PLOS | GENETICS

RESEARCH ARTICLE

Parallel reorganization of protein function in the spindle checkpoint pathway through evolutionary paths in the fitness landscape that appear neutral in laboratory experiments

Alex N. Nguyen Ba[1,2¤], Bob Strome[1], Selma Osman[1], Elizabeth-Ann Legere[3], Taraneh Zarin[1], Alan M. Moses[1,2,3]*

1 Department of Cell & Systems Biology, University of Toronto, Toronto, Ontario, Canada, **2** Center for Analysis of Evolution and Function, University of Toronto, Toronto, Ontario, Canada, **3** Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada

¤ Current address: Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, United States of America
* alan.moses@utoronto.ca

Check for updates

OPEN ACCESS

Citation: Nguyen Ba AN, Strome B, Osman S, Legere E-A, Zarin T, Moses AM (2017) Parallel reorganization of protein function in the spindle checkpoint pathway through evolutionary paths in the fitness landscape that appear neutral in laboratory experiments. PLoS Genet 13(4): e1006735. https://doi.org/10.1371/journal.pgen.1006735

Editor: Justin C. Fay, Washington University School of Medicine, UNITED STATES

Received: November 7, 2016

Accepted: April 5, 2017

Published: April 14, 2017

Copyright: © 2017 Nguyen Ba et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: ANNB is funded by a postgraduate scholarship from the Natural Sciences and Engineering Research Council of Canada (NSERC). AMM is supported by a NSERC Discovery grant and Canadian Institutes of Health research (grant MOP-119579).The funders had no role in study

Abstract

Regulatory networks often increase in complexity during evolution through gene duplication and divergence of component proteins. Two models that explain this increase in complexity are: 1) adaptive changes after gene duplication, such as resolution of adaptive conflicts, and 2) non-adaptive processes such as duplication, degeneration and complementation. Both of these models predict complementary changes in the retained duplicates, but they can be distinguished by direct fitness measurements in organisms with short generation times. Previously, it has been observed that repeated duplication of an essential protein in the spindle checkpoint pathway has occurred multiple times over the eukaryotic tree of life, leading to convergent protein domain organization in its duplicates. Here, we replace the paralog pair in *S. cerevisiae* with a single-copy protein from a species that did not undergo gene duplication. Surprisingly, using quantitative fitness measurements in laboratory conditions stressful for the spindle-checkpoint pathway, we find no evidence that reorganization of protein function after gene duplication is beneficial. We then reconstruct several evolutionary intermediates from the inferred ancestral network to the extant one, and find that, at the resolution of our assay, there exist stepwise mutational paths from the single protein to the divergent pair of extant proteins with no apparent fitness defects. Parallel evolution has been taken as strong evidence for natural selection, but our results suggest that even in these cases, reorganization of protein function after gene duplication may be explained by neutral processes.

Author summary

Parallel evolution of protein domain organization following gene duplication has been demonstrated in the spindle checkpoint pathway leading to the hypothesis that this organization is likely to be adaptive. We test this hypothesis by reconstructing budding yeast

design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

strains with a spindle checkpoint pathway containing a protein with ancestral domain organization, and systematically perform stepwise duplication, degeneration and complementation of the duplicated protein. We show that, under laboratory conditions where the spindle checkpoint pathway is necessary for growth, degeneration of the ancestral pathway organization to the extant sub-functionalized proteins is consistent with a neutral model of duplication-degeneration-complementation.

Introduction

Gene duplication and sub-/neo-functionalization are processes that both increase genomic complexity and are thought to be the major sources of genetic novelty in organisms [1]. Models seeking to explain the retention of paralog genes include relaxation of selection constraints leading to functional novelty through transcriptional changes, splicing divergence or functional repartitioning of protein domains. Particularly interesting are cases of repeated, parallel evolution of these paralog genes that has been taken to be strong evidence for natural selection (several examples reviewed in [2]). In most model organisms, the spindle checkpoint pathway includes the paralogous Bub1 and Mad3 proteins. Surprisingly, although Bub1 and Mad3 are highly diverged, it was recently reported that this paralog pair has arisen from at least nine independent duplication events throughout the tree of life [3]. Perhaps even more striking is that in each event, the duplication almost always leads to a Bub1 homolog, containing a kinase domain and a Mad3 homolog containing a pseudo-substrate APC inhibitor motif [3]. This reorganization of protein function has therefore been thought to be adaptive, leading to a more complex spindle checkpoint pathway [3,4]. However, theoretical work suggests that a causal link between increased genetic complexity and adaptation may not be as clear as commonly assumed [5,6]. For example, if degeneration and complementation of the ancestral bi-functional protein must always lead to Bub1 and Mad3 division of function (i.e., kinase function separated from the pseudo-substrate inhibitor motif), then the repeated observation of Bub1 and Mad3 protein organization might be due to a rate of degeneration that is higher than the rate of reconstitution.

