into *S. pombe* cells (*h*$^-$*, leu1-32*), and leucine prototroph transformants were selected on solid Edinburgh Minimal Medium (EMM2) plates. An empty-vector control strain was created analogously. Of the 113 lincRNAs, 67 were represented by two independently cloned vectors (Supplemental Dataset 1).

High-throughput phenotyping of deletion mutants on solid media

The deletion mutants were broadly phenotyped using a colony-based phenomics platform as described (Kamrad, Rodríguez-López, *et al.*, 2020; Kamrad, Bähler and Ralser, 2021). Mutants were assayed on solid media with a variety of 55 unique stressors, using different

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

concentrations and, in some cases, combinations of stressors. In total we assayed 134 different conditions, with the viability dye phloxine B being included in 66 of these conditions. Supplemental Figure 1 provides a description of the conditions used for phenotyping. Supplemental Dataset 1 contains the concentrations of all the stressors used. Cells were grown for 24 hours on yeast extract supplement (YES) plates in 384-colony format containing a wild-type control grid, followed by pinning cells onto plates containing the stressors, using reduced pressure (4% pinning pressure to transfer a small amount of biomass). Plates were wrapped in plastic to avoid drying out and incubated for ~40 hours at 32°C, unless stated otherwise, before image acquisition and phenotype assessment.

Image acquisition and quantitation, data normalisation and processing as well as hit calling was performed using our *pyphe* pipeline which is available here: www.github.com/Bahler-Lab/pyphe (Kamrad, Rodríguez-López, *et al.*, 2020). Images of plates were acquired with a flatbed scanner (Epson V800 Photo), controlled by *pyphe-scan* through SANE. Images for quantifying colony area (growth) were taken by transmission scanning, using the *--mode Grey* argument. For quantifying redness/viability, images were taken by reflective scanning using *--mode Color*. Images were acquired at 300 or 600 dpi resolution. For colour images, to determine colony redness for viability, we used an opaque fixture to hold the plates in place, the white cover was installed in the scanner lid, and the scanner was covered by a cardboard box to prevent external light interfering with image acquisition. Images were inspected individually and excluded if one or more of the following applied: several colonies were missing due to pinning errors (usually in the corners), evidence of contamination, white background had not been inserted during colour scanning, and/or plate had slipped significantly during scanning so that a whole row/column of colonies was missing from the image. The overall number of excluded plates was low and generally did not result in significant data loss in the final dataset due to the large number of replicate plates.

For image quantification, greyscale transmission images for colony area quantitation were analysed with the R package *gitter* (Wagih and Parts, 2014) using the following parameters: *plate.format=384, inverse="TRUE", remove.noise="TRUE", autorotate="TRUE"*. Images for which *gitter* failed (very few) were excluded from further analysis. Colour images for redness/viability quantification were analysed with *pyphe-quantify* using default parameters.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

For data normalisation and processing, an experimental design table was prepared for each dataset which listed for each plate the path to the data file produced during image quantification, plate layout information, the condition as well as other meta-data (*e.g*. batch number, replicate counter and free-text comments). Data from all images of the same dataset was parsed and processed simultaneously using *pyphe-analyse*, producing a single data report table in tidy format per experiment, containing all data associated with a single measured colony on each line. For analysis of colony areas, the following parameters were used: *--format gitter --load_layouts --gridnorm standard384 --rcmedian --check.* For colony redness analysis the options were: *-format pyphe-quantify-redness --load_layouts --rcmedian --check.*

Pyphe performs some automated quality control. Specifically, during grid normalisation, missing reference grid colonies are flagged and all neighbouring colonies are set to NA. *Pyphe* also checks data for negative and infinite fitness values (rare artefacts of normalisation procedures). For the colony size datasets, additional quality control of the data was performed as follows: missing colonies (colony size 0 reported by gitter and fitness 0 reported by *pyphe-analyse*) were set to NA as these are pinning errors; colonies with a circularity (reported by *gitter*) below 0.85 were set to NA; plates with a CV of >0.2 for wild-type controls were set to NA. For viability datasets, the only QC step was to exclude plates with a wild-type CV of >0.05.
For statistical analysis, tables reporting summary statistics and p-values for each lincRNA gene and condition were obtained with *pyphe-interpret.* Hits were called separately for control conditions (where we tested for difference in means between each lincRNA mutant and wild-type control in the same condition) and all other conditions (where we tested for difference in means between each lincRNA mutant in test condition *vs* corresponding control condition). Welch's t-test, which does not assume homogeneity of variances, was used and the obtained p-values were corrected for multiple testing for each condition separately using the Benjamini-Hochberg method (Benjamini and Hochberg, 1995).

The dataset for clustering (Supplemental Dataset 5; Figure 5) was derived from Supplemental Datasets 2 and 4 by subtracting 1 from the median effect sizes and dividing by the standard deviation of the wild type control for each condition. Conditions were then aggregated by choosing the strongest response across all repeats of the same stressor (the stressor is indicated in the 'stress_description' column in the knock-out_condition_metadata sheet of Supplemental Dataset 1. As not all lincRNA mutants were phenotyped in all conditions, clustering was restricted to a set of 41 core stressors. LincRNA or coding-gene mutants with

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

less than 5 responses were excluded, leaving 194 mutants in total, including 16 ncRNAs). The final dataset only contained 17 NA values which were imputed with 0. Hierarchical clustering was done with *scipy (Virtanen et al., 2020)*, using the Ward method and the Euclidean distance metric. Clusters were obtained by cutting the dendrogram using the fcluster function with the 'maxclust' method. Functional enrichments in Clusters 1 to 3 were analysed using AnGeLi (Bitton *et al.*, 2015), with all protein-coding genes as background list.

Data for the phenotypic correlation network (Fig. 5B) were generated from phenotypes for all lincRNA and coding-gene deletions, using a ternary system: resistant, sensitive and no phenotype encoded as 1, -1 and 0, respectively. The network was generated following general instructions (Contreras-López *et al.*, 2018)(Shannon *et al.*, 2003). Briefly, we used Pearson correlations to calculate the network and filtered on absolute *r* values above 0.6 and adjusted *p* <0.01. Clustering of the network in Cytoscape was done using community clustering (GLay) from the clustermaker extension (Morris *et al.*, 2011).

High-throughput microscopy and flow cytometry for cell-size and cell-cycle phenotypes

Strains, frozen in glycerol in 384-colony format, were revived in YES solid plates, resuspended into 150 µl of liquid YES in 96 well plates and incubated at 32°C for 16 hours. Then, 100 µl from these pre-cultures were added to 1.5 ml of preheated (32°C) liquid YES in 96 deep-well plates and incubated at 32°C for 8 hours. Cells were collected by centrifugation, cell pellets were resuspended in 70% ice-cold ethanol, and stored in the dark at 4°C until further processing. As cell-size and cell-cycle phenotype controls, we used two temperature-sensitive cell-cycle mutants: *cdc10-129* and *wee1-50*. These mutants were grown in 50 ml YES at 25°C, centrifuged and re-suspended in 50 ml of prewarmed (37°C) YES and incubated for 4 hours at 37°C to block cell-cycle progression. After 4 hours, 1 ml of the samples was fixed for microscopy and flow cytometry. The remaining cells were centrifuged and re-suspended in 50 ml of prewarmed YES (25°C), incubated at 25°C, and samples collected and fixed after 20 and 60 min. Over 80% of the 110 lincRNA mutants screened for cellular phenotypes were assayed in at least 2 independent biological repeats.

For cell-size and cell-cycle phenotypes, fixed cells were washed in 50 mM sodium citrate buffer, spun down at 3000g for 5 min, re-suspended in 50 mM sodium citrate containing 0.1 mg/ml

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

RNAse A, and incubated at 37°C for 2 hours. Cells were then spun down at 3000g for 5 min and re-suspended in 500 µl of 50 mM sodium citrate + 1 µM SYTOX Green (ThermoFisher Scientific, Cat. no. S7020). Immediately prior to analysing samples using either high-throughput (HTP) flow cytometry or HTP microscopy, cells in the deep well plates were sonicated for 40 sec at 50W (JSP Ultrasonic Cleaner model US21) to increase the efficiency of singlets.

For HTP-image acquisition, cells were further stained with a 1:1000 dilution of CellMask™ Deep Red Plasma membrane dye (ThermoFisher Scientific, Cat. no. C10046), according to the manufacturer's instructions. Then, 2.5 µl of fixed and stained cells were transferred from 96-well plates into a poly-lysine-coated 384-well Perkin Elmer Cell Carrier Ultra imaging plate (PerkinElmer, Cat. no. 6057500), pre-filled with 25 µl of 1µM SYTOX Green using a Biomek Fx robot. Cells were spun down for 3 min at 200g before imaging. Imaging was performed on a Perkin Elmer Opera Phenix microscope, using a water immersion 63x lens to capture confocal stacks of 12 planes in both Alexa488 (SYTOX Green) and Alexa647(CellMask) channels, with 63 microscopic fields being captured per sample. The images were projected and analysed using the associated Phenix software Harmony, for the automated identification of mono- and bi-nucleated cells and respective cell length. Features were exported for further analysis using R studio.

For HTP flow cytometry, 250 µl of cells were transferred into 96-well plates and 30,000 cells were measured in a Fortessa X20 Flow cytometer (BD Biosciences), using the HTS plate mode on the DIVA software and a 488 nm excitation laser to capture the SytoxGreen DNA staining. Populations of interest were gated as described (Knutsen *et al.*, 2011) using the FlowJo software version 10.3.0. Features of interest (populations with different DNA content) were then exported for further analysis using R studio. The determined percentage of cells in each cell-cycle phase per sample was used to validate the HTP-imaging data. For correlation with the HTP-imaging (bi-nucleated cells), S- and G1-phase cell populations were grouped together (Supplemental Dataset 2).

Data analysis was carried out in R (v.3.5.2), using the package tidyverse for data manipulation, visualisation and statistical analysis. All tests were two-sided unless otherwise stated. For HTP-imaging analysis, cell density was checked for each sample in the multi-well plate and given a score of 0 to 5, where 0 is very low to no density (<50 cells/well) and 5 is at too high density; samples scoring 0 and 5 were excluded from analysis. For cell-size analysis, the median cell

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

size of bi-nucleated cells for each mutant was used to calculate fold-changes relative to wild-type values, applying the Wilcox test to determine significant differences ($p < 0.05$), only considering cells showing a ≥5% difference in size compared to wild-type cells. For cell-cycle analysis by HTP imaging, the percentage of bi-nucleated cells per microscopic field (63 fields/sample) was used to calculate the median value per lincRNA mutant, followed by fold-change analysis calculated by normalising the percentage of bi-nucleated cells in each sample relative to wild-type values, applying the Wilcox test to determine significant differences ($p < 0.05$), only considering cells showing a ≥20% difference in bi-nucleated cells compared to wild-type.

High-throughput phenotyping of overexpression mutants on solid media

Overexpression strains were arrayed in 384 format together with the empty-vector control strain and a grid of the wild-type strain (*972 h-*) for normalisation. Strains were revived from glycerol stocks in YES and grown for 2 days at 32°C. Colonies were then transferred to new YES plates, grown for 1 day, and pinned onto EMM2 (with NH_4Cl but without amino acid supplements) with or without the specified drugs/supplements. YES medium contains thiamine that represses the *nmt1* promoter and leucine that compromises the maintenance of the overexpression plasmid (which contains the *LEU2* marker). We screened the lincRNA overexpression library for colony growth phenotypes in 47 conditions (Supplemental Datasets 1 and 6). Each overexpression strain was represented by at least 12 colonies across 3 different plates and experiments were repeated at least 3 times. Each condition was assayed in three independent biological repeats, together with control EMM2 plates, resulting in at least 36 data points per strain per condition. Plates were incubated at 32°C if not stated otherwise for the condition. Plates were imaged as described for deletion mutants after 40 or 64 hours in order to capture as many hits as possible.

Image acquisition and quantification, data normalisation and processing as well as hit calling were performed using the *pyphe* pipeline as described above for grayscale transmission images to quantify colony sizes. During grid correction, 24,683 colonies were excluded due to missing grid colonies, and 2,539 missing colonies were set to “NA” (pinning errors), and data from 290 of 2772 plates were discarded because they either showed a fraction of unexplained variance (FUV) above 1 or a control CV of >0.5. The final dataset contained 917,368 data points. The

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

colonies which passed the above quality control steps were normalised with the grid first, and the resulting colony sizes were additionally normalised to the control condition (EMM2) for the conditions with stressors. All data from the *pyphe* analysis are provided in Supplemental Dataset 6. The hits were defined by adjusted (Benjamini-Hochberg) p-values ≤0.01 and median effect size (MES) ≥5% compared to empty-vector control.

