bioRxiv preprint doi: https://doi.org/10.1101/082735; this version posted November 24, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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-NC-ND 4.0 International license.

Characterization of Dependencies Between Growth and Division in Budding Yeast

Michael B. Mayhew*
Computational Engineering Division, Lawrence Livermore National Laboratory, Livermore, California USA

Edwin S. Iversen
Department of Statistical Science, Duke University, Durham, North Carolina USA

Alexander J. Hartemink*
Department of Computer Science and
Program in Computational Biology & Bioinformatics,
Duke University, Durham, North Carolina USA

Cell growth and division are processes vital to the proliferation and development of life. Coordination between these two processes has been recognized for decades in a variety of organisms. In the budding yeast *Saccharomyces cerevisiae*, this coordination or 'size control' appears as an inverse correlation between cell size and the rate of cell-cycle progression, routinely observed in G_1 prior to cell division commitment. Beyond this point, cells are presumed to complete $S/G_2/M$ at similar rates and in a size-independent manner. As such, studies of dependence between growth and division have focused on G_1. Moreover, coordination between growth and division has commonly been analyzed *within* the cycle of a single cell without accounting for correlations in growth and division characteristics *between* cycles of related cells. In a comprehensive analysis of three published time-lapse microscopy datasets, we analyze both intra- and inter-cycle dependencies between growth and division, revisiting assumptions about the coordination between these two processes. Interestingly, we find evidence (1) that $S/G_2/M$ durations are systematically longer in daughters than in mothers, (2) of dependencies between $S/G_2/M$ and size at budding that echo the classical G_1 dependencies, and, (3) in contrast with recent bacterial studies, of negative dependencies between size at birth and size accumulated during the cell cycle. In addition, we develop a novel hierarchical model to uncover inter-cycle dependencies, and we find evidence for such dependencies in cells growing in sugar-poor environments. Our analysis highlights the need for experimentalists and modelers to account for new sources of cell-to-cell variation in growth and division, and our model provides a formal statistical framework for the continued study of dependencies between biological processes.

I. INTRODUCTION

Cell division is a process fundamental to all life, and dysregulation of the process is common in diseases like cancer. Since it underlies so many biological phenomena, cell division is highly coordinated with other cellular processes. For instance, in the budding yeast *Saccharomyces cerevisiae*, cell division is known to be coordinated with cell growth [1–6] (reviewed in [7]). This dependence between growth and division is most noticeable in daughter cells that, owing to the asymmetric manner of budding yeast division, are born smaller than their mothers (Figure 1). Consequently, daughters undergo longer G_1 phases to reach a 'critical size' at START, the point of cell-cycle commitment [8, 9]. Generally, a correlation has been observed between the birth mass of a cell and its time spent in G_1, with smaller cells at birth taking longer to complete G_1. This 'size control' is important for the maintenance of a consistent size distribution in the cell population from generation to generation (size homeostasis).

Studies of coordination between growth and division in budding yeast have focused primarily on G_1. Indeed, after having reached a sufficient and roughly similar size, both mother and daughter cells are presumed to proceed through the S, G_2, and M phases at similar rates [9, 10]. These tenets of the traditional model of budding yeast size control set certain expectations for $S/G_2/M$ duration: (1) G_1 duration is size-dependent, while $S/G_2/M$ duration is size-independent, and (2) $S/G_2/M$ duration is roughly constant from cell to cell. Assuming a consistent single-cell growth rate across cells, these tenets are sufficient to maintain a consistent size distribution in the population. Recent studies in bacteria have revealed an alternative size control model by which cells add a relatively constant amount of volume over the cell cycle, regardless of their birth size [11–13]. Recent time-lapse microscopy datasets tracking characteristics of individual cells provide evidence with which to test these models and further characterize coordination between

*Corresponding author

bioRxiv preprint doi: https://doi.org/10.1101/082735; this version posted November 24, 2016. 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-NC-ND 4.0 International license.

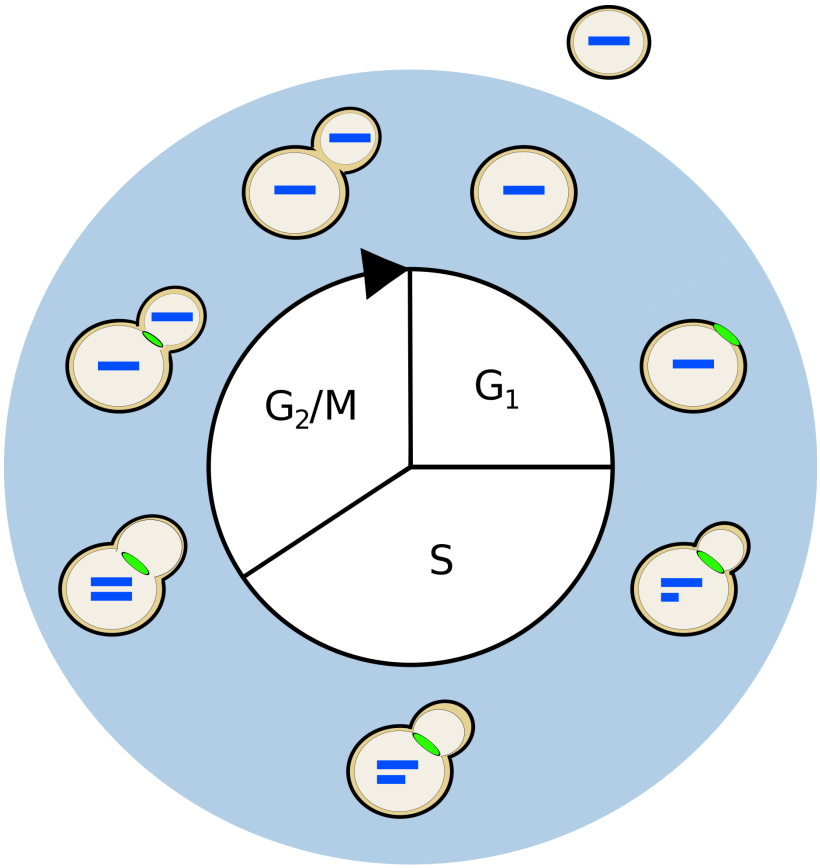

FIG. 1: A diagram of the haploid budding yeast cell cycle. The process of cell division begins at the top of the diagram and proceeds clockwise. Cell division consists of four phases: G_1, S, G_2, and M; the latter two have been merged since classical studies demonstrate that they largely overlap in yeast. Along the outer ring of the diagram are depicted progressive stages of division, as reflected by the different markers of cell-cycle progression. Each bar inside the cell represents a single copy of DNA. The feature at the neck joining the mother and daughter cell represents the myosin ring. The myosin ring appears late in G_1, marking the location where the bud will emerge, and disappears with cytokinesis, indicating the separation of the mother and daughter cytoplasms. After cell wall separation, the mother and daughter cells are free to undergo more rounds of division. In budding yeast, division is asymmetric and daughters (shown outside) are born smaller than their mothers.

growth and division in budding yeast.

These time-lapse datasets also allow investigation of correlations *between* measurements made at different cell cycles, an important gap in our understanding of coordination between growth and division. In multicellular systems, coordination of division among cells has important implications for higher-scale phenomena like development, differentiation, and tissue organization [14–18]. In unicellular organisms like the budding yeast, *Saccharomyces cerevisiae*, inter-cycle dependencies between growth and division are plausible [19] and might affect more classically studied intra-cycle dependencies or characteristics. For example, a mother cell with some advantage in cell division or growth might transmit that advantage to her progeny, resulting in a fast-dividing or fast-growing daughter cell. However, it is unclear the extent to which cell-cycle progression is correlated across cell cycles in budding yeast, if at all.

Statistical modeling provides a powerful and principled foundation for characterizing these correlations in lineages of proliferating cells. Indeed, correlation and biological lineage analysis have been intertwined since the development of the correlation coefficient by Galton and Pearson for data on mother and daughter pea seed size [14, 19–22]. Statistical models of correlation in cellular characteristics involve specifying multivariate probability distributions on observations of each branch of the lineage tree, and these approaches have been successfully applied to bacterial and mammalian cell lineage data [23, 24]. In addition, modern Bayesian inference techniques for regression and model averaging provide a framework for evaluating the plausibility of a variety of different models of correlation between growth and division [25].

Here, we provide in-depth statistical analysis to address four main biological questions: (1) Is $S/G_2/M$ duration approximately constant across cells or does it vary among mothers and daughters?; (2) Is $S/G_2/M$ duration independent of cell size?; (3) Is size at birth independent of size accumulated over the cell cycle?; and (4) Is there evidence

bioRxiv preprint doi: https://doi.org/10.1101/082735; this version posted November 24, 2016. 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-NC-ND 4.0 International license.