Whether sub-functionalization of a bi-functional protein is adaptive or neutral cannot be easily distinguished by sequence analysis alone. Both models predict complementary conservation of function. However, precise quantitative fitness measurements in organisms with short generation time, such as yeasts, can be used to address questions about these two models of sequence evolution (adaptive vs neutral). Using comparative approaches that delineate functional regions of proteins, we found that this gene duplication leads to sequence signatures that indicate repartitioning of the ancestral function in the extant paralogs. To test whether this increase in complexity (from a pathway containing a bi-functional gene to a pathway containing two specialized genes) in the spindle checkpoint is adaptive, we employ high-throughput quantitative fitness measurements in laboratory conditions stressful for the spindle checkpoint pathway and we systematically dissect the reorganization in protein architecture of an 'ancestral' bi-functional gene. To our surprise, we could not detect any increase in fitness for the stepwise functional reorganization (duplication, degeneration and complementation (DDC)).

Results

Subfunctionalization in the spindle checkpoint network

Preservation of duplicate genes as explained by the duplication-degeneration-complementation model is a neutral process by which repartitioning of the functions of the ancestral protein

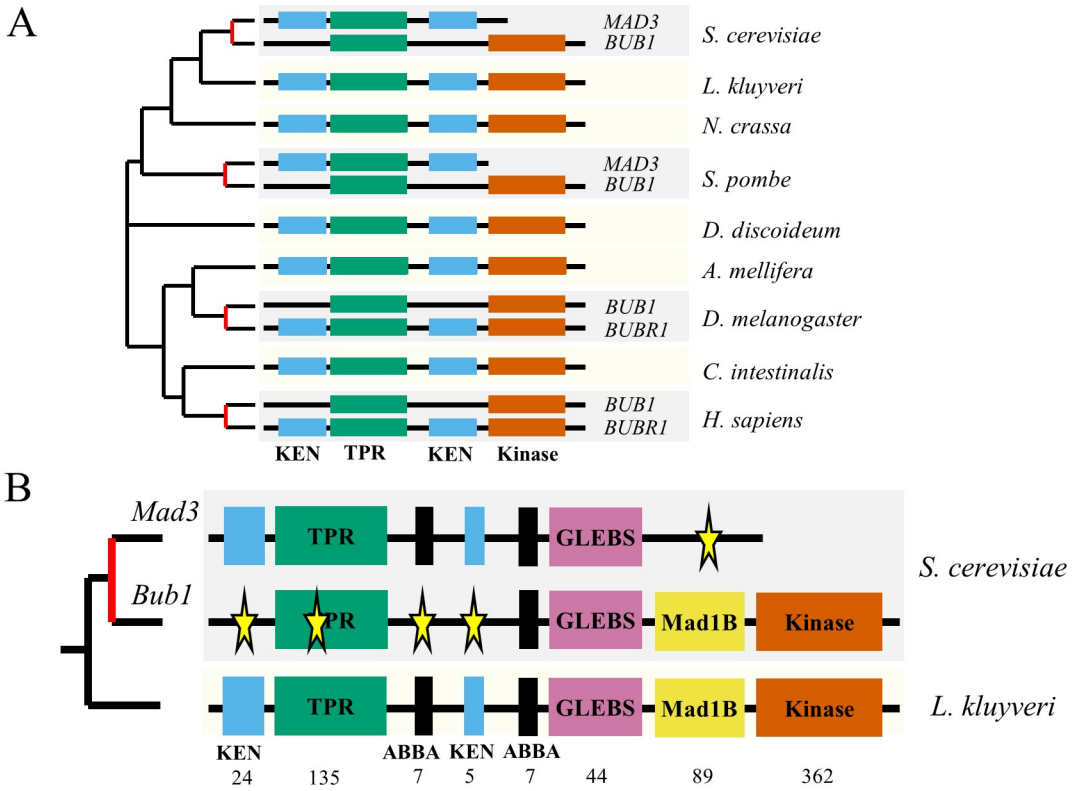

Fig 1. Sequence analysis of the Mad3/Bub1 paralogs. A) Independent duplication events in the bi-functional gene lead to similar domain arrangements. Red line on the phylogenetic tree indicates an inferred duplication event (species highlighted in dark grey). Several species retain the bi-functional gene (species highlighted in light grey). We note that duplicate protein with the KEN boxes in human has been shown to contain a pseudokinase and that this 'MAD3'-like homolog has been named BUBR1 in higher eukaryotes. Schematics are not to scale. B) Amino-acid resolution analysis reveals several regions with changes in constraints in both Bub1 and Mad3 proteins. KEN: KEN-box, TPR: Tetratricopeptide-domain, ABBA: Cdc20 binding motif, GLEBS: Gle2-binding-sequence-domain (binds Bub3), MAD1B: Mad1-binding region, Kinase: Kinase-domain. Numbers indicate the amino acid size of each segment of *L. kluyveri*. The whole protein is 982 amino acids. Stars indicate detected changes in evolutionary constraints.

https://doi.org/10.1371/journal.pgen.1006735.g001