For clustering analyses, we first filtered the relative log2 MES data (relative to empty-vector control) for genes with 5 or more hits followed by conditions with 5 or more hits, resulting in 59 lincRNA mutants with MES data for 29 conditions. Then we discretized the data, classing mutants as either sensitive (-1), resistant (+1), or similar to their fitness in the corresponding control condition (0). We performed hierarchical clustering with the complete method using the Canberra distance metric, and plotted the heatmap (Figure 6B) with the ComplexHeatmap r-package (Gu, Eils and Schlesner, 2016).

For correlation analyses between deletion *vs* overexpression data, we filtered the phenotyping data for the 104 shared lincRNA mutants and the 22 shared stress conditions between the two mutant types. As the overexpression strains could only be assayed on minimal media while the deletion strains were mainly assayed on rich media, we matched conditions based on the added drug/stressor only, disregarding the media background. In case of multiple related conditions (*e.g.* same stress in different doses), the strongest response was used (maximum median effect size).

Acknowledgments

We thank Sarah Aldous, Natalia Borbarán Bravo, Mike Jiang, Janina Karl, Patrycja Krawczyk, Sara Kurz, Kira Kyne, Shirley Liu, and Florence Young for help with deletion and overexpression constructs. This work was supported by a Wellcome Trust Senior Investigator Award (095598/Z/11/Z) and BBSRC Research Grants (BB/R018219/1 and BB/R009597/1) to JB. ST was supported by a Boehringer Ingelheim Fonds PhD Fellowship. The Francis Crick Institute receives its core funding from Cancer Research UK (FC001134), the UK Medical Research Council (FC001134) and the Wellcome Trust (FC001134).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

References

Abdelmohsen, K. *et al.* (2013) 'Senescence-associated lncRNAs: senescence-associated long noncoding RNAs', *Aging Cell*, 12(5), pp. 890–900.

Ard, R. and Allshire, R. C. (2016) 'Transcription-coupled changes to chromatin underpin gene silencing by transcriptional interference', *Nucleic Acids Res*, 44(22), pp. 10619–10630.

Ard, R., Allshire, R. C. and Marquardt, S. (2017) 'Emerging Properties and Functional Consequences of Noncoding Transcription', *Genetics*, 207(2), pp. 357–367.

Ard, R., Tong, P. and Allshire, R. C. (2014) 'Long non-coding RNA-mediated transcriptional interference of a permease gene confers drug tolerance in fission yeast', *Nat Commun*, 5, p. 5576.

Atkinson, S. R. *et al.* (2018) 'Long noncoding RNA repertoire and targeting by nuclear exosome, cytoplasmic exonuclease, and RNAi in fission yeast', *RNA* , 24(9), pp. 1195–1213.

Bader, A. S. *et al.* (2020) 'The roles of RNA in DNA double-strand break repair', *British J Cancer*, 122(5), pp. 613–623.

Balarezo-Cisneros, L. N. *et al.* (2021) 'Functional and transcriptional profiling of non-coding RNAs in yeast reveal context-dependent phenotypes and in trans effects on the protein regulatory network', *PLoS Genet* , 17(1), p. e1008761.

Bassett, A. R. *et al.* (2014) 'Considerations when investigating lncRNA function in vivo', *eLife*, 3, p. e03058.

Batista, P. J. and Chang, H. Y. (2013) 'Long noncoding RNAs: cellular address codes in development and disease', *Cell*, 152(6), pp. 1298–1307.

Benjamini, Y. and Hochberg, Y. (1995) 'Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing', *J Royal Stat Soc: Series B*, pp. 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x.

Bester, A. C. *et al.* (2018) 'An Integrated Genome-wide CRISPRa Approach to Functionalize lncRNAs in Drug Resistance', *Cell*, 173(3), pp. 649–664.e20.

Bitton, D. A. *et al.* (2015) 'AnGeLi: A Tool for the Analysis of Gene Lists from Fission Yeast', *Frontiers Genet*, 6, p. 330.

Bolognesi, B. and Lehner, B. (2018) 'Reaching the limit', *eLife*. doi: 10.7554/eLife.39804.

Breitkreutz, B.-J. *et al.* (2008) 'The BioGRID Interaction Database: 2008 update', *Nucleic Acids Res*, 36(Database issue), pp. D637–40.

Brochado, A. R. and Typas, A. (2013) 'High-throughput approaches to understanding gene function and mapping network architecture in bacteria', *Curr Opin Microbiol*, pp. 199–206. doi: 10.1016/j.mib.2013.01.008.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Cabili, M. N. *et al.* (2011) ‘Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses’, *Genes & Dev*, 25(18), pp. 1915–1927.

Cao, H., Wahlestedt, C. and Kapranov, P. (2018) ‘Strategies to Annotate and Characterize Long Noncoding RNAs: Advantages and Pitfalls’, *Trends Genet*, 34(9), pp. 704–721.

Carter-O’Connell, I. *et al.* (2012) ‘Genome-wide characterization of the phosphate starvation response in Schizosaccharomyces pombe’, *BMC Genomics*, 13, p. 697.

Clément-Ziza, M. *et al.* (2014) ‘Natural genetic variation impacts expression levels of coding, non-coding, and antisense transcripts in fission yeast’, *Mol Syst Biol*, 10, p. 764.

Contreras-López, O. *et al.* (2018) ‘Step-by-Step Construction of Gene Co-expression Networks from High-Throughput Arabidopsis RNA Sequencing Data’, *Methods Mol Biol*, 1761, pp. 275–301.

Davie, E., Forte, G. M. A. and Petersen, J. (2015) ‘Nitrogen regulates AMPK to control TORC1 signaling’, *Curr Biol*, 25(4), pp. 445–454.

Derrien, T. *et al.* (2012) ‘The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression’, *Genome Res*, 22(9), pp. 1775–1789.

DeWeerdt, S. (2019) ‘RNA therapies explained’, *Nature*, 574(7778), pp. S2–S3.

Ding, D.-Q. *et al.* (2012) ‘Meiosis-specific noncoding RNA mediates robust pairing of homologous chromosomes in meiosis’, *Science*, 336(6082), pp. 732–736.

Duncan, C. D. S. and Mata, J. (2014) ‘The translational landscape of fission-yeast meiosis and sporulation’, *Nat Struct Mol Biol*, 21(7), pp. 641–647.

Eser, P. *et al.* (2016) ‘Determinants of RNA metabolism in the Schizosaccharomyces pombe genome’, *Mol Syst Biol*, 12(2), p. 857.

Fauquenoy, S. *et al.* (2018) ‘Repression of Cell Differentiation by a cis-Acting lincRNA in Fission Yeast’, *Current biology: CB*, 28(3), pp. 383–391.e3.

Fox, A. H. *et al.* (2018) ‘Paraspeckles: Where Long Noncoding RNA Meets Phase Separation’, *Trends Biochem Sci*, 43(2), pp. 124–135.

Garg, A. *et al.* (2018) ‘A long noncoding (lnc)RNA governs expression of the phosphate transporter Pho84 in fission yeast and has cascading effects on the flanking prt lncRNA and pho1 genes’, *J Biol Chem*, pp. 4456–4467. doi: 10.1074/jbc.ra117.001352.

Grech, L. *et al.* (2019) ‘Fitness Landscape of the Fission Yeast Genome’, *Mol Biol Evol*, 36(8), pp. 1612–1623.

Gu, Z., Eils, R. and Schlesner, M. (2016) ‘Complex heatmaps reveal patterns and correlations in multidimensional genomic data’, *Bioinformatics* , 32(18), pp. 2847–2849.

Harris, M. A. *et al.* (2013) ‘FYPO: the fission yeast phenotype ontology’, *Bioinformatics* , 29(13),

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

pp. 1671–1678.

Hauser, R. M., Henshall, D. C. and Lubin, F. D. (2018) 'The Epigenetics of Epilepsy and Its Progression', *The Neuroscientist*, 24(2), pp. 186–200.

Heisler, J. *et al.* (2014) 'Morphological Effects of Natural Products on *Schizosaccharomyces pombe* Measured by Imaging Flow Cytometry', *Natural Products Bioprosp*, 4(1), pp. 27–35.

Hillenmeyer, M. E. *et al.* (2008) 'The Chemical Genomic Portrait of Yeast: Uncovering a Phenotype for All Genes', *Science*, pp. 362–365. doi: 10.1126/science.1150021.

Hirota, K. *et al.* (2008) 'Stepwise chromatin remodelling by a cascade of transcription initiation of non-coding RNAs', *Nature*, 456(7218), pp. 130–134.

Hoffman, C. S. and Winston, F. (1990) 'Isolation and characterization of mutants constitutive for expression of the fbp1 gene of *Schizosaccharomyces pombe*', *Genetics*, 124(4), pp. 807–816.

Hon, C.-C. *et al.* (2017) 'An atlas of human long non-coding RNAs with accurate 5' ends', *Nature*, 543(7644), pp. 199–204.

Jeffares, D. C. *et al.* (2015) 'The genomic and phenotypic diversity of *Schizosaccharomyces pombe'*, *Nat Genet*, 47(3), pp. 235–241.

Joung, J. *et al.* (2017) 'Genome-scale activation screen identifies a lncRNA locus regulating a gene neighbourhood', *Nature*, 548(7667), pp. 343–346.

Kamrad, S., Rodríguez-López, M., *et al.* (2020) ', a python toolbox for assessing microbial growth and cell viability in high-throughput colony screens', *eLife*, 9. doi: 10.7554/eLife.55160.

Kamrad, S., Grossbach, J., *et al.* (2020) 'Pyruvate kinase variant of fission yeast tunes carbon metabolism, cell regulation, growth and stress resistance', *Mol Syst Biol*, 16(4), p. e9270.

Kamrad, S., Bähler, J. and Ralser, M. (2021) 'High-throughput, high-precision colony phenotyping with pyphe', *Research Square*. doi: 10.21203/rs.3.rs-401914/v1.

Kim, D.-U. *et al.* (2010) 'Analysis of a genome-wide set of gene deletions in the fission yeast Schizosaccharomyces pombe', *Nat Biotechnol*, 28(6), pp. 617–623.

Kim, J. *et al.* (2016) 'Long noncoding RNAs in diseases of aging', *Biochimica Biophysica Acta*, 1859(1), pp. 209–221.

Knutsen, J. H. J. *et al.* (2011) 'Cell-cycle analysis of fission yeast cells by flow cytometry', *PLoS ONE*, 6(2), p. e17175.

Kopp, F. and Mendell, J. T. (2018) 'Functional Classification and Experimental Dissection of Long Noncoding RNAs', *Cell*, 172(3), pp. 393–407.

Kumar, V. *et al.* (2013) 'Human disease-associated genetic variation impacts large intergenic

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

non-coding RNA expression’, *PLoS Genet*, 9(1), p. e1003201.

Leong, H. S. *et al.* (2014) ‘A global non-coding RNA system modulates fission yeast protein levels in response to stress’, *Nat Commun*, 5, p. 3947.

Lie, S. *et al.* (2018) ‘The contribution of non-essential *Schizosaccharomyces pombe* genes to fitness in response to altered nutrient supply and target of rapamycin activity’, *Open Biol* , p. 180015. doi: 10.1098/rsob.180015.

Lin, N. *et al.* (2014) ‘An evolutionarily conserved long noncoding RNA TUNA controls pluripotency and neural lineage commitment’, *Mol Cell*, 53(6), pp. 1005–1019.

Liu, S. J. *et al.* (2017) ‘CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells’, *Science*, 355(6320). doi: 10.1126/science.aah7111.

Lock, A. *et al.* (2019) ‘PomBase 2018: user-driven reimplementation of the fission yeast database provides rapid and intuitive access to diverse, interconnected information’, *Nucleic Acids Res*, 47(D1), pp. D821–D827.

López-Maury, L., Marguerat, S. and Bähler, J. (2008) ‘Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation’, *Nature Rev Genet*, 9(8), pp. 583–593.

Malecki, M. and Bähler, J. (2016) ‘Identifying genes required for respiratory growth of fission yeast’, *Wellcome Open Res*, 1, p. 12.

Marguerat, S. *et al.* (2012) ‘Quantitative analysis of fission yeast transcriptomes and proteomes in proliferating and quiescent cells’, *Cell*, 151(3), pp. 671–683.

Maundrell, K. (1993) ‘Thiamine-repressible expression vectors pREP and pRIP for fission yeast’, *Gene*, pp. 127–130. doi: 10.1016/0378-1119(93)90551-d.

Moreno, M. B., Durán, A. and Ribas, J. C. (2000) ‘A family of multifunctional thiamine-repressible expression vectors for fission yeast’, *Yeast* , 16(9), pp. 861–872.

Moriya, H. (2015) ‘Quantitative nature of overexpression experiments’, *Mol Biol Cell*, 26(22), pp. 3932–3939.