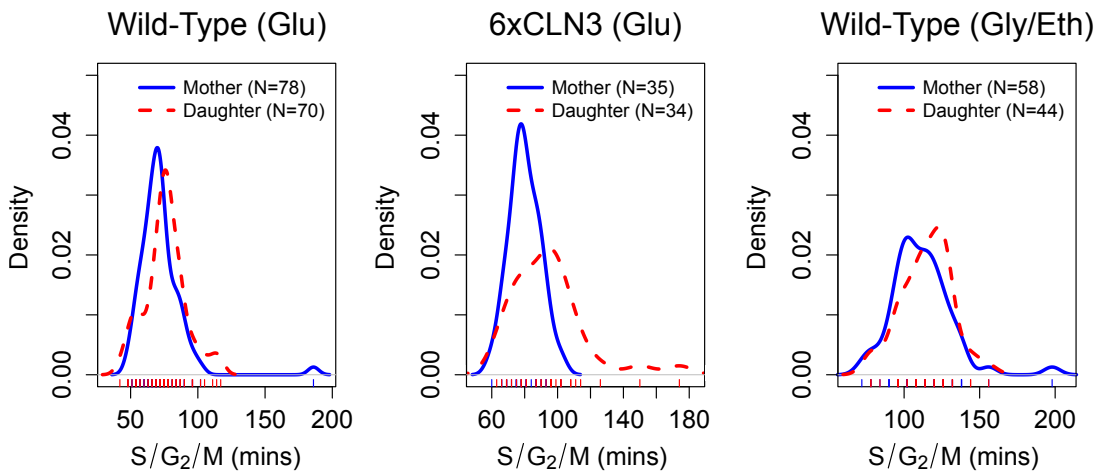

FIG. 2: Density plots of $S/G_2/M$ durations for mother and daughter cells in the three different experimental conditions. Rug plots appear below each density plot.

for inter-cycle dependencies in cell-cycle progression that accompany known intra-cycle dependencies? We carry out our analysis on three recently published microscopy datasets comprising different genetic and nutrient environment conditions [2]. We conduct a Bayesian regression analysis to further investigate the dependence between cell growth and division within and between cycles and comprehensively evaluate the plausibility of different models of correlation. We introduce a novel hierarchical statistical model [26, 27] of budding yeast cell division at the single-cell level to formally characterize inter-cycle correlations in cell-cycle progression, to enable pooling of information across replicate cell lineages, and to separate cell-to-cell variation from variation arising due to measurement error. Our analysis offers fresh biological and methodological insights on the extent and nature of coordination between cell division and cell growth, as well as a novel framework for formally characterizing dependencies within and between cells in these and other biological processes.

II. SINGLE-CELL ANALYSIS OF BUDDING YEAST SIZE CONTROL MODELS

A. Single-cell measurements of *Saccharomyces cerevisiae* growth and division

Single-cell data of haploid budding yeast were acquired from a previously published study [2]. The study followed cell-cycle progression and growth in 26 wild-type lineages (782 cells) grown in glucose, 19 6×CLN3 lineages (376 cells) grown in glucose, and 21 wild-type lineages (518 cells) grown in glycerol/ethanol. Only those cells (or a subset thereof where specified) with fully observed cell-cycle durations were retained for subsequent processing and analysis, resulting in 213 wild-type cells in glucose, 99 6×CLN3 cells, and 157 wild-type cells in glycerol/ethanol.

Cell-cycle progression was measured using the times of occurrence of two landmark cell-cycle events for each yeast cell on the plate: the appearance and disappearance of the myosin ring, visualized by tagging Myo1p with green fluorescent protein (GFP) (Figures 1). The myosin ring is a contractile structure that appears late in G_1, just prior to the appearance of the bud [28] (Figure 1). The disappearance of the myosin ring marks the end of cytokinesis and hence the separation of the shared cytoplasm of the budded cell into mother and daughter cytoplasms (Figure 1). Cell growth was monitored with the red fluorescent protein, DsRed, which was placed under the control of the promoter of ACT1, the constitutively expressed actin gene. In this way, total red fluorescence in a cell served as a proxy for total protein content or cell mass. Red fluorescence in a cell was quantified at each time point and suitably normalized across all cells in a microcolony (see Supplement). Cells and their progeny were continuously monitored until high cell density prevented further accurate measurements.

B. Size-dependent differences in $S/G_2/M$ duration between mothers and daughters

Budding yeast cells born at a smaller than average size tend to undergo longer G_1 phases to reach a sufficient size for cell-cycle entry, manifesting as a dependence between cell size and division within a given cell cycle. Under the assumption that mothers and daughters complete G_1 at roughly the same size, the cells could maintain a consistent

bioRxiv preprint doi: https://doi.org/10.1101/082735; this version posted November 24, 2016. 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-NC-ND 4.0 International license.

Dataset	Estimate (mins)	P-value
Wild-Type (Glu)	−6 (−9, 0)	0.021
6×CLN3	−12 (−18, −3)	0.001
Wild-Type (Gly/Eth)	−6 (−12, 0)*	0.166

TABLE I: Comparision of $S/G_2/M$ duration between mother and daughter cells. Cell counts are the same as in Figure 2. Estimates of differences in $S/G_2/M$ duration and 95% confidence intervals (in parentheses) in minutes are shown in parentheses next to each p-value. P-values less than 0.001 are shown in scientific notation. *Before rounding of interval estimates for display purposes (estimates of upper bound of intervals near 0 had 5 significant digits), interval for wild-type cells in glycerol/ethanol included 0 while interval for glucose did not (hence differences in p-values).

	Wild-Type Glucose		6×CLN3 Glucose		Wild-Type Gly/Eth	
Parameter	Estimate	(P-value)	Estimate	(P-value)	Estimate	(P-value)
Intercept	−0.26	(1.03×10^{-6})	−0.06	(0.366)	0.09	(0.068)
Coefficient	−1.16	(2.19×10^{-8})	−0.68	(0.047)	−0.36	(0.061)

TABLE II: Effect of daughter-mother differences in size at budding on differences in log $S/G_2/M$ duration. Shown are parameter estimates with p-values assessing whether estimates were significantly different from 0 (in parentheses). Mother and daughter fitted masses at budding were computed from linear regressions on the logarithm of each cell's growth traces. See part D of section II for more details. P-values less than 0.001 are shown in scientific notation. Daughter-mother pairs: wt (glucose) N=44; 6×CLN3 N=16; wt (gly/eth) N=26.

cell size distribution from generation to generation provided the amount of time they spent in $S/G_2/M$ was similar on average. As such, we tested whether combined $S/G_2/M$ duration is roughly constant and shared across mother and daughter cells [9].

To assess differences in the observed mother and daughter $S/G_2/M$ duration, we performed two-sided nonparametric Wilcoxon rank sum tests in the three datasets (Figure 2 and Table I). For this analysis, we separated mother and daughter $S/G_2/M$ durations. We used a subset of the mother cells since some $S/G_2/M$ durations were associated with the same mother in different, consecutive cell cycles (see Supplement for details). We found significantly shorter $S/G_2/M$'s for mothers compared with daughters for both 6×CLN3 cells and wild-type cells growing in glucose. We also found suggestive (but not significant) differences in $S/G_2/M$ for wild-type cells growing in glycerol/ethanol.

One potential explanation for these differences in $S/G_2/M$ duration is that daughter cells smaller than their mothers at the onset of $S/G_2/M$ might require more time to complete the budded period. We tested this hypothesis by pairing mother cells from the previous analysis with their first daughters. In this way, mother-daughter pairs only consisted of daughter cells whose immediate mother was not also a daughter cell herself. We then fit a linear regression of the daughter-mother difference in log $S/G_2/M$ duration on the differences in estimated mass at budding (Figure 3). In this regression model, the intercept represents the conditional expected difference between daughter and mother log $S/G_2/M$ durations independent of size differences and the coefficient is the slope of the difference in daughter and mother fitted masses at budding. In two of the three experimental settings, the difference between daughter and mother mass at budding was a significant negative predictor of differences in $S/G_2/M$ duration (Table II). Thus, daughters smaller than their mothers at the point of cell-cycle entry tended to spend more time in $S/G_2/M$. Collectively, these results provide evidence that $S/G_2/M$ duration is not similar between mother and daughter cells, and, surprisingly, we identify a growth-related contribution to this difference in how quickly cells progress through the rest of their cell cycle after G_1.

C. Significant dependencies observed between size at birth and size accumulated

Recent studies in bacteria and budding yeast have suggested a compelling alternative model for size control called the adder model [6, 11–13] which posits that the size added by cells during the cell cycle is roughly constant from cell to cell and independent of the cell's size at birth. So, we set out to evaluate evidence for this hypothesis in our time-lapse measurements. For this analysis, we evaluated dependencies between the observed birth mass (M_0) and mass accumulated between birth and division ($M_{add} = M_{div} - M_0$) for mothers and daughters separately. Interestingly, we find strong negative dependencies between mass at birth and mass accumulated over the cell's life in every cell type and every condition (see Figure 4). Moreover, the correlations we observe are all significant with two-sided hypothesis tests of the Spearman rank coefficients (Table III).

bioRxiv preprint doi: https://doi.org/10.1101/082735; this version posted November 24, 2016. 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-NC-ND 4.0 International license.

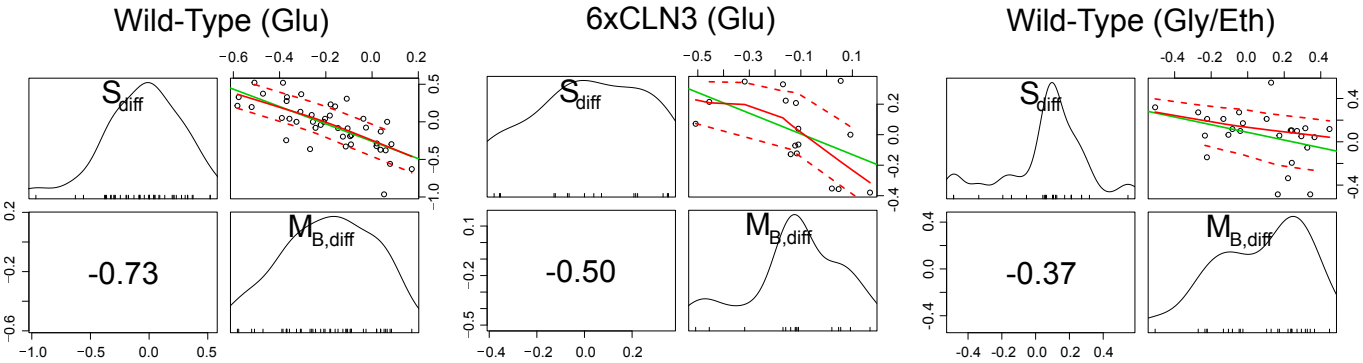

FIG. 3: Scatter and marginal plots of mother-daughter differences in $S/G_2/M$ duration and mass at budding. Marginal density plots of each variable are shown on the diagonal. In the lower off-diagonal panel appears the Spearman rank correlation coefficient. A scatter plot of the data points along with a best linear fit line (green), loess smoothed fit line (solid red) and loess spread lines (dashed red; root mean squared positive and negative residuals). The loess span was 1.0. Daughter-mother pair counts are those listed in Table II.