occurs within the two extant duplicates [7]. Although sub-functionalization is a neutral process, it can lead to adaptation in the case of adaptive conflicts: mutations that are precluded from occurring in the 'ancestral' gene but are adaptive when the functions of the ancestral gene are separated. This resolution of adaptive conflicts can sometimes explain the fitness benefit of mutations that 'specialize' multifunctional proteins [8]. Nevertheless, complementation by functional repartitioning or complete functional preservation in one of the two copies is still expected in the cases where mutations leading to neo-functionalization have occurred.

It has previously been shown that Bub1 and Mad3 protein reorganization had occurred several times over the eukaryotic tree of life, leading to similar sequence profiles (Fig 1A, and see [3]). To obtain an amino acid resolution view of the evolution of the paralogous proteins, we performed sequence analysis of the duplication that occurred in the whole-genome duplication of budding yeasts ([9], see Methods). Briefly, a phylogenetic hidden Markov model is used to identify short or large regions of conservation relative to the flanking amino acid sequence. This algorithm is applied to single copy proteins from species that diverged before the duplication to identify short motifs or domains that are under selective constraints. We then map these regions to the duplicates and test whether there is evidence of relaxation in constraints in

the clade post-duplication using likelihood ratio tests. This analysis revealed several protein regions, in addition to the KEN boxes and kinase domain identified in previous studies [3], that have repartitioned from the ancestral sequence to the paralogs (Fig 1B).

This more detailed view of the changes in constraints allowed us to propose why some of the changes were correlated (S1A Fig for an example). For example, Mad3 is known to interact with Cdc20 via its N-terminal KEN box motif and its tetracopeptide repeat domain (TPR) [10]. According to 3D structural information of Mad3 (S1B Fig), specific residues in the TPR domain (shown in yellow in S1B Fig) appear to contact the KEN box (shown in red in S1B Fig) and may stabilize the interaction of Mad3 with Cdc20 (S1C Fig) [11]. Bub1, on the other hand, does not possess the N-terminal KEN box but still requires the TPR for proper function [12]. We can detect changes in constraints on these specific residues of the TPR of Bub1, suggesting that the degeneration of residues in the TPR domain or the KEN box will disrupt the same function (binding to Cdc20) and that loss of selection constraints on that function will lead to degeneration of both. Indeed, the same correlated changes in constraints exemplified here are also observed in the mammalian, drosophila, and fission yeast Bub1 and Mad3, which occurred through independent gene duplication events (S1D Fig). The changes in constraints at these specific residues are unlikely to disrupt the other functions of the TPR.