Morris, J. H. *et al.* (2011) ‘clusterMaker: a multi-algorithm clustering plugin for Cytoscape’, *BMC Bioinformatics*, 12, p. 436.

Nurse, P. and Hayles, J. (2019) ‘Using genetics to understand biology’, *Heredity*, pp. 4–13. doi: 10.1038/s41437-019-0209-z.

Oda, A. *et al.* (2015) ‘Dynamic transition of transcription and chromatin landscape during fission yeast adaptation to glucose starvation’, *Genes Cells*, 20(5), pp. 392–407.

Parker, S. *et al.* (2018) ‘Large-scale profiling of noncoding RNA function in yeast’, *PLoS Genet*, 14(3), p. e1007253.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Pauli, A. *et al.* (2012) ‘Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis’, *Genome Res*, 22(3), pp. 577–591.

Popadin, K. *et al.* (2013) ‘Genetic and epigenetic regulation of human lincRNA gene expression’, *Amer J Human Genet*, 93(6), pp. 1015–1026.

Prelich, G. (2012) ‘Gene overexpression: uses, mechanisms, and interpretation’, *Genetics*, 190(3), pp. 841–854.

Rallis, C. *et al.* (2014) ‘Systematic screen for mutants resistant to TORC1 inhibition in fission yeast reveals genes involved in cellular ageing and growth’, *Biol Open*, 3(2), pp. 161–171.

Rallis, C. and Bähler, J. (2016) ‘Cell-based screens and phenomics with fission yeast’, *Critical Rev Biochem Mol Biol*, pp. 86–95. doi: 10.3109/10409238.2015.1103205.

Rallis, C., Codlin, S. and Bähler, J. (2013) ‘TORC1 signaling inhibition by rapamycin and caffeine affect lifespan, global gene expression, and cell proliferation of fission yeast’, *Aging Cell*, pp. 563–573. doi: 10.1111/acel.12080.

Rhind, N. *et al.* (2011) ‘Comparative functional genomics of the fission yeasts’, *Science*, 332(6032), pp. 930–936.

Rinn, J. L. and Chang, H. Y. (2012) ‘Genome Regulation by Long Noncoding RNAs’, *Annu Rev Biochem*, pp. 145–166. doi: 10.1146/annurev-biochem-051410-092902.

Rodríguez-López, M. *et al.* (2016) ‘A CRISPR/Cas9-based method and primer design tool for seamless genome editing in fission yeast’, *Wellcome Open Res*, 1, p. 19.

Rodríguez-López, M. *et al.* (2020) ‘The GATA Transcription Factor Gaf1 Represses tRNAs, Inhibits Growth, and Extends Chronological Lifespan Downstream of Fission Yeast TORC1’, *Cell Rep*, 30(10), pp. 3240–3249.e4.

Schlackow, M. *et al.* (2017) ‘Distinctive Patterns of Transcription and RNA Processing for Human lincRNAs’, *Mol Cell*, 65(1), pp. 25–38.

Schwer, B. *et al.* (2017) ‘Defining the DNA Binding Site Recognized by the Fission Yeast Zn 2 Cys 6 Transcription Factor Pho7 and Its Role in Phosphate Homeostasis’, *mBio*. doi: 10.1128/mbio.01218-17.

Shah, S. *et al.* (2014) ‘lncRNA recruits RNAi and the exosome to dynamically regulate pho1 expression in response to phosphate levels in fission yeast’, *Genes & Dev*, 28(3), pp. 231–244.

Shannon, P. *et al.* (2003) ‘Cytoscape: a software environment for integrated models of biomolecular interaction networks’, *Genome research*, 13(11), pp. 2498–2504.

Siam, R., Dolan, W. P. and Forsburg, S. L. (2004) ‘Choosing and using *Schizosaccharomyces pombe* plasmids’, *Methods* , 33(3), pp. 189–198.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Sideri, T. *et al.* (2014) ‘Parallel profiling of fission yeast deletion mutants for proliferation and for lifespan during long-term quiescence’, *G3*, 5(1), pp. 145–155.

Tanaka, N. *et al.* (1998) ‘Isolation and characterization of an invertase and its repressor genes from Schizosaccharomyces pombe’, *Biochem Biophys Res Commun*, 245(1), pp. 246–253.

Touat-Todeschini, L. *et al.* (2017) ‘Selective termination of lncRNA transcription promotes heterochromatin silencing and cell differentiation’, *EMBO J*, 36(17), pp. 2626–2641.

Tuck, A. C. *et al.* (2018) ‘Distinctive features of lincRNA gene expression suggest widespread RNA-independent functions’, *Life Sci Alliance*, 1(4), p. e201800124.

Ulitsky, I. (2016) ‘Evolution to the rescue: using comparative genomics to understand long non-coding RNAs’, *Nature Rev Genet*, 17(10), pp. 601–614.

Ulitsky, I. and Bartel, D. P. (2013) ‘lincRNAs: genomics, evolution, and mechanisms’, *Cell*, 154(1), pp. 26–46.

Vachon, L. *et al.* (2013) ‘Functional characterization of fission yeast transcription factors by overexpression analysis’, *Genetics*, 194(4), pp. 873–884.

Van Treeck, B. *et al.* (2018) ‘RNA self-assembly contributes to stress granule formation and defining the stress granule transcriptome’, *Proc Natl Acad Sci*, 115(11), pp. 2734–2739.

Vassiliadis, D. *et al.* (2019) ‘A genome-wide analysis of carbon catabolite repression in *Schizosaccharomyces pombe*’, *BMC Genom*, 20(1), p. 251.

Virtanen, P. *et al.* (2020) ‘SciPy 1.0: fundamental algorithms for scientific computing in Python’, *Nature Methods*, 17(3), pp. 261–272.

Wagih, O. and Parts, L. (2014) ‘gitter: a robust and accurate method for quantification of colony sizes from plate images’, *G3*, 4(3), pp. 547–552.

Watanabe, Y. and Yamamoto, M. (1994) ‘*S. pombe mei2+* encodes an RNA-binding protein essential for premeiotic DNA synthesis and meiosis I, which cooperates with a novel RNA species meiRNA’, *Cell*, 78(3), pp. 487–498.

Wei, S. *et al.* (2019) ‘Systematic evaluation of *C. elegans* lincRNAs with CRISPR knockout mutants’, *Genome Biol*, 20(1), p. 7.

Wilhelm, B. T. *et al.* (2008) ‘Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution’, *Nature*, 453(7199), pp. 1239–1243.

Wood, V. *et al.* (2019) ‘Hidden in plain sight: what remains to be discovered in the eukaryotic proteome?’, *Open Biol*, 9(2), p. 180241.

Yague-Sanz, C. *et al.* (2020) ‘Nutrient-dependent control of RNA polymerase II elongation rate regulates specific gene expression programs by alternative polyadenylation’, *Genes & Dev*, 34(13-14). doi: 10.1101/gad.337212.120.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Yamashita, A. (2019) ‘meiRNA, A Polyvalent Player in Fission Yeast Meiosis’, *Non-coding RNA*, 5(3). doi: 10.3390/ncrna5030045.

Yamashita, A., Shichino, Y. and Yamamoto, M. (2016) ‘The long non-coding RNA world in yeasts’, *Biochim Biophys Acta*, 1859(1), pp. 147–154.

Yildirim, E. *et al.* (2013) ‘Xist RNA is a potent suppressor of hematologic cancer in mice’, *Cell*, 152(4), pp. 727–742.

Yoshikawa, K. *et al.* (2011) ‘Comprehensive phenotypic analysis of single-gene deletion and overexpression strains of Saccharomyces cerevisiae’, *Yeast* , 28(5), pp. 349–361.

Zhou, Y. *et al.* (2015) ‘The fission yeast MTREC complex targets CUTs and unspliced pre-mRNAs to the nuclear exosome’, *Nat Commun*, 6, p. 7050.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

FIGURE LEGENDS

Figure 1: Phenomics assays of lincRNA mutants.

(A) Representation of all non-coding RNAs across the three *S. pombe* chromosomes (Atkinson *et al.*, 2018). LincRNAs analysed in this study are coloured in red (if they showed no phenotypes) or purple (if they showed phenotypes in at least one condition when deleted).

(B) Schematic overview of experimental design and workflow for phenotyping and data analyses.

(C) Colony size (growth) and redness (viability) provide orthogonal readouts with strong biological signals. These two readouts are not correlated (r_{Pearson}= -0.022). Both methods are highly reproducible with overall coefficients of variation of 0.050 and 0.007 for size and redness, respectively (based on 3514 wild-type control colonies across all plates). The lower relative distribution spreads of control values (wild-type *vs* entire dataset) indicates a strong biological signal. Fractions of unexplained variance were 0.56 for size and 0.40 for redness.

Figure 2. Colony growth and viability of deletion mutants in benign conditions.

(A) Volcano plot for colony size of lincRNA mutants (green) and coding-gene mutants (grey) growing in rich medium (top graph) and minimal medium (bottom graph). The dashed lines show the significance thresholds. Strains with lower fitness (smaller colonies) are <0 on the x-axis, and those with higher fitness are >0. We applied a significance threshold of 0.05 after Benjamini-Hochberg correction for multiple testing and a difference in fitness of abs(log2(mutant/wild type))>log2(0.05) to call hits based on colony size; this difference is similar to the median coefficient of variation (CV).

(B) Volcano plot for colony viability (phloxine B redness score) of lincRNA mutants (green) and coding-gene mutants (grey) growing in rich medium (top graph) and minimal medium (bottom graph). The dashed lines show the significance thresholds. Strains showing lower fitness (redder colonies) are above zero on the x-axis, and those with higher fitness are below zero. We determined quantitative redness scores and applied a significance threshold of 0.05 after Benjamini-Hochberg correction for multiple testing and an effect size threshold of abs($\log_2$(mutant/wild-type))>$\log_2$(0.015) to identify colonies that are more or less red than wild-type colonies. The labels indicate the identity of the significant lincRNA genes.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 3. Cell-size and cell-cycle traits of lincRNA mutants.

(A) LincRNA deletion mutants showing ≥5% difference in median cell size ($p_{\mathrm{Wilcoxon}} < 0.05$), compared to wild-type (wt) cells and the conditional cell-size mutants *wee1-50* and *cdc10-129*, captured at 60 min after release to permissive temperature. The sizes of bi-nucleated cells were measured in 63 microscope fields using high-throughput microscopy.

(B) Representative cells from (A), with bi-nucleated cells in red.

(C) Plot of cell growth *vs* cell length of bi-nucleated cells for all lincRNA mutants analyzed here. The data on log2 growth of mutant relative to wild-type cells in rich medium are from the colony-based screen (Figure 2A). The length data of bi-nucleated cells grown in rich medium are from the high-throughput microscopy (A).

(D) lincRNA deletion mutants showing ≥20% difference in percentage of bi-nucleated cells ($p_{\mathrm{Wilcoxon}} < 0.05$) compared to wt cells as in (A). The median proportion of bi-nucleated cells was quantified from the proportion of bi-nucleated cells in each microscope field, captured for each lincRNA mutant using high-throughput microscopy.

Figure 4. Colony growth and viability of deletion mutants in diverse conditions.

(A) Distributions of significant hits per mutant (left) or per condition (right) for lincRNA mutants with altered colony growth (blue) or viability (orange).

(B) Plot showing the number of growth-phenotype hits agreeing or disagreeing between independently generated lincRNA mutants.

(C) Distributions of significant hits per mutant (left) or per condition (right) for coding-gene mutants with altered colony growth (blue) or viability (orange).

(D) Top Venn diagram: numbers of lincRNA mutants that showed a phenotype for both colony growth (big or small) and viability (red or white, with red colonies having lower viability) in 67 conditions. Bottom Venn diagram: numbers of coding-gene mutants showing a phenotype for both colony growth and viability in 67 conditions.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 5. Functional signatures in lincRNA phenotype profiles.

(A) Hierarchical clustering of discretized data for 16 lincRNA mutants (green) and 178 coding-gene mutants (grey), as indicated above the columns. Clustering was limited to the core conditions where phenotypes for all mutants were available, including growth phenotypes (brown) and viability phenotypes (purple), as indicated to the right of rows. Only mutants with at least five hits across the 41 conditions are shown. Resistant (dark red) and sensitive (dark blue) phenotypes are indicated for corresponding mutant-condition combinations. Hierarchical clustering of both mutants and conditions was performed with the Ward method using Euclidean distances. Based on the dendrogram, the genes were divided into 3 clusters indicated in different colours (top row). A detailed version of this cluster specifying the conditions and mutants is provided in Supplemental Figure 6C.

(B) Cytoscape gene network representing phenotype correlations between lincRNA and coding-gene mutants. Yellow and blue edges show positive and negative phenotype correlations, respectively. The lincRNAs are shown in green and the protein-coding genes in grey, including a pink border if their function is unknown. Clusters discussed in the main text are highlighted in color.