Dataset	Cell Type	$\mathbf{M}_{add,full}$ P-value	$\mathbf{M}_{add,sg2m}$ P-value
Wild-Type (Glu)	Mothers (N=78)	0.011	0.188
	Daughters (N=70)	2.10×10^{-8}	0.183
6×CLN3	Mothers (N=35)	2.22×10^{-4}	0.003
	Daughters (N=34)	1.69×10^{-4}	0.001
Wild-Type (Gly/Eth)	Mothers (N=58)	3.82×10^{-7}	0.001
	Daughters (N=44)	4.49×10^{-5}	0.172

TABLE III: Test of no correlation between size at birth and mass accumulated over the cell cycle. The third column shows p-values for the correlation between mass at birth (M_0) and mass added over the cell's entire cell cycle ($M_{add,full}$). The fourth column shows p-values for the correlation between birth mass and mass added during $S/G_2/M$ ($M_{add,sg2m}$). P-values less than 0.001 are shown in scientific notation.

One possible explanation for the dependence we observe is that it is driven primarily by a negative correlation between mass at birth and size accumulated during G_1 (more classical size control dependence) and that mass at birth and size accumulated during $S/G_2/M$ are uncorrelated. However, we observe significant negative associations between mass at birth and size accumulated during $S/G_2/M$, particularly in 6×CLN3 cells (Table III). For 6×CLN3 cells, the dependencies we observe might indicate a compensatory mechanism during $S/G_2/M$ to overcome disabled G_1 size control and ensure robust cell size at division. In aggregate, we find no evidence for adder model effects in our time-lapse datasets.

D. Post-G_1 dependence between cell-cycle progression and cell growth

As aforementioned, budding yeast daughter cells tend to spend more time in G_1 than their mothers to reach a sufficient size for cell-cycle entry. This reflects a dependence between G_1 duration and cell size at birth. It has been hypothesized that G_1 is the primary period during which cell-cycle progression depends on cell size and that $S/G_2/M$ progression is largely independent of size, subject instead to a timing mechanism [10]. Moreover, analyses of coordination between growth and division have focused primarily on dependencies *within* rather than *across* cell cycles. However, given that budding yeast cells divide asymmetrically, leading to partitioning of organelles and other cellular contents between mothers and daughters, it is plausible that cell-cycle progression might depend on the growth and division characteristics of the cell's mother as well as on the size of the cell itself.

Classically, one would analyze the correlation between a cell-cycle interval (e.g. G_1) and the cell's size at the beginning of that interval. Here, by conditioning on more predictor variables, we can estimate the relative effects of a cell's size and the growth and division characteristics of its mother on the cell's current cell-cycle progression. To analyze these effects, we first computed growth characteristics of a cell and its immediate antecedent cell. Using the single-cell growth traces of each cell j and its immediate predecessor cell $Pa(j)$ (mother cycle that preceded cycle

bioRxiv preprint doi: https://doi.org/10.1101/082735; this version posted November 24, 2016. 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-NC-ND 4.0 International license.

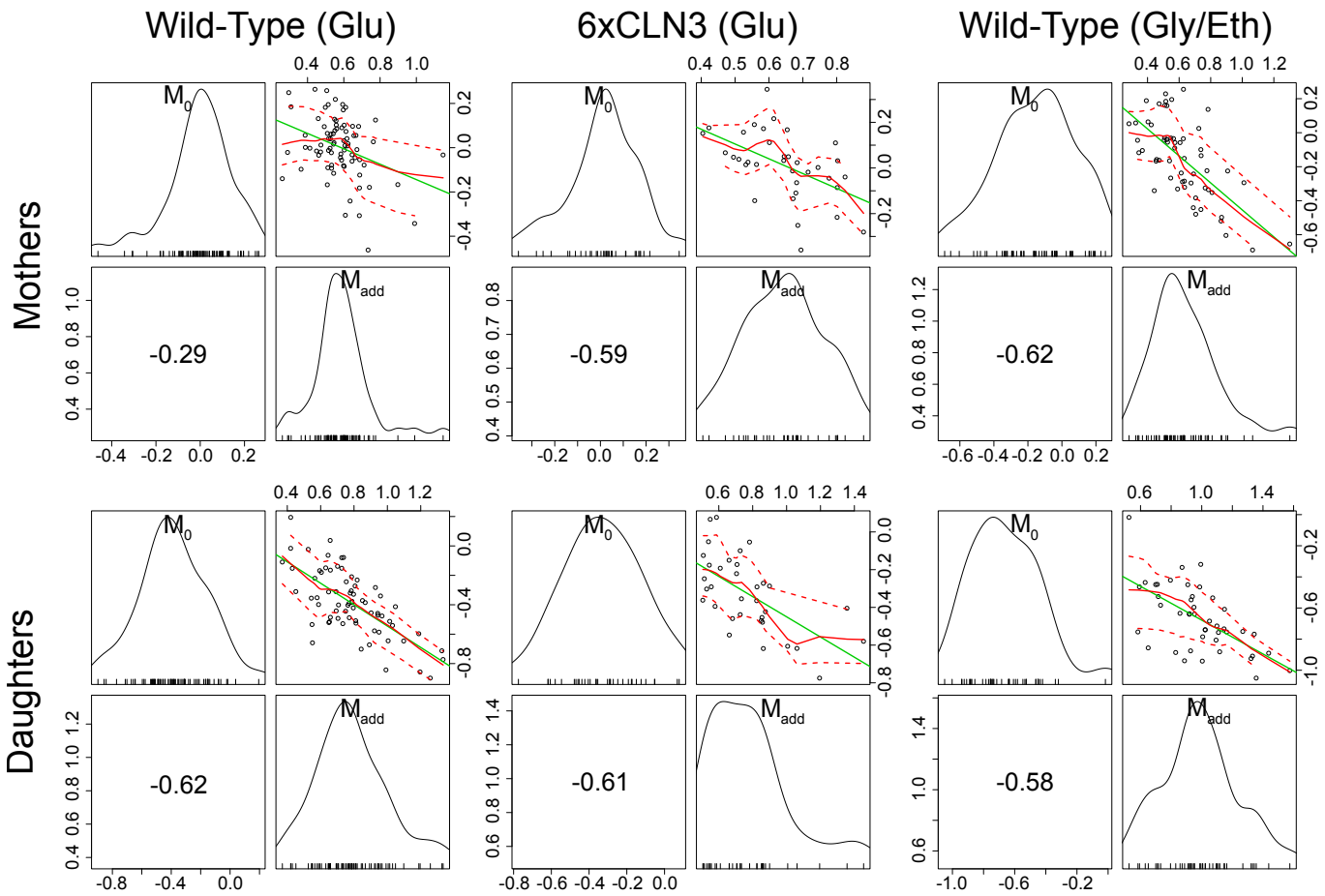

FIG. 4: Scatter plots and univariate density plots of mass at birth and mass accumulated from birth to division for mother and daughter cells in all three experimental conditions. The layout and content of each of the six panels is the same as in Figure 3. Cell counts are as in Table III.

of cell j) in lineage i, we estimated the following growth-related variables ($\hat{\alpha}_{i,j}$, $\hat{\alpha}_{i,Pa(j)}$, $\hat{M}_{0,i,j}$, $\hat{M}_{0,i,Pa(j)}$, $\hat{M}_{B,i,j}$, and $\hat{M}_{B,i,Pa(j)}$) [2], by assuming exponential single-cell growth kinetics and fitting a separate linear model to the logarithm of each cell's growth trace (see Supplement). The fitted intercept of this linear model gave the estimated birth mass ($\hat{M}_{0,i,j}$) of cell j from each lineage i while the slope gave the estimated mass accumulation rate ($\hat{\alpha}_{i,j}$). We also retained the fitted mass at budding of each cell ($\hat{M}_{B,i,j}$). We then fit linear regression models for each of the three experimental datasets of log S/G$_2$/M durations on these cell-level estimates as well as on the log S/G$_2$/M duration of the cell's predecessor ($S_{i,Pa(j)}$): either the duration from its previous cycle (for mothers) or from its mother's cycle (for daughters).

In this setting, a model represents a particular pattern of dependence between growth and division and is determined by the set of predictor variables included in the regression. As we have 7 different predictor variables (not counting the included intercept term, μ_S), there are $2^7 = 128$ possible regression models. To infer the most plausible model of dependence between size and cell-cycle progression while explicitly accounting for uncertainty in the model specification, we conducted Bayesian model averaging [25]. Since we didn't have strong prior information about the dependencies between the collection of division and growth variables, we assumed that each regression model was equally plausible *a priori*. We then computed posterior probabilities of each model for both mother cells (Figure 5) and daughter cells (Figure 6). To carry out this analysis, we started from all mother cells in the lineage. To account for potential correlations between different cell cycles of the same mother cell, we only retained every other cell cycle of the mother cell starting with its most recent cycle. This procedure resulted in 53 wild-type pairs in glucose, 25 6×CLN3 pairs, and 46 wild-type pairs in glycerol/ethanol. To investigate effects on dependence between mothers and daughters, we used the mother-daughter pairs of a previous analysis (cell counts provided in Table II).

When we consider mother-to-mother dependencies, we find a strong association for wild-type mothers between S/G$_2$/M duration and growth characteristics, particularly mass at budding (Figure 5, left and right panels). Mass at

bioRxiv preprint doi: https://doi.org/10.1101/082735; this version posted November 24, 2016. 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-NC-ND 4.0 International license.