On the other hand, the pattern of evolution on the ABBA motifs in the yeast paralogs is more complicated. The ABBA motif is known to be required for binding to Cdc20 [13], and indeed, both Bub1 and Mad3 retain at least one copy of the ABBA motif (Fig 1B). Nevertheless, we can clearly detect motif turnover in the first ABBA motif in Bub1 (Fig 1B). This suggests two possible explanations: 1) the two ABBA motifs serve the same function and the loss of the first motif was compensated in the Bub1 protein by another motif, 2) the two ABBA motifs may bind Cdc20 for different functional reasons and the Bub1 protein underwent a single change in constraint on this motif. Since we have not identified a newly conserved motif in the Bub1 protein (which could compensate for the loss of the first ABBA motif), we believe that the second possibility is more parsimonious.

Consistent with important functions for the domains identified in the ancestral fungal protein, there were no conserved regions that were lost in both paralogs (Fig 1B). The repartitioning of functional regions suggests that the duplication lead to sub-functionalization. We therefore sought to experimentally verify that sub-functionalization had occurred in the paralog pair. As a proxy for the 'ancestral gene', we obtained the gene from *Lachancea kluyveri* which diverged prior to the whole-genome duplication event (see Discussion). We refer to this gene as the 'single-copy protein (SCP)', and transformed it into *S. cerevisiae*.

We first assessed the localization of the single-copy protein because it was known that Bub1 and Mad3 localize to different subcellular compartments: Bub1 is localized to the kinetochores in a Mps1 dependent manner during specific phases of the cell-cycle [14] and Mad3 is constitutively localized in the nucleus [15]. If Bub1 and Mad3 were products of a sub-functionalization event, we would also expect the *L. kluyveri* protein to localize to both subcellular compartments. To test this, we tagged the three proteins with GFP and visualized their localization by fluorescence microscopy. To account for possible differences in growth or imaging conditions that may influence the apparent localization of the proteins, we designed an assay where the tagged single-copy protein could be visualized alongside the *S. cerevisiae* tagged protein in the same field of view using the identical GFP tag. Briefly, we expressed additional cytoplasmic fluorescent proteins of other colors (non-GFP) that allow us to differentiate which cells carried the *L. kluyveri* protein fusion or the *S. cerevisiae* fusion and obtained fluorescent micrographs of mixed cultures (see Methods). As expected, we observed a distinct localization pattern for the single-copy protein, which was quantitatively different from either Bub1 or Mad3 (Fig 2A). Upon closer inspection, we found that the localization of the single-copy