Figure 6. Growth phenotypes of lincRNA overexpression mutants in different conditions.

(A) Distributions of significant phenotype hits per strain (left) and condition (right) for lincRNA overexpression strains with altered growth under benign and stress conditions. Overall, 113 overexpression strains were phenotyped under 47 different conditions, based on 31 distinct environmental factors. We applied a significance threshold of $p \leq 0.01$, after correction for multiple testing, and a difference in fitness of ≥5% to call hits based on colony size.

(B) Hierarchical clustering of discretized relative log2 median effect sizes for lincRNA overexpression strains (rows), using only the strains and conditions with at least 5 hits (59 mutants, 29 conditions). Resistant (red) and sensitive (blue) phenotypes are indicated for strain-condition combinations. The sub-cluster highlighted in red is discussed in the main text.

(C) Comparison of phenotype data from lincRNA deletion vs overexpression mutants. Plot showing maximum median effect sizes for 104 lincRNA mutants represented in both deletion and overexpression libraries, phenotyped under 22 shared conditions. The pairwise Pearson correlation coefficient is indicated. To aid visualization, ten extreme outliers were removed out of 2288 data points.

available under aCC-BY 4.0 International license.

Figure 1

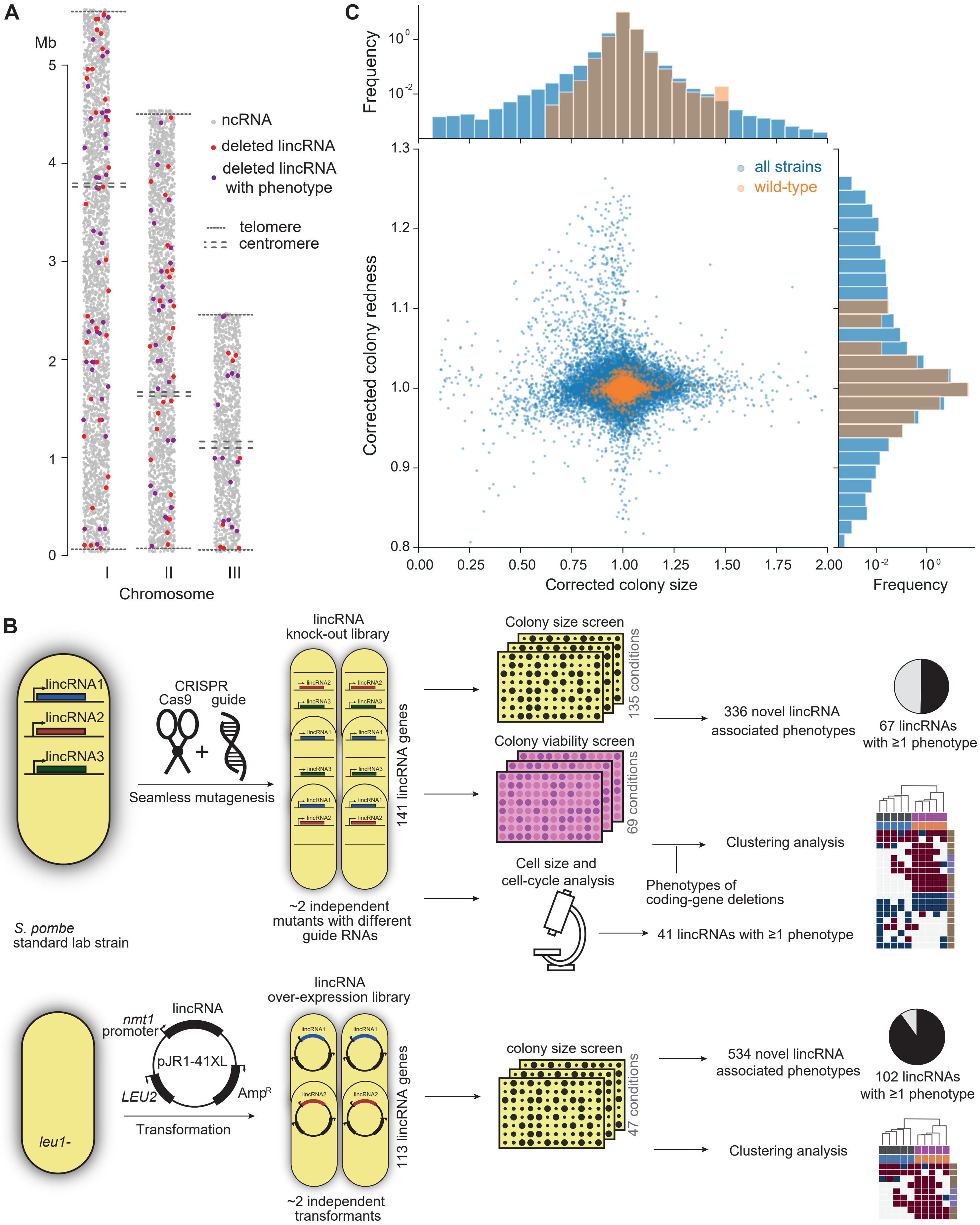


Figure 2

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

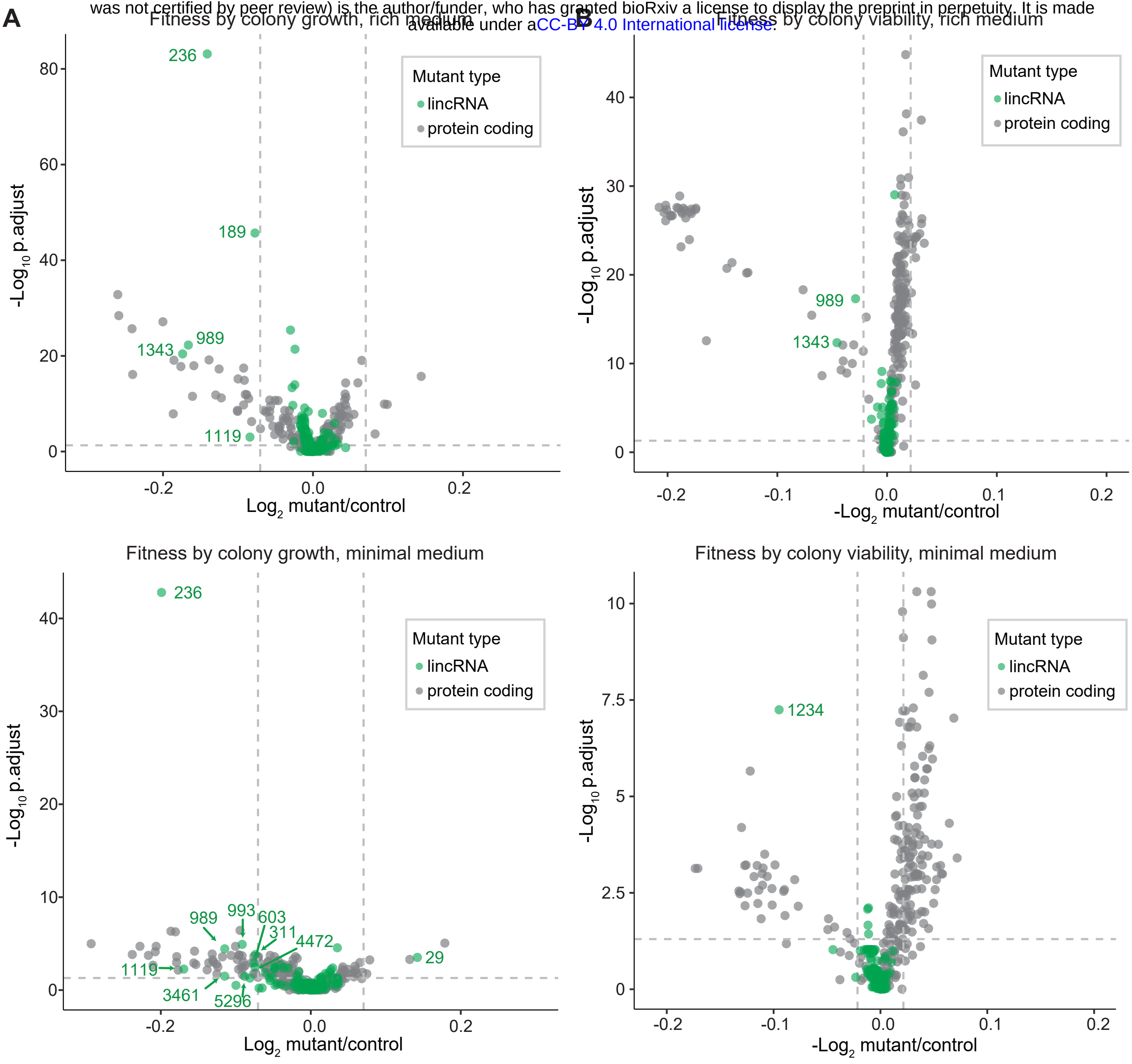


Figure 3

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

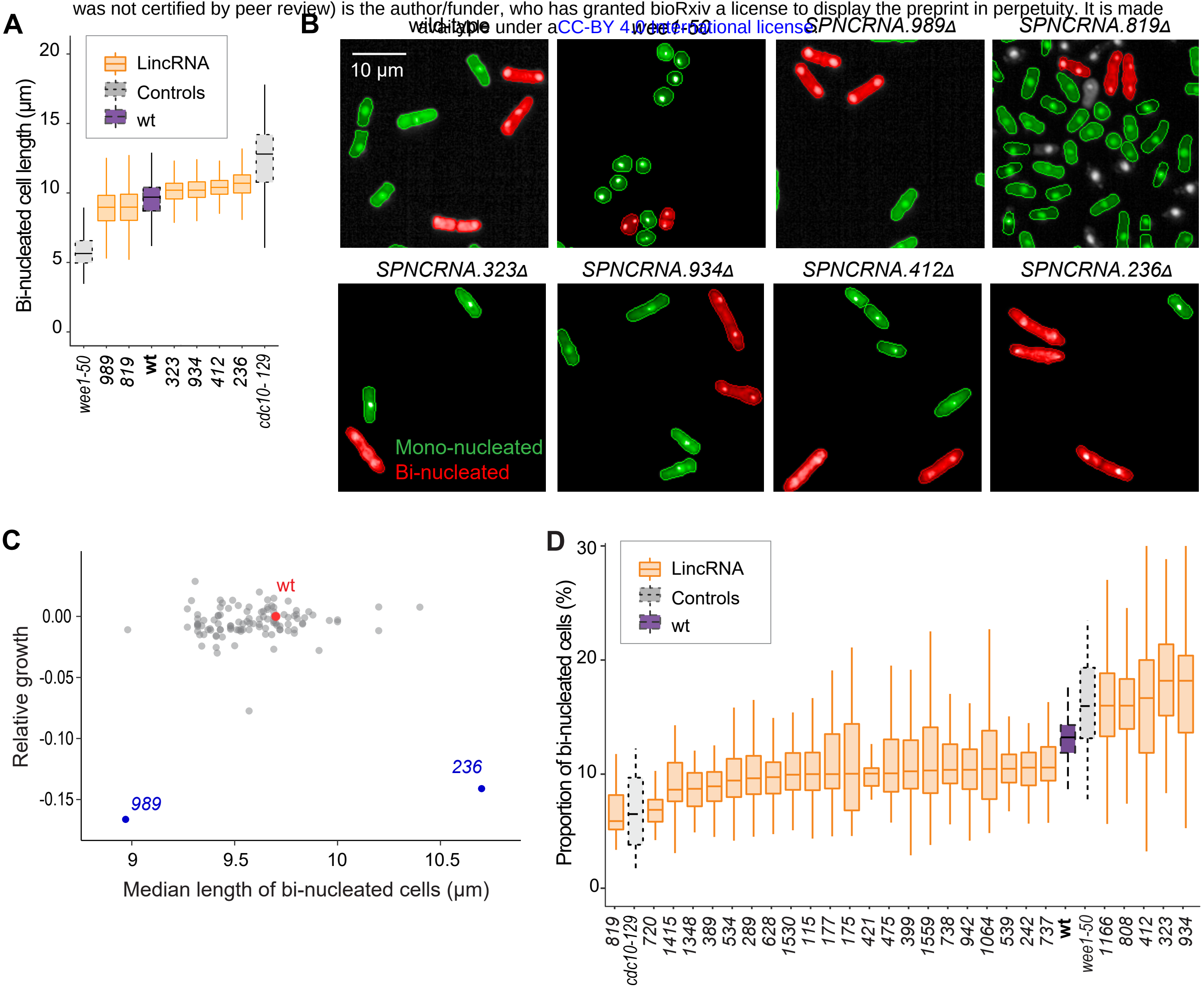


Figure 4

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

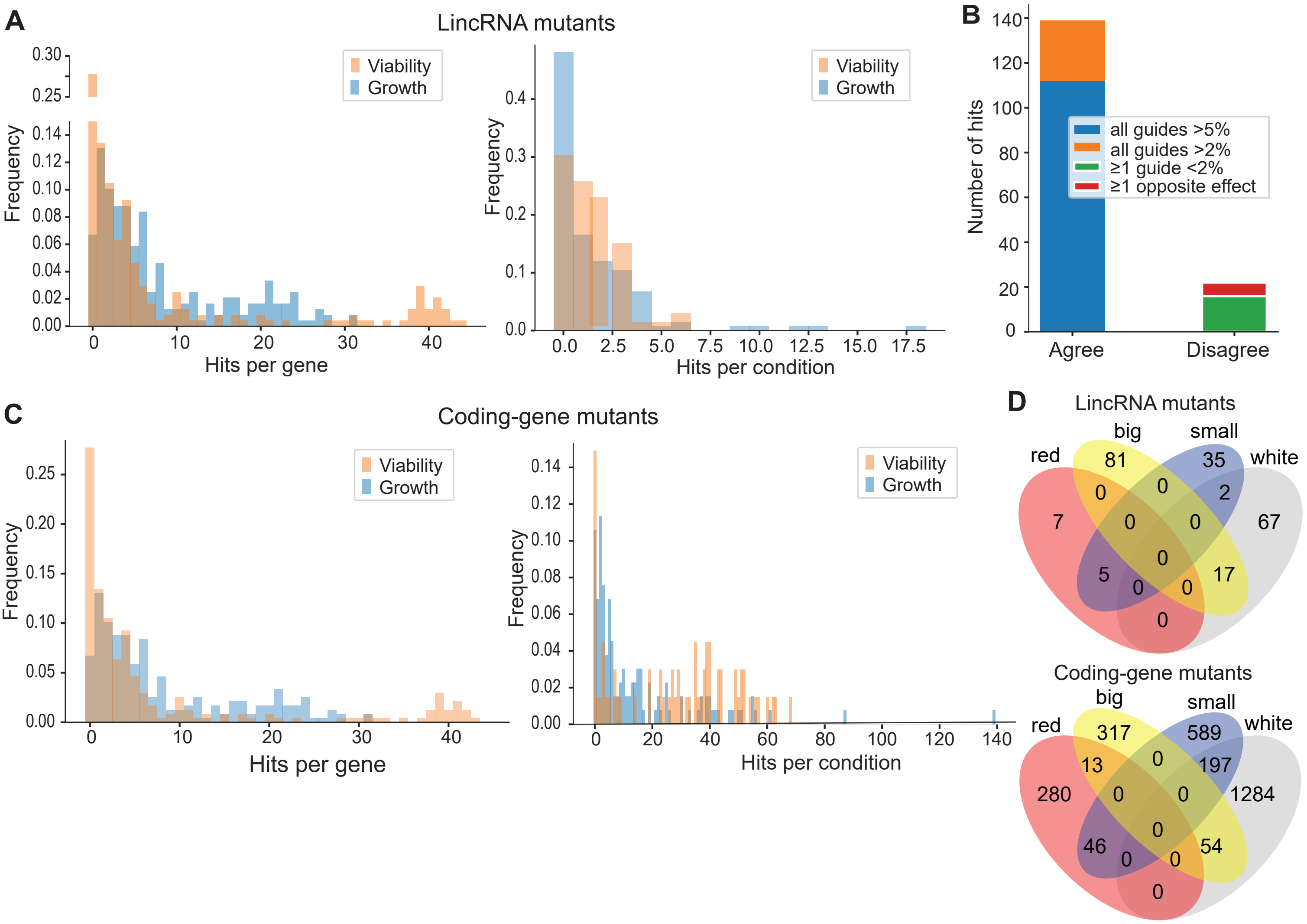


Figure 5

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

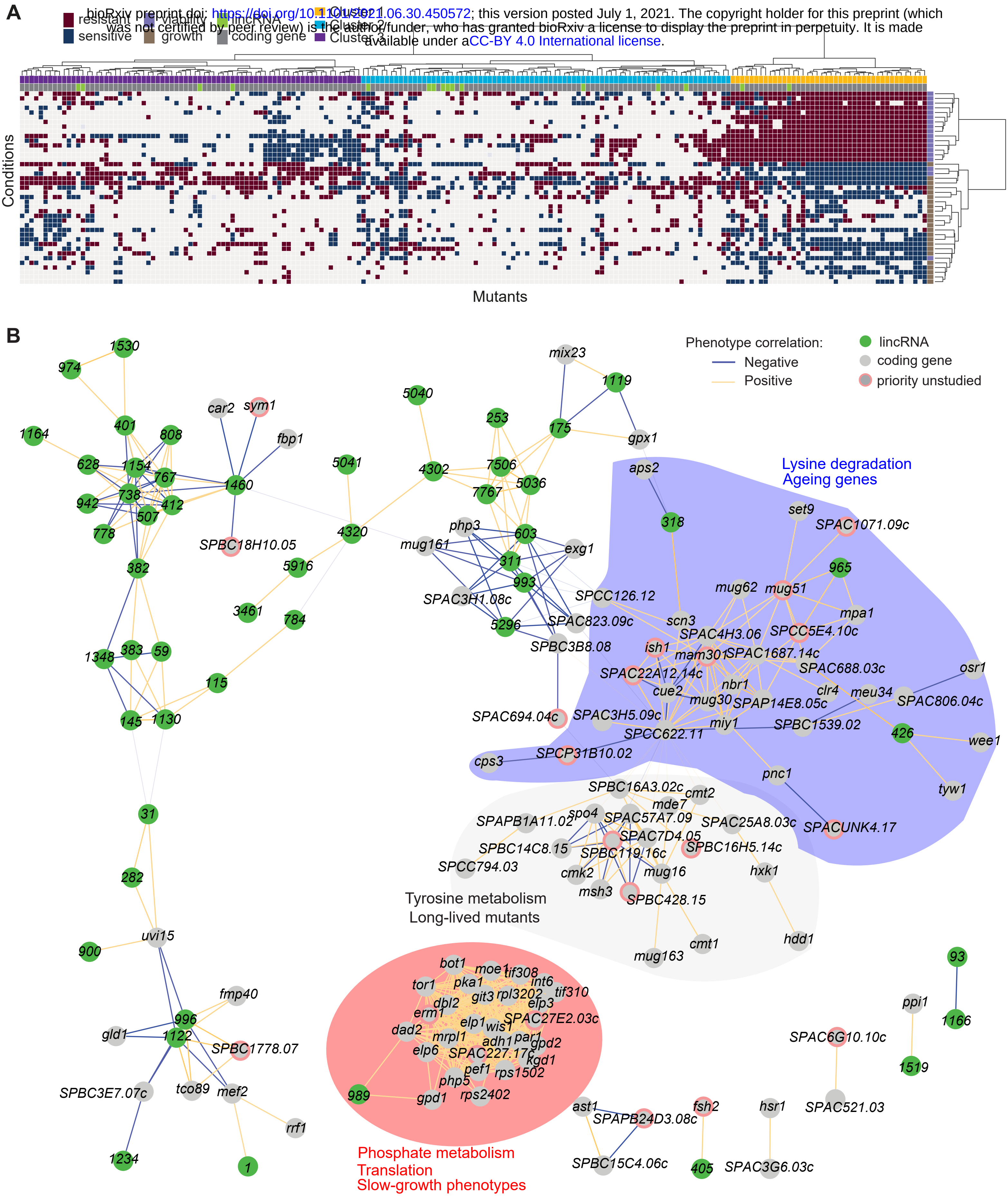
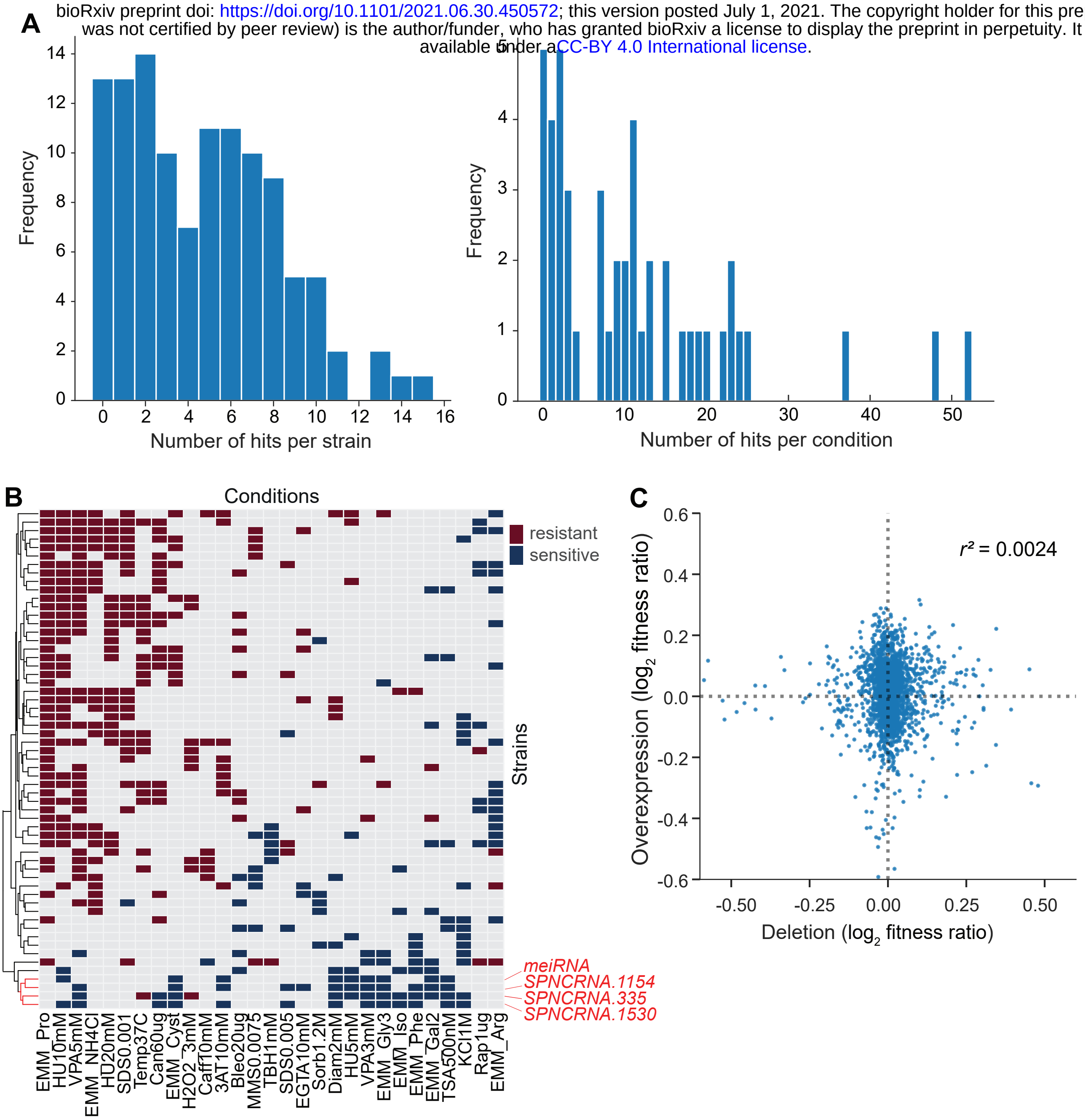