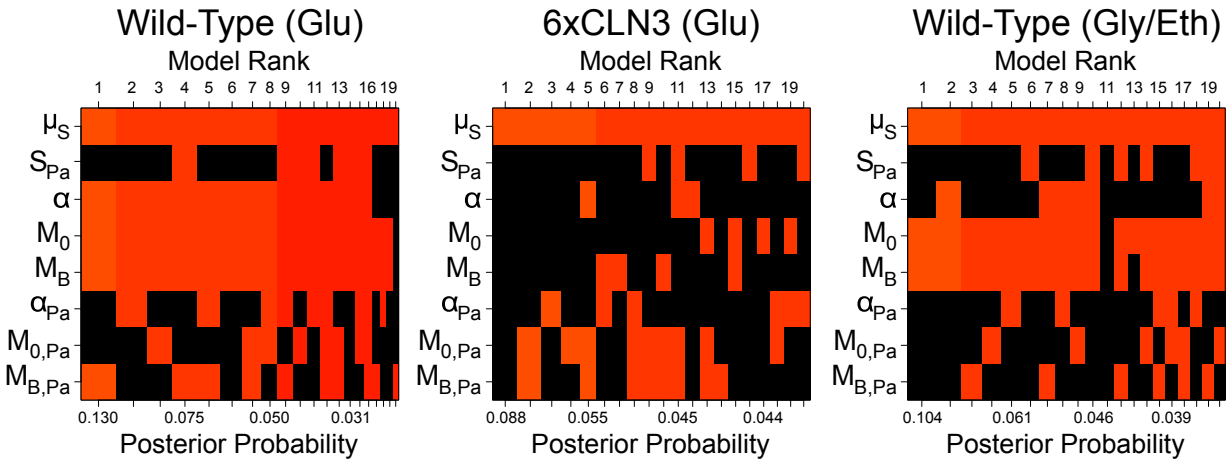

FIG. 5: Bayesian adaptive sampling results for mother cells. Shown are the top 20 models (columns ranked by posterior probability) where colored squares indicate that the corresponding predictor variable is included in the model and black squares indicate exclusion of the variable.

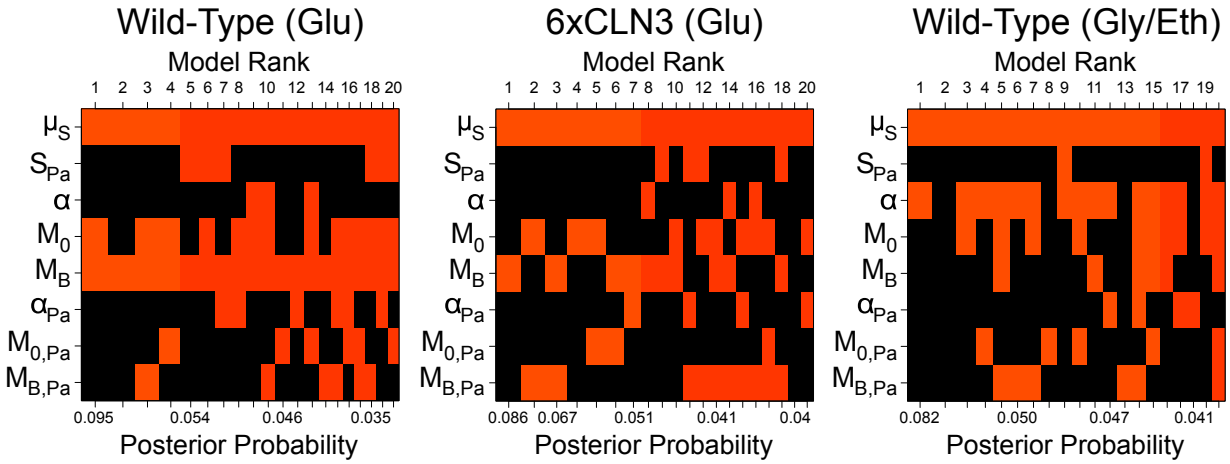

FIG. 6: Bayesian adaptive sampling results for daughter cells. Panel layout and content is the same as in Figure 5.

budding (M_B) was included in nearly every enumerated model with non-zero posterior probability (Figure 5) indicating that mass at budding was an informative predictor of mother S/G$_2$/M duration (∼82% posterior probability of inclusion; $\log_{10}$(Bayes factor) = 0.657 or 'substantial evidence' for inclusion [29]). Mass at birth (M_0; ∼74% posterior probability) and mass accumulation rate (α; ∼71% posterior probability) also tended to be included as predictors, reinforcing the dependence between current cell growth characteristics and S/G$_2$/M duration. Likewise, for a mother growing in glycerol/ethanol, we detect an association between her current S/G$_2$/M duration and mass at birth (∼64% posterior probability) and budding (∼61% posterior probability; Figure 5, right panel). In addition, the posterior means (averaged across all models) of the included regression coefficients for mass at budding for wild-type mothers in glucose (−1.129) and glycerol/ethanol (−0.262) were consistent with classical G$_1$ size control (larger mass at budding corresponds to less time spent in S/G$_2$/M). This pattern of dependence has not previously been observed in mother cells, potentially due to the fact that we are conditioning on multiple growth and division characteristics for the cell's current and previous cycles. Instead, standard analyses have focused primarily on the cell's current cycle. We did not see such patterns of dependence for 6×CLN3 mother cells. Importantly, we also did not find strong evidence for dependence between the S/G$_2$/M duration of a wild-type or 6×CLN3 mother in her current cycle and the growth or division characteristics of the mother in her previous cycle (variables with Pa subscript in Figures 5 and 6) indicating that, given the cell's current growth and division characteristics, her S/G$_2$/M duration can be considered independent of her previous growth and cell-cycle progression.

Extending this analysis to mother-to-daughter associations, we again discovered patterns of dependence between a cell's S/G$_2$/M duration and mass at budding (M_B): for wild-type daughter cells growing in glucose (Figure 6, left panel). Mass at budding was included as an explanatory variable for the wild-type daughter's S/G$_2$/M duration (in

bioRxiv preprint doi: https://doi.org/10.1101/082735; this version posted November 24, 2016. 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-NC-ND 4.0 International license.

glucose) in nearly all models with non-zero probability (∼93% posterior probability of inclusion; $\log_{10}$(Bayes factor) = 1.118 or 'strong evidence' for inclusion according to [29]; Figure 6, left panel). As in the previous analysis, the model-averaged posterior mean of the included regression coefficient was −0.616, an estimate consistent with classical G1 size control. We also found mild associations between $S/G_2/M$ duration and other growth characteristics in all three conditions: ∼52% posterior probability of inclusion of mass at budding as a predictor in 6×CLN3 cells; ∼57% and ∼50% posterior probabilities of inclusion of birth mass (M_0) as a predictor in both wild-type cells in glucose and 6×CLN3 cells; ∼58% posterior probability of inclusion of mass accumulation rate (α) as a predictor in wild-type cells in glycerol/ethanol. For the last association in glycerol/ethanol, we further note that the model-averaged posterior mean of the included regression coefficient for mass accumulation rate was −18.391, indicating that daughters with larger mass accumulation rates spent less time in $S/G_2/M$. Collectively, our findings for both mother and daughter cells run counter to the notion that $S/G_2/M$ duration is independent of size.

III. HIERARCHICAL MODELING OF CORRELATION IN BUDDING YEAST CELL DIVISION AT THE SINGLE-CELL LEVEL

The regression framework used in the previous section helped identify dependencies between growth and division within and across cycles. However, we limited our analysis to rigidly defined mother-mother and mother-daughter pairs and did not take advantage of the inherent hierarchical organization of the data (i.e. cells make up lineages, multiple lineages are observed for each experimental condition). Moreover, simply computing sample-based estimates of inter-cell correlations would preclude separation of cell-to-cell variation in cell-cycle progression from variation due to measurement error. Hierarchical models provide a formal framework to represent such structure and naturally pool information across replicate lineages as well as allow for estimation of cell-specific and noise-related sources of variation. An important property of these models is the potential to reduce parameter estimation error relative to sample-based approaches by shrinking cell-specific parameter estimates towards sample (population) estimates [26]. For these reasons, and to more effectively characterize dependencies between cells, we opted to analyze single-cell cell-cycle durations with a novel hierarchical model.

A. Observing a cellular branching process

To motivate our model, we consider how the single-cell data are acquired in a single time-lapse movie. A single yeast cell (the founder cell; cell 1 in Figure 7) growing on the agarose slab is identified at the onset of the time-lapse experiment. The time at which the founder cell begins its cell cycle is generally not observed, so the full cell-cycle duration of the founder cell cannot be determined. Moreover, the identity of the founder cell (mother or daughter) is unknown. Each founder cell generates a lineage consisting of two fully observed sub-lineages (Figure 7). The initial cell of the first sub-lineage is simply the founder cell in its first cell cycle after division from its daughter. The initial cell of the second sub-lineage is the founder cell's first daughter.

For each cell, we observe the time since the start of the time-lapse experiment of the appearance and disappearance of the cell's myosin ring. Hereafter, we refer to these times as budding and division times, respectively. In our notation, the budding time for cell j from lineage i is $B_{i,j}$ and the division or cycle time is $C_{i,j}$ (Figure 8). We transformed these times to cell-specific budding ($B^{rel}_{i,j}$) and division ($C^{rel}_{i,j}$) durations (Figure 8; details in Supplement). To refer to durations specific to each cell, we adopted the binary indexing scheme of Di Talia et al. [2] and classified cells into four categories: mother origin, daughter origin, mother, or daughter (Figure 7).

We view the budding and division durations from each lineage (as in Figure 8) as noisy observations from an underlying branching process. A diagram illustrating a sample branching process for the observed lineage in Figure 8 appears in Figure 9. The branches of the tree in Figure 9 represent the expected values of the observed cell-cycle durations for each cell. The expected value of a cell's division duration is given by the full branch length while the cell's expected budding duration is some fraction of the branch length. To identify inter-cycle correlation in cell-cycle progression, we assume that the branch lengths in the process may depend on one another. Further details of model construction follow.

B. Likelihood and error model for budding and division observations

The first (or lowest) level of our hierarchical model gives a likelihood for the observed budding and division durations conditioned on cell-specific and population-level parameters of the model (discussed below). This first level captures noise or error in observations while higher levels of the hierarchical model will capture cell-to-cell variability. In

bioRxiv preprint doi: https://doi.org/10.1101/082735; this version posted November 24, 2016. 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-NC-ND 4.0 International license.