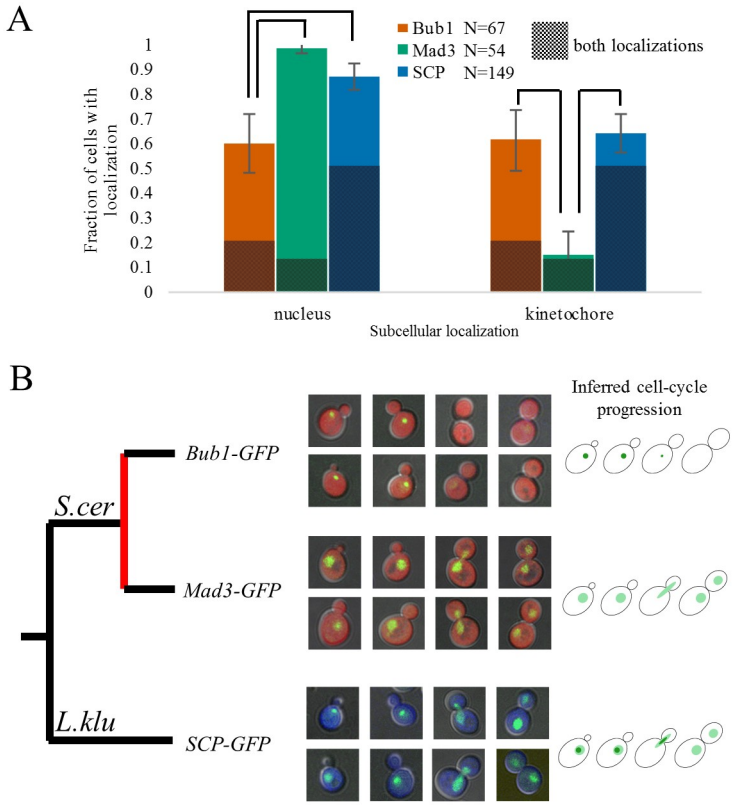

Fig 2. Localization of the Bub1/Mad3 paralogs. A) Green-fluorescent protein tagged gene cloned from *L. kluyveri* (single-copy protein, SCP) localizes differently than either the Bub1 or Mad3 proteins from *S. cerevisiae*. In both cases, the single-copy protein is in a background expressing a cytoplasmic blue fluorescent protein, while the *S. cerevisiae* protein is in a background expressing a red fluorescent protein. Quantification of the localization pattern was performed by visual inspection of the green fluorescent protein of cells grown in co-culture, and then the strains were classified according to their blue or red fluorescence. Cells without obvious GFP localization were discarded prior to the analysis (in total 26 cells were discarded from the Bub1-GFP strain, 7 from the Mad3-GFP strain, and 11 from the SCP-GFP strain). The shaded section of the bars represents cells for which we found localization in both nucleus and kinetochore puncta and are therefore present in both bar graphs. Error bars represent 95% confidence intervals of the fraction as obtained by bootstrap replicates. Lines over the bar chart represent $p<0.05$ (Fisher's exact test corrected for multiple testing). B) Green-fluorescent protein tagged genes cloned from *L. kluyveri* (SCP) localizes as a mixture of Bub1 and Mad3 protein from *S. cerevisiae*. Bub1 is clearly visible as puncta in S/G2 phase of the cell cycle, presumably when it localizes to the kinetochores, while Mad3 is seen as nuclear throughout the cell-cycle. The single-copy protein is also seen as nuclear throughout the cell-cycle, however is also seen as puncta in S/G2.

https://doi.org/10.1371/journal.pgen.1006735.g002

protein appears to be a mixture of the localization patterns of Bub1 and Mad3. Consistent with this, when ordering by bud size as marker of cell stage, we observed that the *L. kluyveri* protein localized constitutively to the nucleus, and also showed a punctate localization within the nucleus in the early cell-cycle (Fig 2B).

Because Bub1 and Mad3 are required for the spindle-checkpoint pathway, we tested whether the single-copy protein could rescue defects of cells lacking Bub1 or Mad3 with respect to the functionality of the pathway. To do so, we took advantage of a strain carrying a non-essential chromosome containing the ochre suppressor SUP11 [16]. This strain forms red-pigmented colonies upon loss of this chromosome due to the presence of the ade2-101 allele. As has been reported previously [17], we found that strains without a functional Bub1 gene have a very strong increase in chromosome loss rate (29/65 vs 3/131 sectored colonies, p-