Figure 6

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

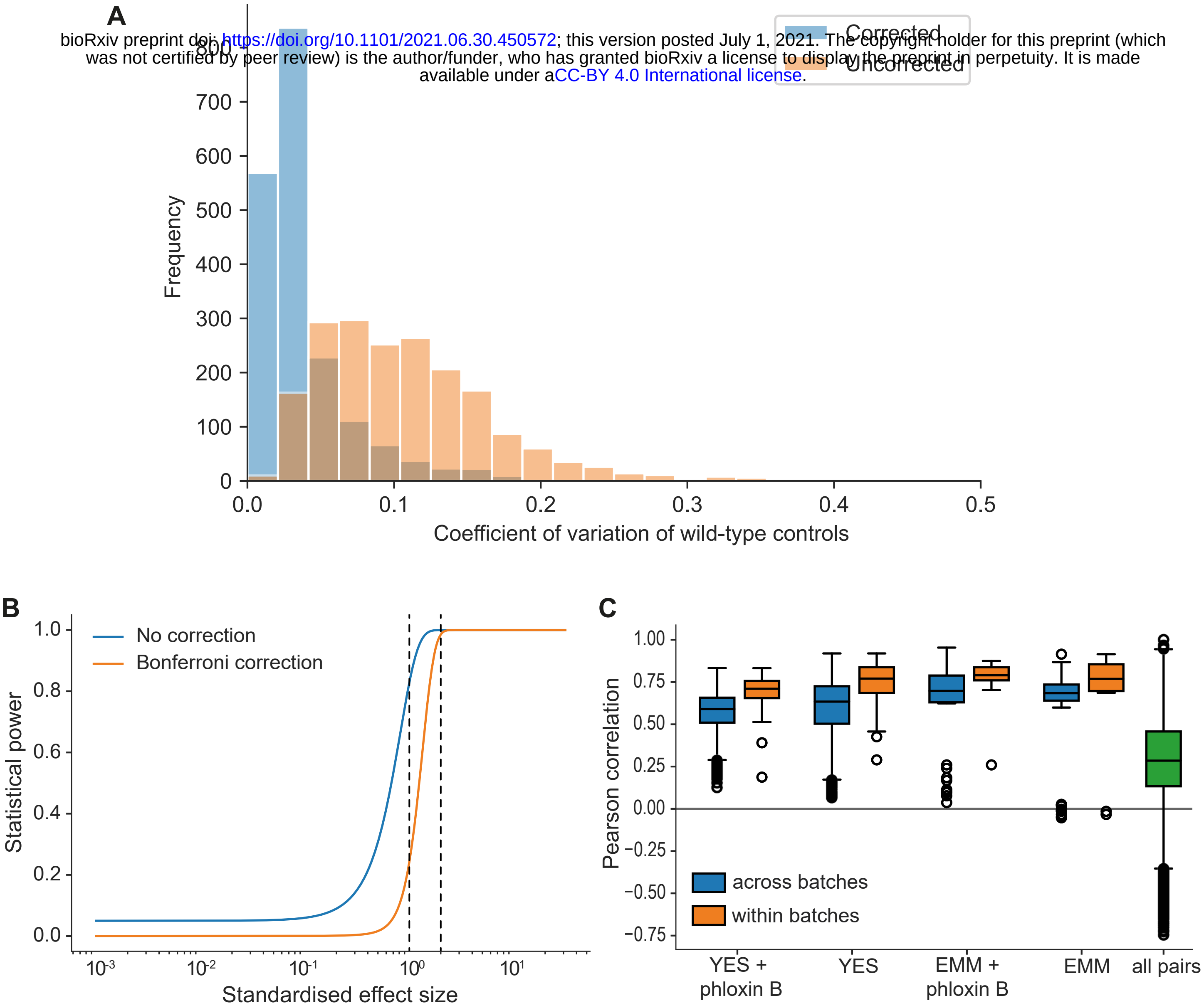


bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Media	**Heavy metals**
YES	Cobalt(II) chloride
EMM	Cadmium chloride
Temperature	**Protein damage/stress/ synthesis/ protease inhibition**
32°C	Tunicamycin
25°C	Cycloheximide
Ageing	Brefeldin A
1 day	Benzamidine
7 days	Bortezomib
Nutrients	**DNA damage**
Canavanine A (arginine starvation)	Bleomycin
3-Amino-1,2,4-triazole (3-AT) for histidine starvation	UV light
Proline as nitrogen source	Methyl methanesulfonate (MMS)
Arginine as nitrogen source	**Nucleic acid metabolism**
glutamine as nitrogen source	Hydroxyurea (HU)
Glycerol as carbon source	Formamide
Galactose as carbon source	**HDAC inhibitors**
Xylose as carbon source	Valproic Acid
Sorbitol as carbon source	Dihydrocoumarin
0.1% glucose as carbon source	Trichostatin A
0.5% glucose as carbon source	**TOR-signalling**
1% glucose as carbon source	Torin1
3% glucose as carbon source	Rapamycin
Chelating Agents	**Other drugs**
EDTA	Caffeine
EGTA	Cysteine
2.,2'-dipyridyl (iron chelator)	Methionine
Osmotic Stress	Sodium orthovanadate
Sodium Chloride	Ethanol
Potassium Chloride	Tea tree oil
Lithium Chloride	Phloxine B
Cell wall/ membrane integrity Stress	Antimycin A
SDS	**Combinations of chemicals**
Calcofluor	KCl + SDS
Oxidative stress	NaCl +MMS
Tert-Butyl hydroperoxide	KCl + calcofluor
Hydrogen peroxide	Glycerol + MMS
Diamide	HU +MMS

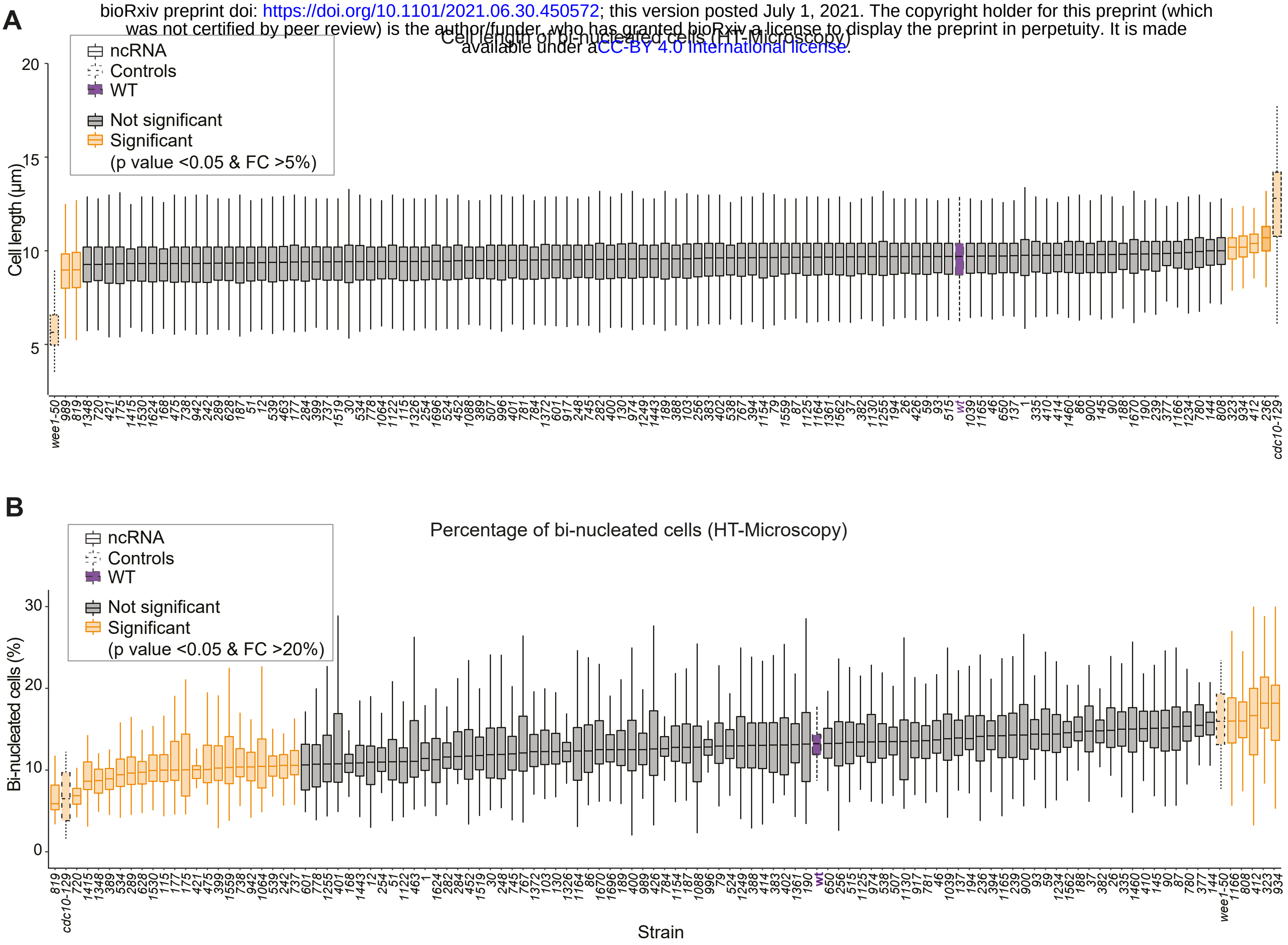
Supplemental Figure 1: Overview of conditions used for phenotyping of lincRNA knock-out library.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



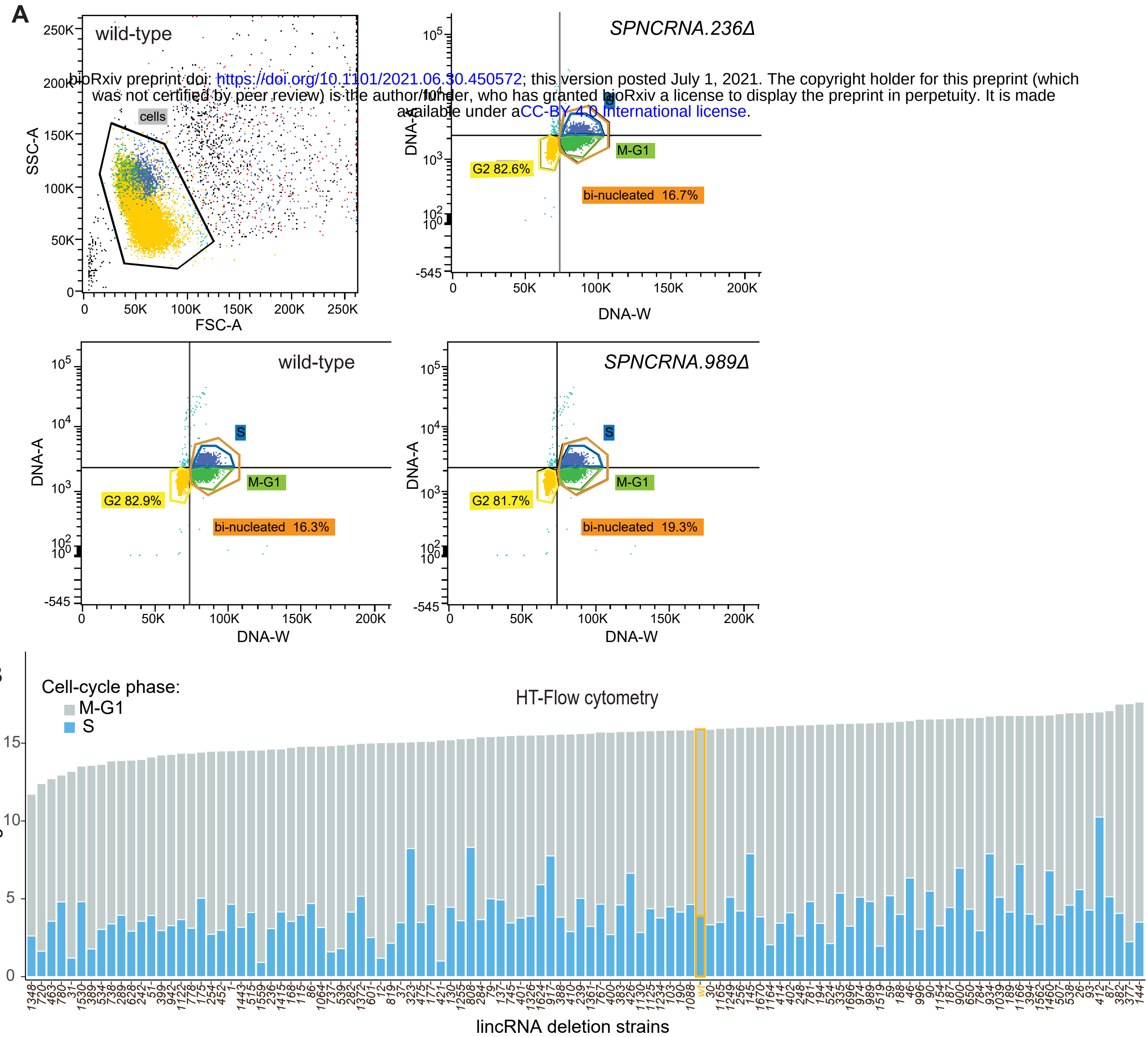
Supplemental Figure 2:
(A) Grid corrections reduce noise in the data. Shown are distributions of coefficients of variation of internal wild-type controls for 1906 plates across the data set (coding and non-coding deletion mutants) before and after correction. The median CV before and after correction are 0.098 and 0.027, respectively. So our normalisation approach typically reduces the noise by about 4-fold.
(B) Plot of the statistical power (1 - chance of non-rejection of wrong null-hypothesis) vs standardised effect size (difference in means divided by standard deviation) using median number of replicates per lincRNA gene and condition (n = 9) and number of replicates for control conditions (rich medium, median = 173.5). The large number of replicates enables statistical detection of differential growth with low false negative rate (type II error rate). The two curves show the statistical power before correction for multiple testing and after Bonferroni correction (which represents a worst case scenario). A 5% difference in fitness, which is a very subtle effect and approximates the standard deviation of our method, is detected with a chance of 83% (no correction) or 24% (Bonferroni correction) (left dashed line). A stronger 10% difference in fitness is detected with ~100% power (no correction) or 98.5% power (Bonferroni) (right dashed line). Two-sided Student's t-tests and a standard deviation of 5% were used throughput with a significance threshold of 0.05.
(C) Four control conditions (rich and minimal media, with or without phloxine B each) were included in most of the ~30 batches acquired over 2 years. Boxplots of Pearson correlations for technical repeats within one batch (orange) vs repeats of the same condition across all batches (blue). While correlation within the same batch is consistently higher, this effect is small considering the biological signal as illustrated by the distribution of pairwise correlations across all conditions (green).