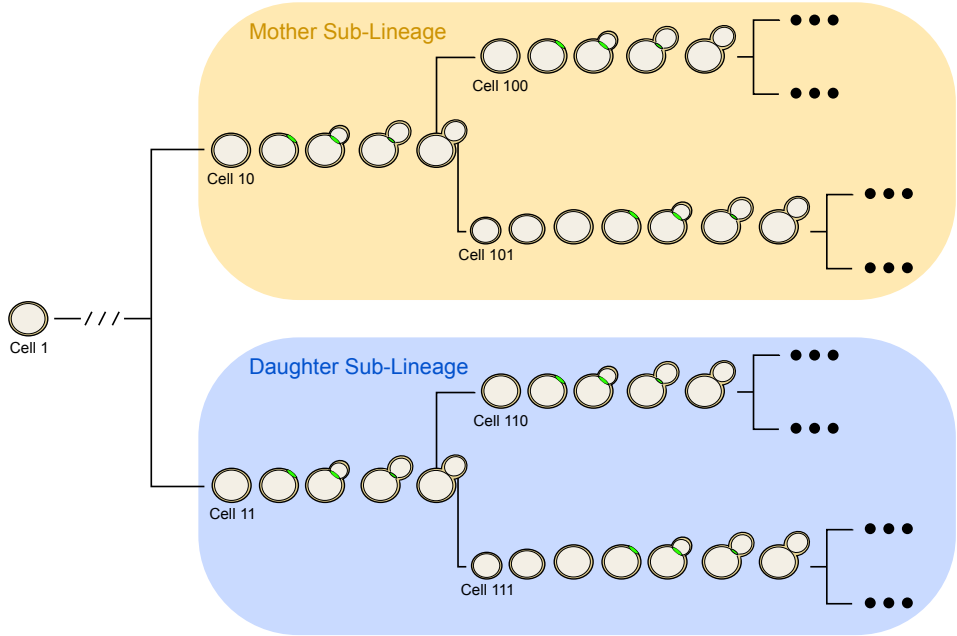

FIG. 7: Illustration of single-cell lineages and classification of cell types. Shown is a typical single-cell lineage tree from the dataset of Di Talia *et al.*, 2007. Arranged along each branch of the lineage tree (individual cell cycles) are images of representative cells undergoing cell-cycle events (e.g. budding; appearance of green myosin ring at bud neck). The founder cell of each lineage was considered cell 1. The mother sub-lineage of the lineage tree consists of the mother origin cell and her progeny. Likewise, a daughter origin cell and her progeny constitute the daughter sub-lineage. We classify cells into two other types: mother cells and daughter cells. The division of any cell produces a new mother cell and daughter cell. Binary cell labels ending in 0 indicate a mother cycle while labels ending in 1 indicate a daughter cycle.

constructing an error model for the budding and division durations, we first assumed that the elapsed times from which the durations for cell j in lineage i were derived were independent and normally distributed with means $\mu_{B_{i,j}}$ and $\mu_{C_{i,j}}$ and variance τ^2. While the observed cell-cycle durations are positively valued, we use normal errors at the lowest level of the hierarchical model rather than alternative error models (e.g. log-normal) as we don't expect multiplicative errors from the manual recording of budding and division events. Rather, as will be discussed in a later section, we constrain the branch lengths or expected values of a cell's durations to be positively valued. After transformation of the elapsed times into durations, and concatenation of the durations for lineage i into a vector, the likelihood of budding and division durations for lineage i is:

$$\begin{pmatrix} \tilde{B}_i^{rel} \\ \tilde{C}_i^{rel} \end{pmatrix} \Big| \, \tilde{\gamma}_i, \tilde{\beta}_i, \Theta_{pop}, \tau^2 \sim \mathsf{MVNorm}(A\tilde{\mu}_i, \tau^2 AA').$$

Here, A is a linear transformation matrix (to convert elapsed times into durations) and $\tilde{\mu}_i$ is a vector of the expected durations from birth to budding and birth to division for lineage i. $\tilde{\gamma}_i$ is a vector containing two types of cell-specific parameters that make up the branching process of a lineage: the expected base cell-cycle durations for each cell in lineage i ($\lambda_{i,j}$s) and extensions to expected daughter cell-cycle durations due to smaller size at birth ($\delta_{i,j}$s). Θ_{pop} is the set of population-level parameters $\{\Lambda, \Delta, \sigma_\lambda^2, \sigma_\delta^2, \psi, \rho, \phi\}$ (see Table IV) that specify correlation and cell-to-cell variation in the $\lambda_{i,j}$s and $\delta_{i,j}$s. The vector $\tilde{\beta}_i$ contains another set of cell-specific parameters: the proportions of cell-cycle time (or the fraction of a cell's λ branch) in the unbudded phase ($\beta_{i,j}$) for each cell in the lineage, another source of variability in cell-cycle progression. We describe the model for the means ($\tilde{\mu}_i$) of the observations and the remaining parameters in the next section.

C. Representing inter-cycle dependence and cell-to-cell variability in cell-cycle progression with an asymmetric branching process

The second level of the hierarchical model consists of cell-specific parameters (e.g. branch lengths) that give the expected value of a cell's budding and division durations for a particular lineage. This second level comprises a

bioRxiv preprint doi: https://doi.org/10.1101/082735; this version posted November 24, 2016. 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-NC-ND 4.0 International license.

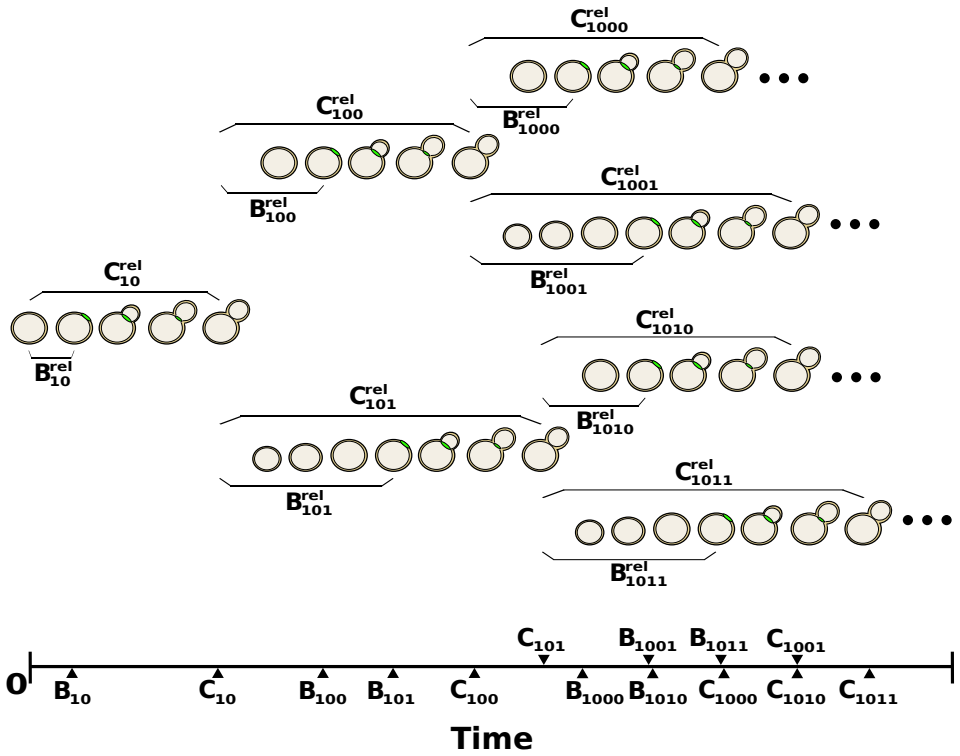

FIG. 8: Sample diagram of budding and division observations arising from the time-lapse microscopy experiments. Here, an example mother origin cell (cell 10) of the sub-lineage proceeds through the cell cycle, undergoing budding and division. Once divided from her daughter (cell 101), the mother origin (now mother cell 100) undergoes another round of division. Budding ($B_{i,j}$) and division ($C_{i,j}$) times were recorded for each cell (as shown on timeline at base of figure). These times were transformed to durations of budding ($B^{rel}_{i,j}$) and division ($C^{rel}_{i,j}$). The ellipses following the lineage leaves indicate that other lineages could be larger. The hierarchical model was fit to these durations.

Parameter	Description
Λ	population average of mother cell-cycle duration (minutes)
Δ	population average of daughter cell G_1 extension duration (minutes)
ψ	correlation between λs from two successive mother cycles
ρ	correlation between a mother's λ and her daughter's λ
ϕ	correlation between a mother's λ and her daughter's δ
σ^2_λ	variance in cell-specific λ branch lengths (minutes^2)
σ^2_δ	variance in cell-specific δ branch lengths (minutes^2)
μ_m	expected proportion of λ spent unbudded in mothers
μ_d	expected proportion of λ spent unbudded in daughters
η_m	prior weight of information for mother unbudded period
η_d	prior weight of information for daughter unbudded period
τ^2	variance in measurement error (minutes^2)

TABLE IV: Population parameters of hierarchical model

branching process in which the branch lengths are potentially correlated and vary from cell to cell (Figure 9). The process accounts for the mother-daughter asymmetry typical of budding yeast in terms of both correlation structure and branch length. The branch lengths for a lineage i are represented by a vector $\tilde{\gamma}_i$, composed of two sets of parameters: $\lambda_{i,j}$s and $\delta_{i,j}$s. The δs were introduced to account for longer daughter cell cycles due to smaller birth sizes. To measure correlation between these branch lengths, we introduced three parameters: ψ, ρ, and ϕ. ψ is the correlation between the λs of a mother cell in two successive cycles, ρ is the correlation between λs of a mother and her daughter cell, and ϕ is the correlation between a mother's λ and her daughter's δ.

As noted in previous work [30, 31] and since the cell-cycle durations are positively valued, we jointly model all

bioRxiv preprint doi: https://doi.org/10.1101/082735; this version posted November 24, 2016. 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-NC-ND 4.0 International license.