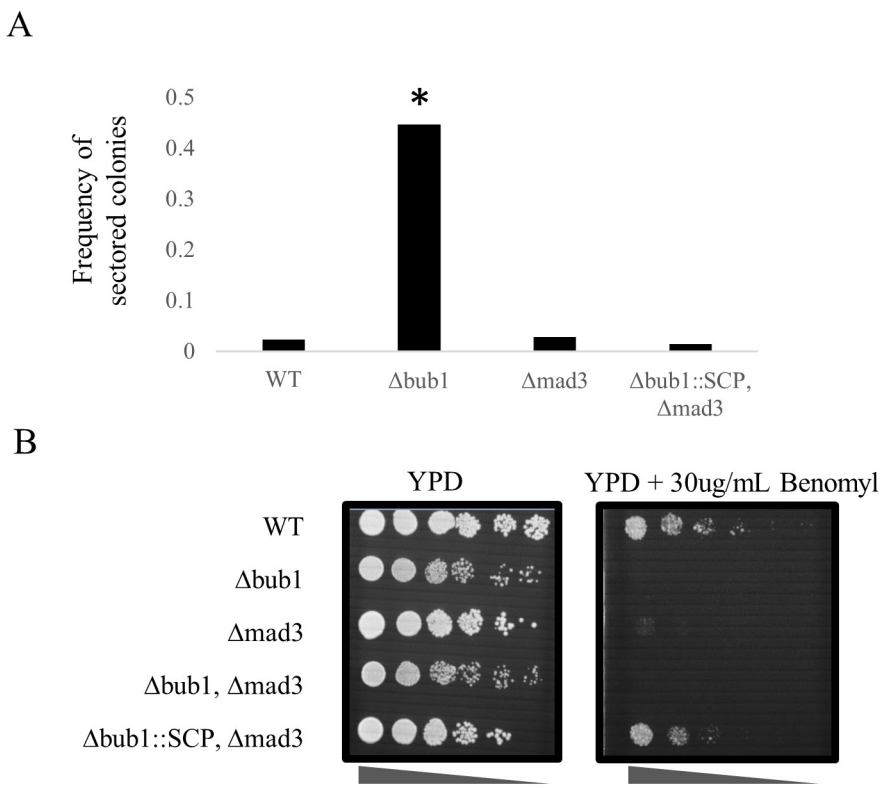

Fig 3. Phenotype rescue of spindle checkpoint mutants. A) Yeast cells carrying the non-essential chromosome loss reporter were plated on agar plates with low adenine supplementation. The relative increase in sectored colonies after two days of growth (sectored colonies/white colonies) is indicative of higher chromosome loss rate, which is a phenotype easily visible in cells lacking Bub1. In contrast, cells carrying the SCP gene have a loss rate that is not distinguishable from wild-type. * indicates statistical significance P<0.001. B) Yeast cells expressing the single-copy protein (SCP) at the BUB1 locus can grow on YPD plates with 30ug/mL benomyl even in the absence of Mad3, as assessed by a 10-fold spot dilution assay. YPD plate was imaged after 2 days. YPD plate containing benomyl was imaged after 3 days.

https://doi.org/10.1371/journal.pgen.1006735.g003

value < 0.0001). However, we were unable to detect an increase in chromosome loss rate for strains lacking a functional Mad3 gene (5/177 vs 3/131 sectored colonies). Nevertheless, cells lacking Bub1 and Mad3, but carrying the single-copy gene driven by the Bub1 promoter in this strain completely abolished the Bub1 chromosome segregation fidelity defect (3/207 vs 3/131 sectored colonies, Fig 3A). We next sought to verify if the single-copy protein could rescue the growth defects of impaired cell-cycle functions of cells lacking Bub1 and Mad3. Cells lacking Bub1 or Mad3 are highly sensitive to benomyl, a microtubule destabilizing drug [18], because cells fail to correctly detect mitotic spindle attachments. If the single-copy protein can perform the functions of both Bub1 and Mad3, we expect that the single-copy protein would rescue the fitness defect of cells lacking Bub1 or Mad3. If the phenotype is not fully rescued, it may indicate neo-functionalization and adaptation in the Bub1 or Mad3 protein (or it may be due to an artifact of expressing a heterologous gene). To test this, we performed spot dilution assays and found that, if expressing the single-copy protein, cells lacking Bub1 and Mad3 can grow in the presence of benomyl with similar growth characteristics to wild-type *S. cerevisiae* cells (Fig 3B). We could not simply test the function of the spindle checkpoint in other species as we observed that yeasts other than *S. cerevisiae* were highly resistant to benomyl (S2 Fig).

At the limits of the resolution of this plate assay (we estimate that this assay can detect at the minimum only a 5% growth difference), and taken together with the localization data, these

results suggest that the single copy protein can rescue Bub1 and Mad3 function in *S. cerevisiae*. Consistent with the DDC model, the sub-functionalization of the Bub1 and Mad3 proteins may confer no fitness advantage to *S. cerevisiae*.