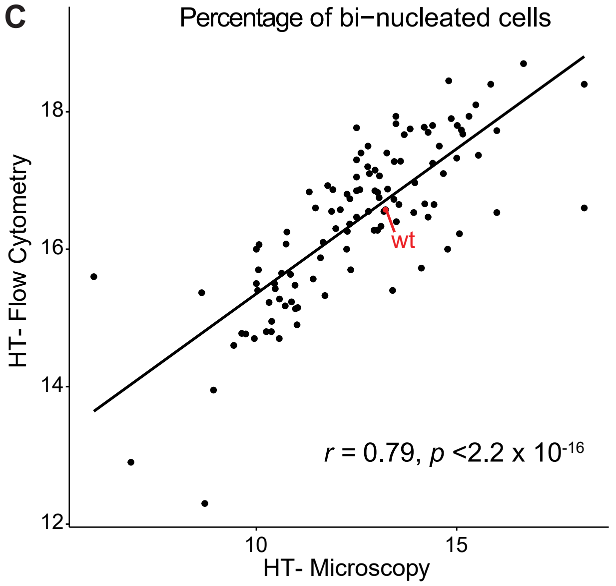
bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



Supplemental Figure 3:
(A) Lengths of bi-nucleated cells growing exponentially in rich medium for different lincRNA deletion mutants.
(B) Proportion of bi-nucleated cells in the population under the same conditions as A.

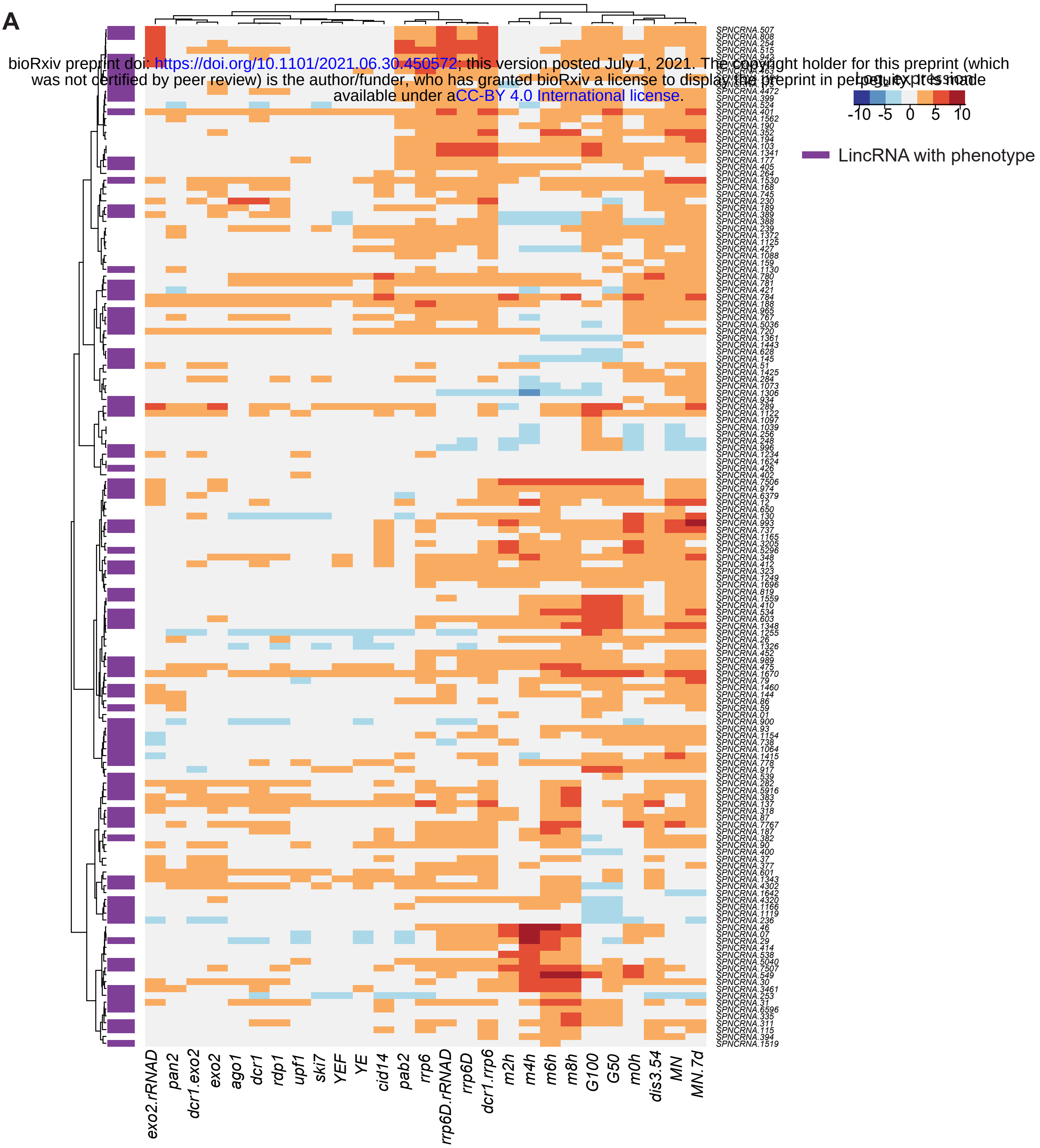
bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

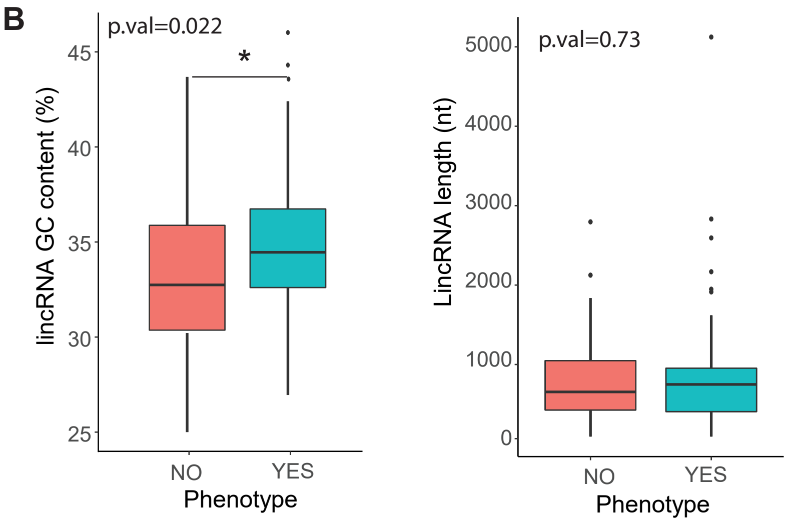




Supplemental Figure 4:
(A) Procedure to identify cell-cycle stages using high-throughput flow cytometry data. Cell doublets are first excluded (top left) and different stages assigned according to the DNA amount (DNA-A) and duration of pulse (DNA-W).
(B) Percentage of cells in different cell-cycle stages according to high-throughput flow cytometry. These data are only semi-quantitative, however, and the proportions of cell-cycle phases are only approximate.
(C) Comparison of percentage of binucleated cells calculated using high-throuput flow cytometry or high-thoughput microscopy. The Pearson correlation coefficient is indiated.

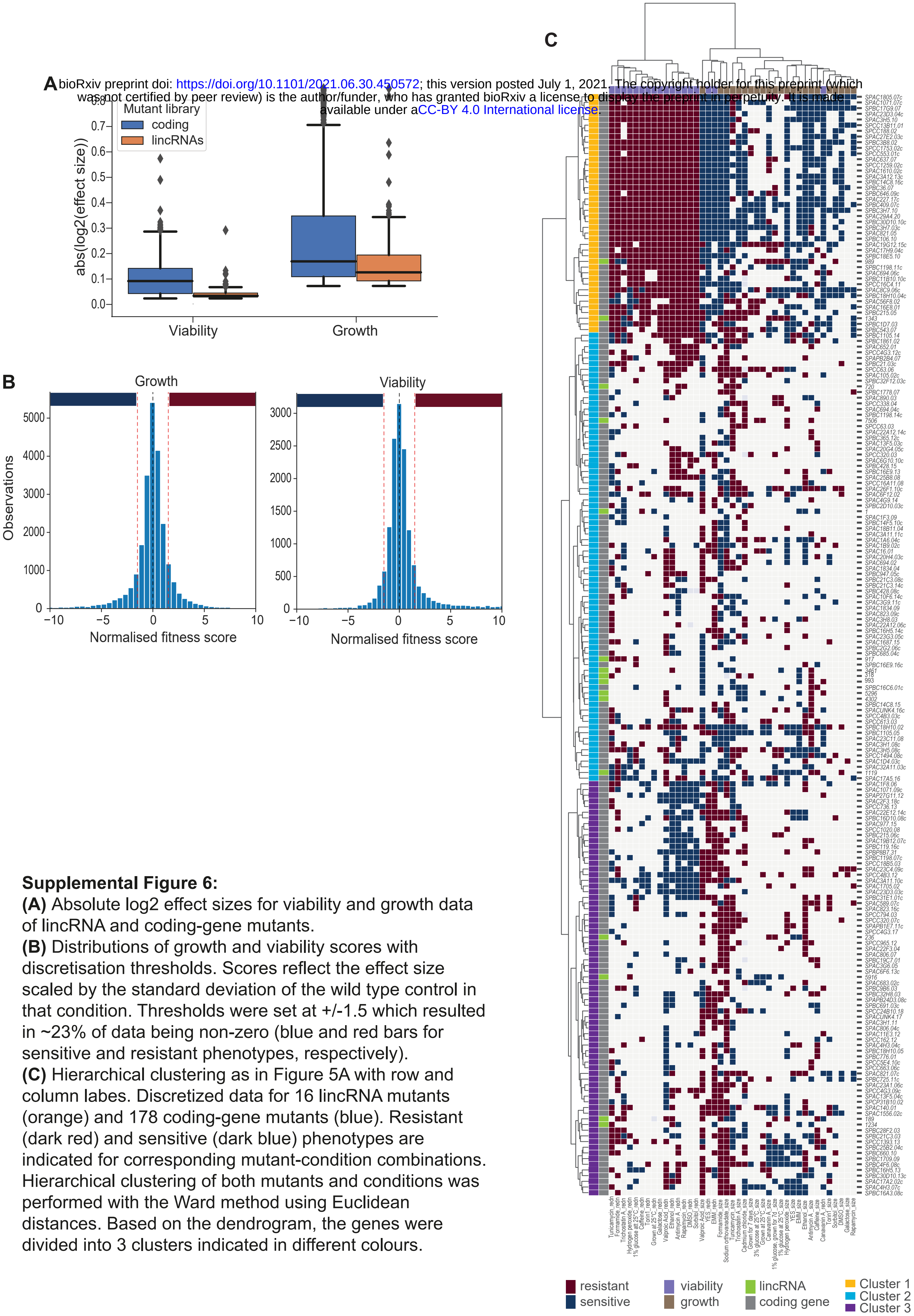
bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.