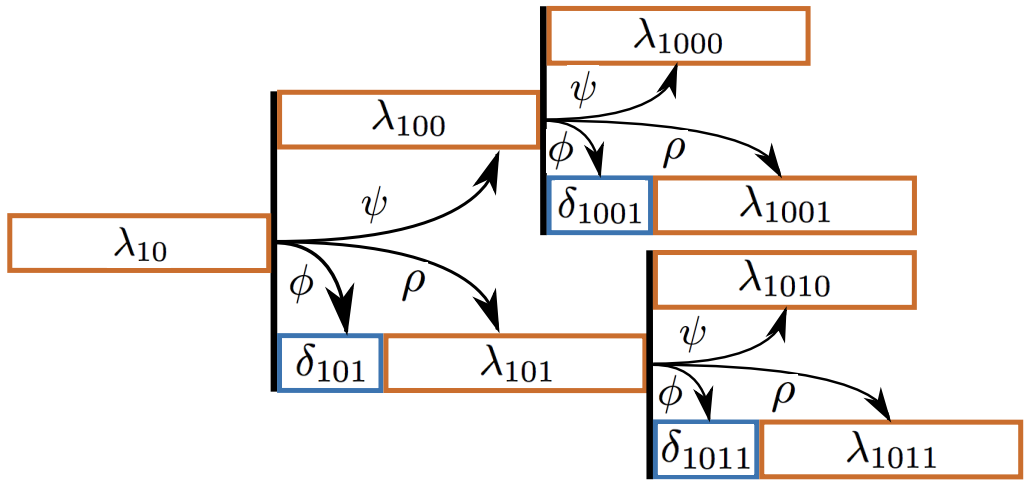

FIG. 9: A diagram of the asymmetric branching process specifying expected cell-cycle durations for each cell. The diagram is drawn to indicate the branch lengths that give rise to the relative budding and division measurements in the sub-lineage of Figure 8. λ_{10} is the expected cell-cycle duration of mother origin cell 10. The expected cell-cycle duration of her subsequent cycle is λ_{100}, which depends on her first cell cycle through the correlation parameter ψ. For the daughter branch, two parameters specify total expected cell-cycle duration: δ_{101} and λ_{101}. In general, the λ component represents a 'baseline' cell-cycle duration for the daughter cell to which the daughter extension component is prepended to account for the longer G_1 observed in daughters. These branch lengths depend on the mother's cell-cycle duration through the correlation parameters ρ and ϕ, respectively.

branch lengths for a lineage i ($\tilde{\gamma}_i$) with a multivariate log-normal distribution. That is:

$$\tilde{\gamma}_i = \mathsf{exp}(\tilde{\gamma}_i^*) \tag{1}$$

with

$$\tilde{\gamma}_i^* \,|\, \Theta_{pop}^* \sim \mathsf{MVNorm}(\mu_{\tilde{\gamma}_i}^*, \Sigma_{\tilde{\gamma}_i}^*). \tag{2}$$

Here, $\tilde{\gamma}_i^*$ is the vector of cell-specific branch lengths on the natural logarithmic scale. These log-scale durations follow a multivariate normal distribution with a structured covariance matrix $\Sigma_{\tilde{\gamma}_i}^*$. In this covariance matrix, we encode a simple model of inter-cycle dependence. With no strong expectations of the extent of inter-cycle correlation structure, we consider the simplest model for inter-cell correlation: that the expected log-scale cell-cycle duration of a newly arisen cell depends solely on the expected log-scale cell-cycle duration of its predecessors in the lineage only through its mother. This assumption constitutes a first-order autoregressive or AR(1) assumption of dependence. As we do not observe the cell-cycle durations of each lineage's founder cell (Figure 7), this correlation structure dictates that the branch lengths of the mother origin ($\lambda_{i,j}$) and daughter origin are correlated with one another, and we model them accordingly (see Supplement). However, in general, we do not directly model correlation between branches of sister cells owing to our AR(1) dependence assumption and to all sub-lineage cell-cycle durations being fully observed. In addition, by jointly modeling the $\lambda_{i,j}^*$s and $\delta_{i,j}^*$s with a multivariate normal distribution, we assume that the log-scale branch lengths of cell j are conditionally normally distributed given the λ^* of the cell's mother (or $\lambda_{i,Pa(j)}^*$).

The mean vector $\mu_{\tilde{\gamma}_i}^*$ consists of parameters Λ^* and Δ^* (counterparts of Λ and Δ on the log scale). $\Sigma_{\tilde{\gamma}_i}^*$ is parameterized by log-scale analogues of ψ, ρ, ϕ, σ_λ, and σ_δ. To infer parameters on the original scale (Table IV), we transform the log-scale analogues (see Supplement).

Given the branch lengths on the original scale, additional parameters describing cell-cycle progression across all replicate lineages in a condition (Θ_{pop}), and a cell's identity (mother or daughter), we can compute the expected value of a cell's observed budding and division durations. For example, if a cell j is a mother cell then:

$$E[B_{i,j}^{rel} \,|\, \tilde{\gamma}_i, \tilde{\beta}_i, \Theta_{pop}] = \beta_{i,j}\lambda_{i,j} \tag{3}$$

and

$$E[C_{i,j}^{rel} \,|\, \tilde{\gamma}_i, \tilde{\beta}_i, \Theta_{pop}] = \lambda_{i,j}. \tag{4}$$

On the other hand, if cell j is a daughter cell, then:

$$E[B_{i,j}^{rel} \,|\, \tilde{\gamma}_i, \tilde{\beta}_i, \Theta_{pop}] = \delta_{i,j} + \beta_{i,j}\lambda_{i,j} \tag{5}$$

bioRxiv preprint doi: https://doi.org/10.1101/082735; this version posted November 24, 2016. 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-NC-ND 4.0 International license.

and

$$E[C^{rel}_{i,j} \mid \tilde{\gamma}_i, \tilde{\beta}_i, \Theta_{pop}] = \delta_{i,j} + \lambda_{i,j}. \tag{6}$$

As with the λ and δ parameters, each cell has a parameter indicating the proportion of its $\lambda_{i,j}$ it spends in the unbudded state: $\beta_{i,j}$. The vector $\tilde{\beta}_i$ comprises these cell-specific parameters for lineage i.

The third level of our hierarchical model (represented by the population-level parameters Θ_{pop}; described in Table IV) encapsulates patterns of cell-cycle progression and inter-cycle dependence shared across replicate lineages in a given experimental condition. Owing to the hierarchical structure of the data, we assume for each experimental condition that the branch lengths ($\lambda_{i,j}$s and $\delta_{i,j}$s) from all lineages are drawn from the same distribution. Specifically, cell-specific branch lengths from each replicate lineage in a given experimental condition arise from the multivariate log-normal distribution parameterized by Λ, Δ and other parameters of Θ_{pop}. We make a similar assumption for the cell-specific $\beta_{i,j}$s except that, owing to observed differences in G_1 progression, we consider the cell's type (mother or daughter) when specifying the distribution of $\beta_{i,j}$ parameters.

$$\beta_{i,j} \sim \begin{cases} \mathsf{Beta}(\mu_m \eta_m, (1-\mu_m)\eta_m) & \text{if cell } j \text{ is a mother} \\ \mathsf{Beta}(\mu_d \eta_d, (1-\mu_d)\eta_d) & \text{if cell } j \text{ is a daughter} \end{cases} \tag{7}$$

More formally, we assume that the branch lengths are exchangeable within a given experimental setting and that the unbudded proportions are exchangeable within a given cell type (mother or daughter) and experimental setting [26].

D. Hierarchical model fitting uncovers variation in cell-cycle progression across experimental settings and between mothers and daughters

As one might expect, fits of the hierarchical model to the three different datasets suggest distinct patterns of cell-cycle progression. As shown in Table V, population average mother cell-cycle duration (Λ) was approximately 88 minutes for wild-type cells in glucose. In contrast, mother cells divided nearly twice as slowly in glycerol/ethanol ($\sim$146 minutes). Since Cln3 is a rate-limiting factor for cell-cycle entry [2, 32], daughters with 6 copies of CLN3 show quite short G_1 extensions (δ) compared to wild-type daughters grown in glucose (Table V). Likewise, the estimated spread in 6×CLN3 daughter G_1 extensions (σ^2_δ) was much smaller compared to the corresponding estimates for wild-type cells, reflecting greater availability of Cln3 protein [2]. In contrast, wild-type daughters in glycerol/ethanol took nearly 95 minutes more to complete G_1 (on average) than their mothers. Consistent with empirical evidence, cell-cycle progression in glycerol/ethanol is generally slower than in glucose [33].

One previously mentioned benefit of our hierarchical approach is the ability to separate cell-to-cell variation in cell-cycle progression from measurement error. As mentioned previously, while inferences for σ_λ do not change much between the two types of cells grown in glucose, σ_δ is dramatically reduced in 6×CLN3 cells reflecting differences in cell-cycle progression one might expect. However, as the cells were grown in similar conditions, differences in cell-cycle progression should have little effect on an experimenter's ability to record budding and division times, and so, measurement error should be similar. Importantly, inferences for τ for all three conditions are similar to one another (95% confidence intervals overlap).

We find corroboration for our previous analysis of differences in $S/G_2/M$ duration between mother and daughter cells: the estimates for μ_m and μ_d (expected unbudded proportions of mother and daughter λ's; Table V). We find that the expected mother budded period $(1-\mu_m)$ is mildly (4% of total cell-cycle duration) shorter than the daughter budded period in wild-type cells growing in glucose, with the posterior probability of $(1-\mu_m) < (1-\mu_d)$ being 0.981. In 6×CLN3 cells, this difference in budded duration increases with mothers showing a nearly 8% shorter average budded duration and the posterior probability of $(1-\mu_m) < (1-\mu_d)$ being 0.996. In glycerol/ethanol, differences between mother and daughter budded durations are less pronounced. Although evidence for daughters spending more time in the budded proportion of the cell cycle was still strong (posterior probability of $(1-\mu_m) < (1-\mu_d)$ being 0.892), daughters only took $\sim$3% longer in their budded durations than their mothers.