A precise and rapid quantitative fitness assay for gene network rewiring

It is possible that the spot dilution assay did not have enough resolution to detect the fitness advantage of the reorganization of protein function in the paralogs. To address this, we designed a method to more precisely quantify relative selection coefficients (*s*, see Methods). Briefly, our assay is a competitive fitness assay that takes advantage of flow-cytometry to provide relative counts of fluorescently labeled cells within a growing population [19] and this assay can be performed in a high-throughput fashion where the combination of alleles of interest are tested in cellular backgrounds expressing different fluorescent proteins as experimental replication. Replicate strains can be rapidly created by the SGA cloning strategy (see Methods and [20], S3 Fig). To determine the resolution limit of this fitness assay, we generated 16 spores from a cross between a strain carrying a wild-type Bub1 allele marked with CaURA3 (the *URA3* gene from *Candida albicans*, which complements the *URA3* gene from *S. cerevisiae*) and a query strain (containing SGA mating type reporters) expressing a green or a red fluorescent protein (8 spores for each cross). These spores were competed to form 64 fitness assays. We reasoned that if any additional single nucleotide polymorphisms (SNPs) between the query strains and parental strain of our mutant arrays had a detectable fitness effect (but masked due to epistasis within their respective backgrounds), or if these SNPs showed non-transitive fitness effects, they would be uncovered within these 16 spores. We compared the relative proportions of cells expressing the green fluorescent protein and red fluorescent protein at the 20th generation to the 40th generation and we calculated the selection coefficient for each competition (see Methods). Because our strains are supposedly genetically identical (except for the possibility of non-shared SNPs), we expect an average selection coefficient of zero and the standard deviation obtained from this test can be used to estimate the resolution of our assay (deviations due to growth conditions or to the non-shared SNPs). We measured the selection coefficients of the wild-type strains and observed a mean selection coefficient of 0.00069, with a standard deviation of 0.0017 (Fig 4A). This indicates that the resolution of our assay is in the order of $s = 0.0033$ (1.96 times the standard deviation) and we believe this represents the difference in growth rate that we can detect. To account for other possible variations that may occur during the course of the study (changes in media, etc) we therefore chose to report as deleterious/beneficial any differences in fitness where both replicates of a competition exceeded a selection coefficient with an absolute value of 0.005 or greater while remaining consistent with all other competitions.

We next compared strains carrying deletions of Bub1 or Mad3, or single-copy protein rescues of these deletions to wild-type strains and found that cells lacking Bub1 have a strong fitness defect when growing on benomyl while cells lacking Mad3 have a more moderate defect ($s<-0.3$ and $s = -0.01317$ respectively, Fig 4A). Remarkably, cells lacking both Bub1 and Mad3, but expressing the single-copy protein at the Bub1 locus, retain wild-type growth rate when challenged with the same benomyl concentration ($s = 0.002$, Fig 4A). Thus, even at the much higher resolution of these quantitative competition experiments, we find no evidence that the duplication and reorganization of the Bub1 and Mad3 proteins (Fig 1) confers a fitness benefit.

No evidence of fitness advantage for the subfunctionalization of the Bub1/Mad3 ancestral protein

The experiments above show that there is apparently no fitness advantage to the Bub1/Mad3 protein organization relative to the single-copy protein. However, the real evolutionary