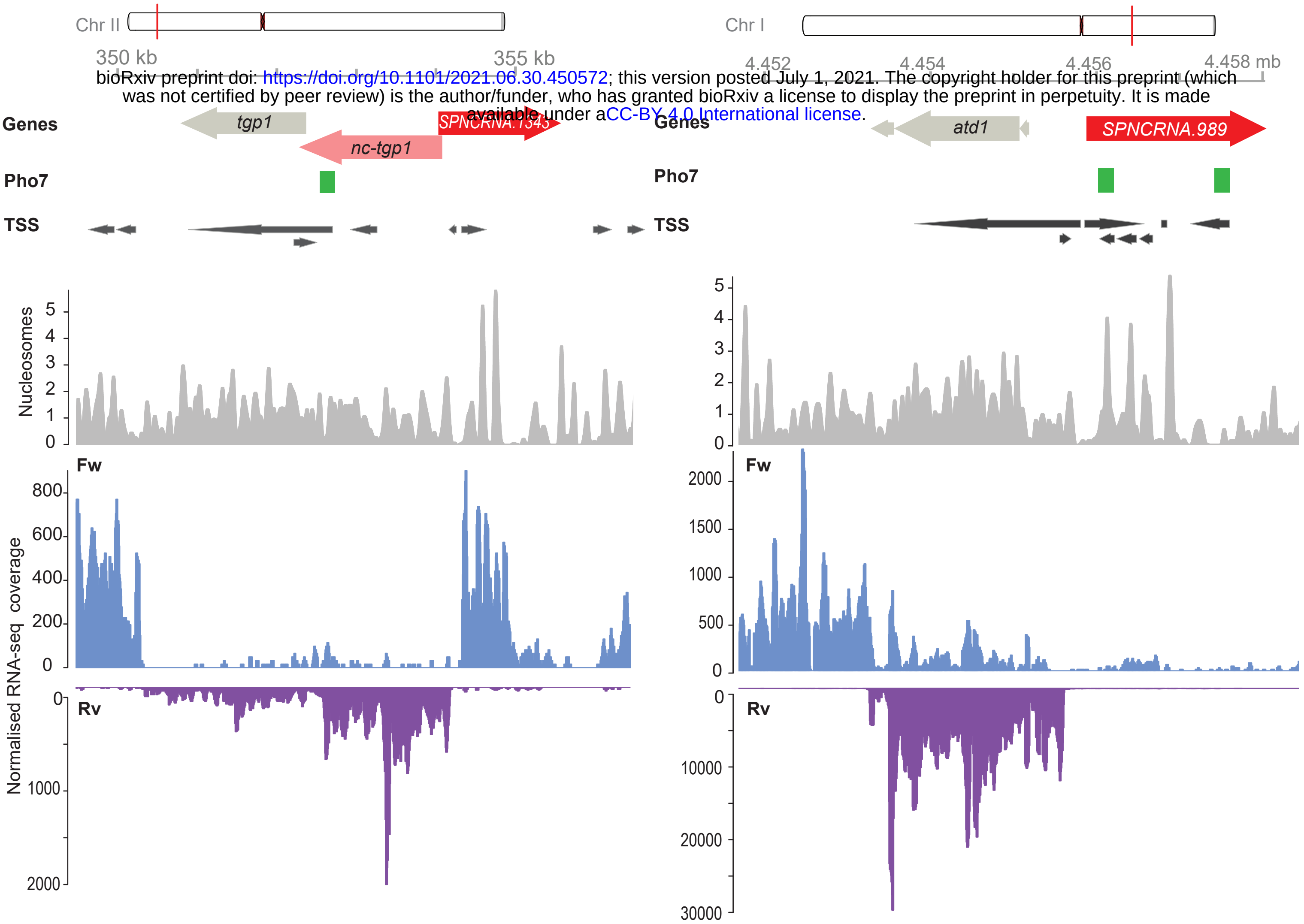
Supplemental Figure 5:
(A) Expression patterns of lincRNAs in this study (data from Atkinson et al, 2018). LincRNA that display any phenotype when deleted are marked in purple.
(B) Comparison of GC content (left) and length (right) for linRNAs that display phenotype when deleted (aquamarine) vs those that do not (coral). The p-values for significance of difference are indicated (Wilcoxon test).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



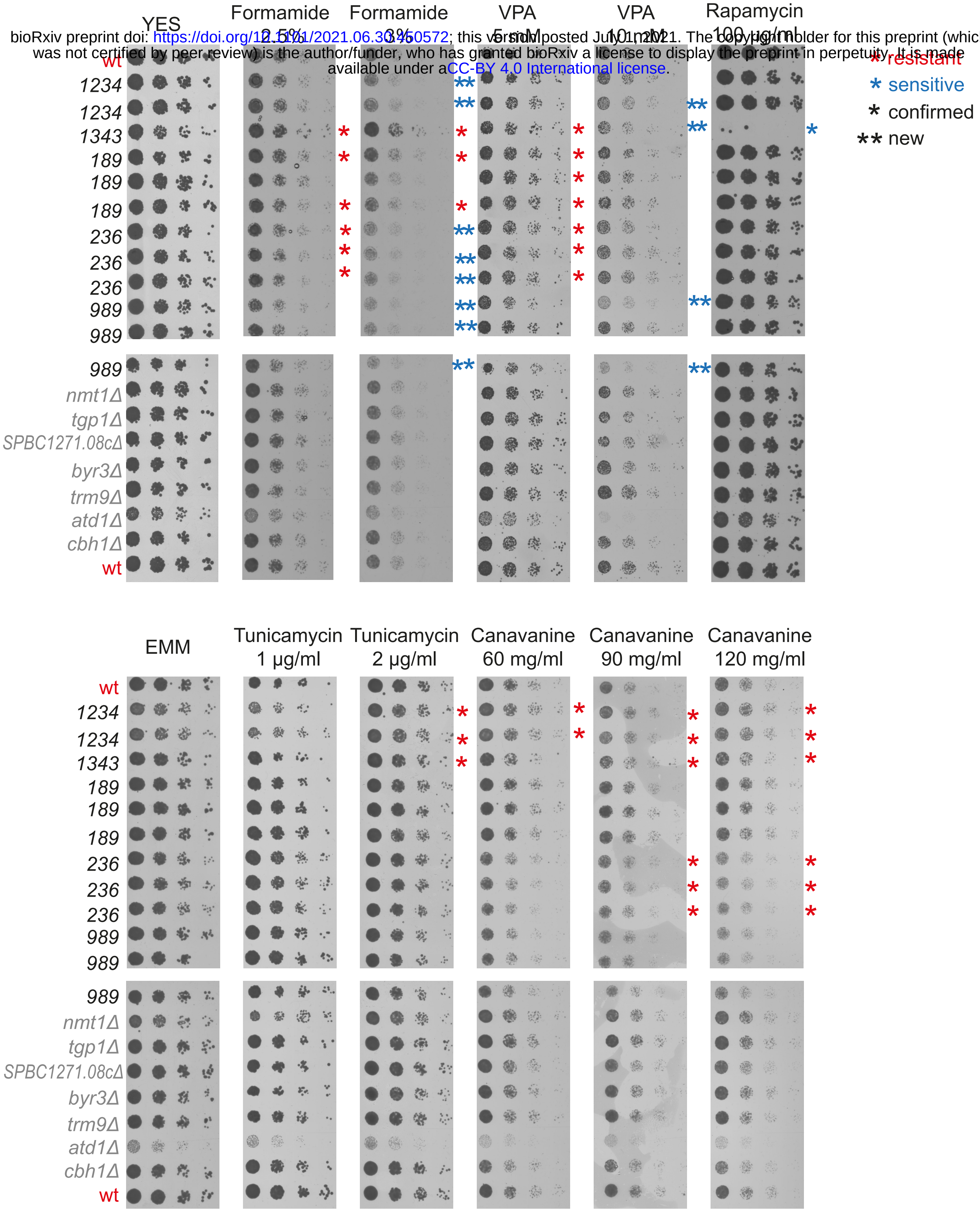
Supplemental Figure 6:
(A) Absolute log2 effect sizes for viability and growth data of lincRNA and coding-gene mutants.
(B) Distributions of growth and viability scores with discretisation thresholds. Scores reflect the effect size scaled by the standard deviation of the wild type control in that condition. Thresholds were set at +/-1.5 which resulted in ~23% of data being non-zero (blue and red bars for sensitive and resistant phenotypes, respectively).
(C) Hierarchical clustering as in Figure 5A with row and column labes. Discretized data for 16 lincRNA mutants (orange) and 178 coding-gene mutants (blue). Resistant (dark red) and sensitive (dark blue) phenotypes are indicated for corresponding mutant-condition combinations. Hierarchical clustering of both mutants and conditions was performed with the Ward method using Euclidean distances. Based on the dendrogram, the genes were divided into 3 clusters indicated in different colours.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



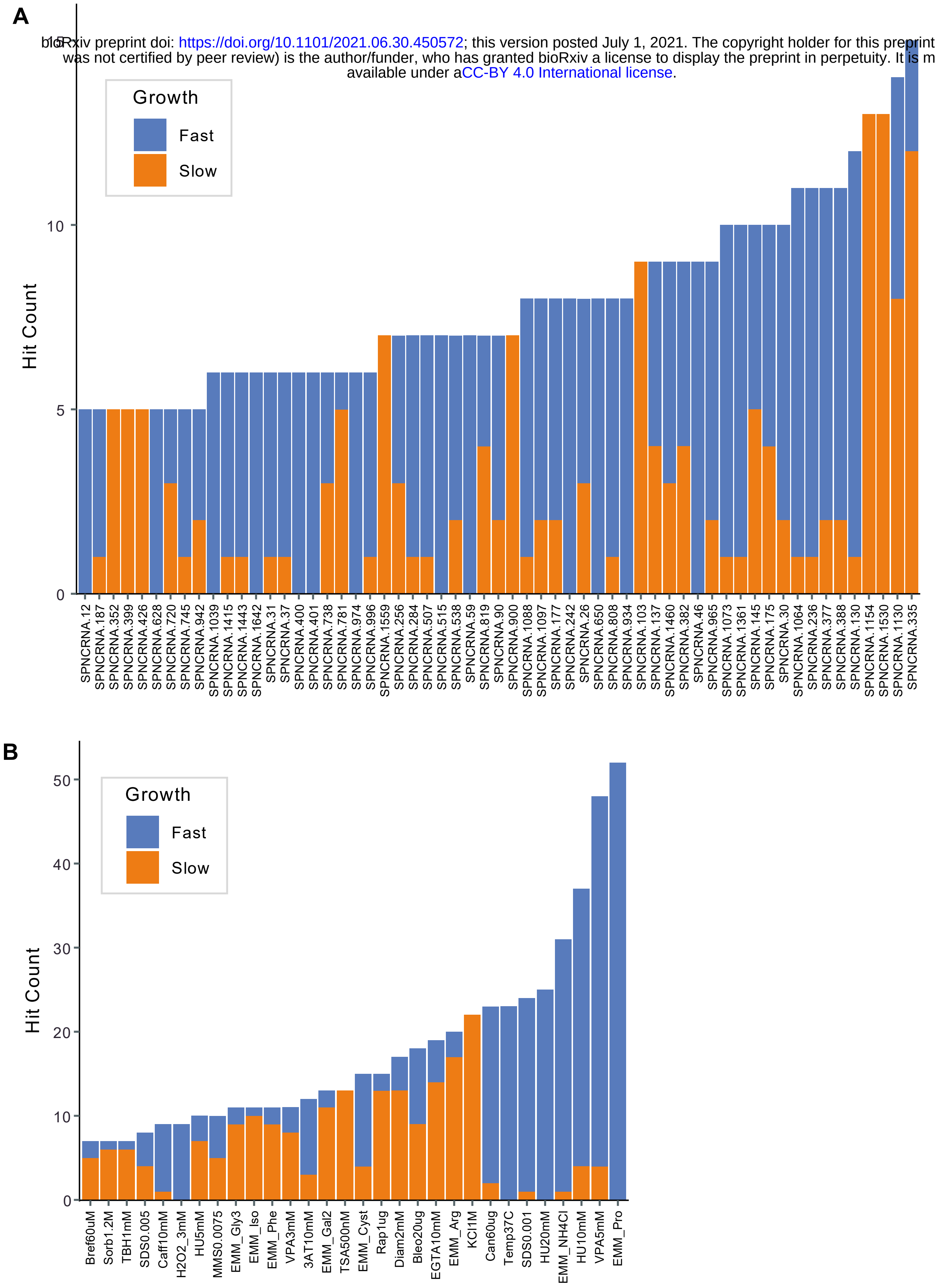
Supplemental Figure 7: Genome browser view of the chomosomal regions surrounding *SPNCRNA.1343* (left) and *SPNCRNA.989* (right). The Pho7 binding sites as described by Schwer et al. 2017, the transcription start sites (TSS) as described by Thodberg et al. 2018, and the nucleosome positioning data as described by Atkinson et al. 2018. RNA-sequencing normalised coverage is separeted in forward (Fw) and reverese (Rv) strands (Maria Rodriguez-Lopez, unpublished data).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



Supplemental Figure 8:
Spot assays with 5-fold serial dilutions to validate selected lincRNA deletion phenotypes from the screen. Deletion mutants of lincRNAs (black) and neighbouring protein-coding genes (grey) grown in YES were serially diluted and plated onto YES (top) or EMM (bottom) plates containing the drugs indicated. Single asterisks indicate validated lincRNA phenotypes, while double asterisks indicate subtle phenotypes in the higher doses of formamide and VPA not detected in the doses used for the screen, suggesting dosage effects in the serial dilution assays.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.30.450572; this version posted July 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



Supplemental Figure 9: Hit counts per strain **(A)** or per condition **(B)** for lincRNA overexpression strains with altered colony growth separated by faster and slower colony growth phenotypes relative to empty vector control (evc). Data for slower and faster growth are stacked. Adjusted p. value ≤0.01 and difference in fitness ≥5% were used as significant thresholds to call hits. Only genes or conditions which produced ≥5 hits are shown.