As part of our deeper investigation of inter-cycle dependence, we generated inferences for the three correlation parameters in the model (ρ, ψ, and ϕ). As shown in Table V, no strong correlations exist in the two strains grown in glucose (wild-type and 6×CLN3) with all 95% posterior confidence intervals overlapping 0. However, we do see moderate mother-to-mother λ correlations (ψ) for wild-type cells grown in glycerol/ethanol. As we did not detect mother-to-daughter correlations in the same conditions, the inferences for ψ suggest that cells in glycerol/ethanol tend to retain the rate of cell-cycle progression with which they are born. This correlation could not be explained by drift in cell-cycle progression due to a cell's replicative age or time spent by the cells on the plate (see Supplement). However, considering our previous analysis indicating that a mother's current $S/G_2/M$ duration is conditionally independent

bioRxiv preprint doi: https://doi.org/10.1101/082735; this version posted November 24, 2016. 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-NC-ND 4.0 International license.

	Wild-Type Glucose		6×CLN3 Glucose		Wild-Type Gly/Eth	
Parameter	**Estimate**		**Estimate**		**Estimate**	
Λ	87.60	(85.08,90.33)	95.80	(92.40,100.79)	145.24	(139.19,152.34)
Δ	25.55	(20.21,31.30)	18.57	(12.20,23.03)	94.02	(81.68,107.78)
μ_m	0.18	(0.16,0.19)	0.15	(0.13,0.17)	0.25	(0.23,0.27)
μ_d	0.13	(0.09,0.17)	0.05	(0.01,0.11)	0.22	(0.17,0.27)
η_m	23.19	(18.45,28.77)	25.48	(19.02,33.19)	22.84	(17.89,28.49)
η_d	16.53	(11.51,23.33)	20.39	(13.48,28.22)	23.37	(16.56,31.56)
ψ	−0.07	(−0.31,0.20)	−0.03	(−0.38,0.50)	0.46	(0.22,0.65)
ρ	−0.18	(−0.41,0.08)	−0.29	(−0.61,0.12)	0.21	(−0.13,0.51)
ϕ	−0.20	(−0.40,0.05)	−0.64	(−0.84,0.01)	−0.02	(−0.26,0.24)
σ_δ	20.00	(14.63,30.32)	6.56	(2.83,10.11)	42.65	(33.16,56.56)
σ_λ	15.30	(13.66,17.20)	15.72	(13.49,18.82)	26.07	(22.54,30.11)
τ	1.31	(0.93,1.91)	1.43	(0.99,2.30)	1.45	(0.96,2.30)

TABLE V: Posterior inferences (modes and 95% highest posterior density intervals) for model parameters. Λ, Δ, σ_δ, σ_λ, and τ are in minutes. Cell-specific unbudded duration parameters, μ_m and μ_d, range from 0 to 1. wt (glucose) N=213; 6×CLN3 N=99; wt (gly/eth) N=157.

of her previous S/G_2/M given current growth characteristics, this cell-cycle dependence is likely mediated by growth characteristics of the mother cell in her current cycle.

To rule out the possibility that this result is an artifact of over-fitting the data, we carried out a leave-one-out cross-validation analysis to evaluate the capacity of models with different numbers and combinations of the correlation parameters to predict the observed cell-cycle durations of the left-out cells. These models in which some or all of the correlation parameters were allowed to vary were compared with a baseline model in which all the correlation parameters were removed (see Supplement). The results of this analysis were consistent with our parameter inferences in that models lacking ψ predicted the observed cell-cycle duration of wild-type cells in glycerol/ethanol more poorly (see Supplement).

Our statistical framework provides a useful tool for experimenters and modelers to characterize cell-to-cell variation in cell-cycle progression as well as other biological processes while taking potential inter-cycle correlation into consideration.

IV. DISCUSSION & CONCLUSIONS

In this novel analysis of time-lapse microscopy data from budding yeast we set out to address questions regarding dependencies within and between the fundamental processes of growth and division. We have found evidence from our analysis contradicting two tenets of budding yeast size control: a relatively constant S/G_2/M duration shared by mother and daughter cells and a lack of dependence between S/G_2/M duration and size. Our hierarchical modeling results and statistical analysis also demonstrated that combined S/G_2/M duration appeared longer on average in daughter cells compared with mother cells. Moreover, we detect a size-related component underlying these differences as the difference in mass at budding of daughters and their mothers was a strong predictor of differences in S/G_2/M duration. In support of our results, at least one classical study with single cells has noted that the budded duration was mildly longer (5–8 minutes on average) for daughters compared with mothers under a range of different growth conditions [34]. These observations are important because experimenters and modelers might otherwise assume approximately similar S/G_2/M durations across cell types or simpler dependence structure between size, G_1 and S/G_2/M that might not be present in their experimental conditions, potentially affecting downstream conclusions about coordination between growth and division.

Our Bayesian regression analysis uncovered patterns of dependence between cell size characteristics and S/G_2/M duration for wild-type cells in two different nutrient conditions. While post-G_1 dependence between growth and division has been noted anecdotally in the literature [7] or postulated in dynamic models of cell-cycle progression [31, 35], our analysis more formally and comprehensively demonstrates this dependence. Importantly, our analysis identifies these dependencies in mother cells as well as daughter cells, on which size control studies have traditionally been focused. While dependence between size and cell-cycle progression has been thoroughly investigated and widely observed in the G_1 phase of the budding yeast cell cycle [1, 2], dual complementary mechanisms of size control have been

bioRxiv preprint doi: https://doi.org/10.1101/082735; this version posted November 24, 2016. 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-NC-ND 4.0 International license.

noted in the fission yeast, *Schizosaccharyomyces pombe*, with a strong size control imposed at the G_2/M boundary and a weaker compensatory size control imposed at the G_1/S boundary [10]. However, while our findings seem at first glance to extend the classical model of size control in budding yeast, we caution that this observed dependence does not necessarily imply a true size control mechanism. Rather, this association could be related to compensation in cell-cycle time due to premature cell-cycle entry or the activation of a cell-cycle checkpoint [36] due to perturbed cell-cycle progression.

Consequently, our analysis has generated experimentally testable hypotheses about the molecular basis of post-G_1 size dependence and insights for future studies of size control. While the dependence we observe in 6×CLN3 cells, for example, is likely not due to activation of the morphogenesis checkpoint (vital for delays in nuclear division due to impaired bud formation) [37], other molecular targets related to DNA replication checkpoints (e.g. Rad53) or cryptic budding yeast size control (e.g. Bck2) should be tested to ascertain their relative effects on dependence between mass at budding and S/G_2/M. Recent experimental work in budding yeast suggests an intriguing mechanistic model for size control in which dilution of Whi5 (by the cell increasing in volume) dictates cell cycle entry at the G_1/S transition [5]. This work could be extended to analyze changes in concentration of Whi5 and other proteins during S/G_2/M to identify potential mechanistic bases for the dependencies we observe.

We do not detect evidence in our datasets for an 'adder' model of size control. Instead, we detect substantial negative dependencies between size at birth and size accumulated over the cell cycle. An important distinction between the current study and previous analyses is the measurement of cell size. In our datasets, cell size was measured via a fluorescent protein-based proxy for cell mass whereas recent work in bacteria and budding yeast has focused on cell volume [6, 11–13]. Elements of cell volume in budding yeast, particularly the vacuoles, are known to undergo dynamic, regulated changes over the course of the cell cycle [38]. On the other hand, it's unclear the extent to which the fluorescent protein construct used in our datasets is the best proxy for cell size without direct measurements of cell mass. Therefore, further experimental and analytical studies are required to reconcile these results and understand whether the type of cell size measurement has any affect on mechanistic conclusions about size control.

In addition to these contributions, we have developed a novel hierarchical model of budding yeast cell-cycle progression at the single-cell level. The model enables natural pooling of replicate lineage data and separation of cell-to-cell variation from measurement noise. In uncovering correlations in cell-cycle progression between wild-type mother cycles in glycerol/ethanol, our model highlights a potential need for considering *between* as well as *within* cell dependencies. If capturing cell-to-cell dependence is an experimental goal, then time-lapse microscopy is preferable over techniques involving fixed and independent samples taken over time from an initially synchronized population of cells. From a statistical perspective, observing more lineages and more generations per lineage allows for better characterization of this dependence. Our analysis also suggests that both single-cell experimentalists and modelers (regardless of organism or biological processes of study) should at least consider the possibility of such dependence to avoid potential confounding of other observed between-cell or within-cell dependencies. Dependence between cell growth and cell division has been thoroughly studied *within* a given cell cycle. However, the possibility that this correlation might be mediated by inter-cycle dependencies brought about by changes in environment or nutrient availability cannot be ignored. Correlations within a cycle could disappear or decrease in magnitude when conditioning on characteristics of the previous cell or generation. Conversely, as in the case of the dependence we observed in our Bayesian regression analyses, conditioning on additional variables from previous cell cycles might not affect an observed correlation, providing greater context for experimental follow-up or model construction. In either case, our analysis demonstrates that both experimentalists and modelers can benefit from considering multi-generational data acquisition and analysis to verify the robustness of their within-cycle correlation inferences.

Our model provides a flexible and extensible platform for analysis of intra-cycle as well as inter-cycle dependencies. The hierarchical specification of our model and our Bayesian approach to inference easily accommodates new lineage information. We also note that the model is not limited to cell division observations and can be adapted to the statistical analysis of dependence in any biological process. In addition, while we made use of the single-cell growth data in our regression analysis, we are finalizing development of extensions to the hierarchical model to formally fit both the growth and division measurements in a joint analysis, making for a powerful tool to estimate correlations between multiple biological processes while accounting for dependencies between cells in a lineage.

This work represents an important step towards understanding the dependencies in cell-cycle progression and cell growth within and across cells in a dividing population. The statistical model-based approaches described here—coupled with ongoing time-lapse microscopy studies—will shed new light on cell-cycle and cell growth regulation and reveal mechanistic insights about the coordination between these two fundamental biological processes.

bioRxiv preprint doi: https://doi.org/10.1101/082735; this version posted November 24, 2016. 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-NC-ND 4.0 International license.

V. ACKNOWLEDGMENTS

We would like to thank Stefano Di Talia and Frederick Cross; Sung Sik Lee and Matthias Heinemann for generously providing their data; and Merlise Clyde, Stefano Di Talia, Steve Haase, Daniel Lew, Nick Buchler, Bruce Futcher, Kurt Schmoller, Ivan Surovtsev, members of the Haase and Hartemink labs, and the anonymous reviewers for comments about the analysis and the manuscript. This work was funded in part by grants from NIH (P50 30 GM081883-01) and DARPA (HR0011-09-1-0040 to A.J.H.). This work was performed in part under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344 (LLNL-JRNL-702334-DRAFT).

---

[1] Johnston, G.C., Pringle, J.R., & Hartwell, L.H. 1977. Coordination of growth with cell division in the yeast *Saccharomyces cerevisiae*. *Exp Cell Res*. **105**, 79-98. (DOI: 10.1016/0014-4827(77)90154-9)
[2] Di Talia, S., Skotheim, J.M., Bean, J.M., Siggia, E.D., & Cross, F.R. 2007. The effects of molecular noise and size control on variability in the budding yeast cell cycle. *Nature*. **448**, 947-951. (DOI: 10.1038/nature06072)
[3] Goranov, A.I., Cook, M., Ricicova, M., Ben-Ari, G., Gonzalez, C., Hansen, C., Tyers, M., & Amon, A. 2009. The rate of cell growth is governed by cell cycle stage. *Genes Dev*. **23**, 1408-1422. (DOI: 10.1101/gad.1777309)
[4] Ferrezuelo, F., Colomina, N., Palmisano, A., Gari, E., Gallego, C., Csikasz-Nagy, A., & Aldea, M. 2012. The critical size is set at a single-cell level by growth rate to attain homeostasis and adaptation. *Nat Commun*. **3**, 1012. (DOI: 10.1038/ncomms2015)
[5] Schmoller, K.M., Turner, J.J., Koivomagi, M. & Skotheim, J.M. 2015. Dilution of the cell cycle inhibitor Whi5 controls budding-yeast cell size. *Nature*. **526**, 268-272. (DOI:10.1038/nature14908)
[6] Soifer, I., Robert, L., & Amir, A. 2016. Single-Cell Analysis of Growth in Budding Yeast and Bacteria Reveals a Common Size Regulation Strategy. *Curr Biol*. **26**, 356-361. (DOI: 10.1016/j.cub.2015.11.067)
[7] Turner, J.J., Ewald, J.C., & Skotheim, J.M. 2012. Cell size control in yeast. *Curr Biol*. **22**, R350-R359. (DOI: 10.1016/j.cub.2012.02.041)
[8] Hartwell, L.H. 1974. *Saccharomyces cerevisiae* cell cycle. *Bacter Rev*. **38**, 164-198.
[9] Hartwell, L.H. & Unger, M.W. 1977. Unequal division in *Saccharomyces cerevisiae* and its implications for the control of cell division. *J Cell Biol*. **75**, 422-435. (DOI: 10.1083/jcb.75.2.422)
[10] Rupes, I. 2002. Checking cell size in yeast. *Trends Genet*. **18**, 479-485. (DOI: 10.1016/S0168-9525(02)02745-2)
[11] Campos, M., Surovtsev, I.V., Kato, S., Paintdakhi, A., Beltran, B., Ebmeier, S.E., Jacobs-Wagner, C. 2014. A Constant Size Extension Drives Bacterial Cell Size Homeostasis. *Cell*. **159**, 1433-1446. (DOI: 10.1016/j.cell.2014.11.022)
[12] Taheri-Araghi, S., Bradde, S., Sauls, J.T., Hill, N.S., Levin, P.A., Paulsson, J., Vergassola, M., Jun, S. 2015. Cell-Size Control and Homeostasis in Bacteria. *Current Biology*. **25**, 385-391. (DOI: http://dx.doi.org/10.1016/j.cub.2014.12.009)
[13] Jun, S., & Taheri-Aghari, S. 2015. Cell-size maintenance: universal strategy revealed. *Trends in Microbiology*. **23**, 4-6. (DOI: http://dx.doi.org/10.1016/j.tim.2014.12.001)
[14] Smith, J.A. & Martin, L. 1973. Do cells cycle? *Proc Natl Acad Sci USA*. **70**, 1263-1267.
[15] Neufeld, T.P., de la Cruz, A.F., Johnston, L.A., & Edgar, B.A. 1998. Coordination of growth and cell division in the *Drosophila* wing. *Cell*. **93**, 1183-1193. (DOI: 10.1016/S0092-8674(00)81462-2)
[16] Hawkins, E.D., Markham, J.F., McGuinness, L.P., & Hodgkin, P.D. 2009. A single-cell pedigree analysis of alternative stochastic lymphocyte fates. *Proc Natl Acad Sci USA*. **106**, 13457-13462. (DOI: 10.1073/pnas.0905629106)
[17] Balbach, S.T., Esteves, T.C., Houghton, F.D., Siatkowski, M., Pfeiffer, M.J., Tsurumi, C., Kanzler, B., Fuellen, G., & Boiani, M. 2012. Nuclear reprogramming: kinetics of cell cycle and metabolic progression as determinants of success. *PLoS ONE*. **7**, e35322. (DOI: 10.1371/journal.pone.0035322)
[18] Pauklin, S., & Vallier, L. 2013. The cell-cycle state of stem cells determines cell fate propensity. *Cell*. **155**, 135-147. (DOI: 10.1016/j.cell.2013.08.031)
[19] Powell, E.O. 1958. An outline of the pattern of bacterial generation times. *J Gen Microbiol*. **18**, 382-417. (DOI: 10.1099/00221287-18-2-382)
[20] Galton, F. 1894. Natural Inheritance (5th ed.). New York: Macmillan and Company.
[21] Pearson, K. 1896. Mathematical contributions to the theory of evolution. III. Regression, heredity and panmixia. *Phil Trans R Soc London*. **187**, 253-318.
[22] Stewart, E.J., Madden, R., Paul, G., & Taddei, F. 2005. Aging and death in an organism that reproduces by morphologically symmetric division. *PLoS Biol*. **3**, e45. (DOI: 10.1371/journal.pbio.0030045)
[23] Cowan, R. & Staudte, R. 1986. The bifurcating autoregression model in cell lineage studies. *Biometrics*. **42**, 769-783.
[24] Huggins, R.M. & Basawa, I.V. 1999. Extensions of the bifurcating autoregressive model for cell lineage studies. *J Appl Prob*. **36**, 1225-1233. (DOI: 10.1239/jap/1032374768)
[25] Clyde, M.A., Ghosh, J., & Littman, M.L. 2011. Bayesian adaptive sampling for variable selection and model averaging. *J Comput Graph Stat*. **20**, 80-101. (DOI: 10.1198/jcgs.2010.09049)
[26] Greenland, S. 2000. Principles of multilevel modelling. *Intl J Epidemiol*. **29**, 158-167. (DOI: 10.1093/ije/29.1.158)
[27] Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., & Rubin, D.B. 2013. *Bayesian Data Analysis*, 3rd edn.

bioRxiv preprint doi: https://doi.org/10.1101/082735; this version posted November 24, 2016. 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-NC-ND 4.0 International license.

Boca Raton: CRC Press.

[28] Bi, E., Maddox, P., Lew, D.J., Salmon, E.D., McMillan, J.N., Yeh, E., & Pringle, J.R. 1998. Involvement of an actomyosin contractile ring in *Saccharomyces cerevisiae* cytokinesis. *J Cell Biol*. **142**, 1301-1312. (DOI: 10.1083/jcb.142.5.1301)

[29] Kass, R.E. & Raftery, A.E. 1995. Bayes factors. *JASA*. **90**, 773-795. (DOI: 10.1080/01621459.1995.10476572)

[30] Dowling, M.R., Kan, A., Heinzel, S., Zhou, J.H.S., Marchingo, J.M., Wellard, C.J., Markham, J.F., Hodgkin, P.D. 2014. Stretched cell cycle model for proliferating lymphocytes. *Proc Natl Acad Sci USA*. **111**, 6377-6382. (DOI: 10.1073/pnas.1322420111)

[31] Oguz, C., Palmisano, A., Laomettachit, T., Watson, L.T., Baumann, W.T., & Tyson, J.J. 2014. A stochastic model correctly predicts changes in budding yeast cell cycle dynamics upon periodic expression of CLN2. *PLoS ONE*. **9**, e96726. (DOI: 10.1371/journal.pone.0096726)

[32] Cross, F. & Blake, C. 1993. The yeast Cln3 protein is an unstable activator of Cdc28. *Mol Cell Biol*. **13**, 3266-3271. (DOI: 10.1128/MCB.13.6.3266)

[33] Broach, J.R. 2012. Nutritional control of growth and development in yeast. *Genetics* **192**, 73-105. (DOI: 10.1534/genetics.111.135731.)

[34] Lord, P.G. & Wheals, A.E. 1981. Variability in individual cell cycles of *Saccharomyces cerevisiae*. *J Cell Sci*. **50**, 361-376.

[35] Charvin, G., Cross, F.R., & Siggia, E.D. 2009. Forced periodic expression of G1 cyclins phase-locks the budding yeast cell cycle. *Proc Natl Acad Sci USA*. **106**, 6632-6637. (DOI: 10.1073/pnas.0809227106)

[36] Sidorova, J.M. & Breeden, L.L. 2002. Precocious S-phase entry in budding yeast prolongs replicative state and increases dependence upon Rad53 for viability. *Genetics*. **160**, 123-136.

[37] McNulty, J.J. & Lew, D.J. 2005. Swe1p responds to cytoskeletal perturbation, not bud size, in *S. cerevisiae*. *Curr Biol*. **15**, 2190-2198. (DOI: 10.1016/j.cub.2005.11.039)

[38] Weisman, L.S. 2003. Yeast vacuole inheritance and dynamics. *Annu Rev Genet*. **37**, 435-460. (DOI: 10.1146/annurev.genet.37.050203.103207)