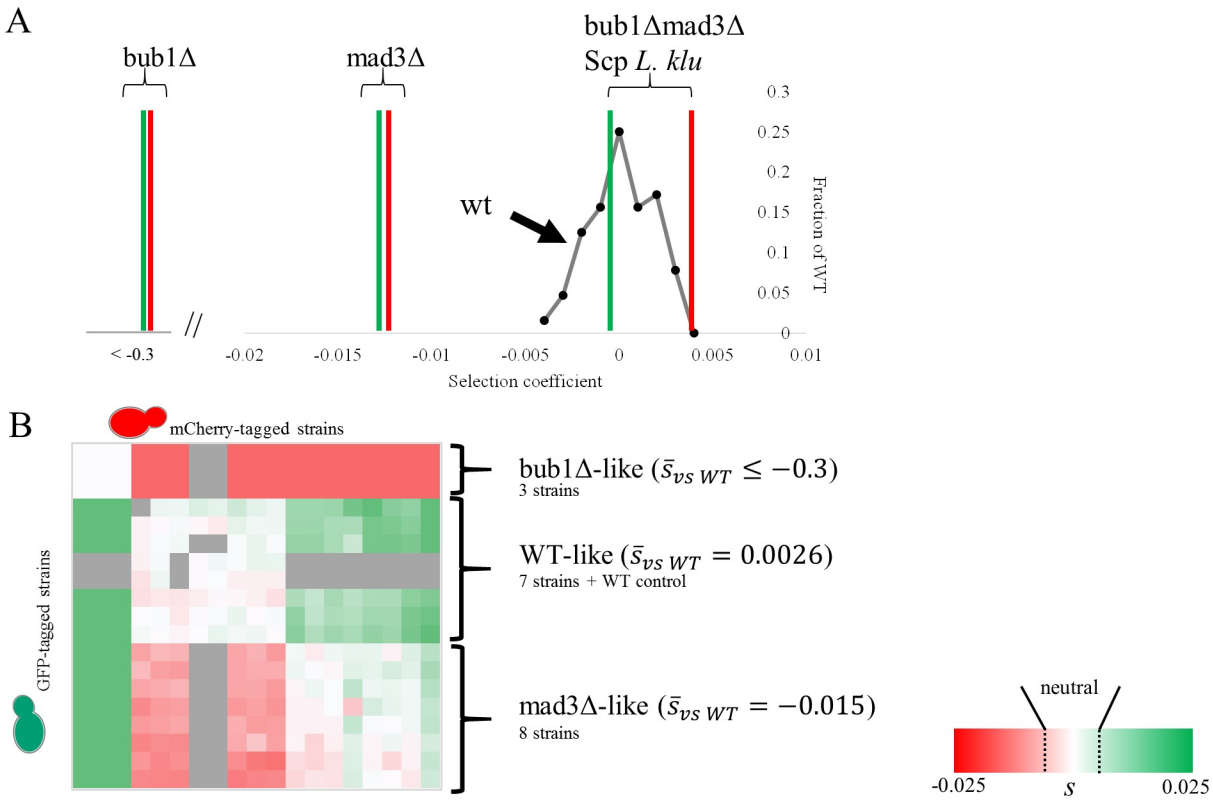

Fig 4. Competitive fitness measurements of reconstructed genotypes. A) The resolution of our assay was assessed by performing 64 competitions of wild-type spores from the SGA process carrying identical marked alleles from a mosaic of BY4741/BY4742 parental background and are assayed using our high-throughput fitness assay (grey curve). We indicate the measured fitness effects of strains carrying mad3Δ, bub1Δ, or the single-copy protein at the *BUB1* locus as red and green vertical bars. The two bars for each genotype represent measurements of relative selection coefficient to the wild-type control from replicate strains carrying swapped fluorescent proteins. B) Heat map clustering of all the competitive fitness assays using Euclidean distance and average linkage of the selection coefficients. The transpose off-diagonals are biological replicates with strains carrying swapped fluorescent proteins. Three major fitness classes are observed, which correspond to genotypes that have fitness comparable to mad3Δ cells, bub1Δ cells, or to the wild-type (as indicated by the mean selection coefficient measured for the strains present in each cluster). Grey cells in the heat map indicate competitions that were not performed.

https://doi.org/10.1371/journal.pgen.1006735.g004

trajectory almost certainly did not replace the single-copy protein directly with the fully formed Bub1 and Mad3. Instead, a series of evolutionary events (likely separated by millions of years) probably occurred from the single-copy protein to the extant paralogs. It is possible that certain steps along this path, particularly at the beginning after the duplication, were advantageous and drove the reorganization of the paralogs.

Because there is insufficient phylogenetic resolution to infer the exact order and number of mutations, based on previous knowledge of the functional elements found in Bub1 and Mad3 (Fig 1), we created strains with genetic make-up of possible evolutionary intermediates that correspond to stepwise mutational events during protein function reorganization. The evolutionary paths assayed include mutations at multiple loci, and therefore possible paths were created using several rounds of the SGA cloning strategy (see Methods). Briefly, we sought to test single and double mutations of the KEN boxes on the SCP at the *BUB1* locus, loss of kinase of the SCP at the *MAD3* locus, non-functionalization of the gene, or "evolution" to the extant protein. These mutations represent key evolutionary steps from the ancestral protein to the extant Bub1 and Mad3 (Fig 5). We introduced these mutations into the SCP and cloned them individually into different starting strains (for example, we generated a strain containing the

SCP with a mutated KEN box at the *BUB1* locus). To combine them, query strains carrying the SGA markers and different fluorophores were crossed to the library in an ordered array and selected such that the final products of several rounds of mating were otherwise genetically identical haploid spores carrying different combinations of marked alleles. We then performed an all-by-all competitive fitness assay, and found that genotypes cluster in three distinct fitness classes, which correspond to cells lacking Mad3 function (Δmad3-like, with an average selection coefficient of -0.015 relative to wild-type), cells lacking Bub1 function (Δbub1-like, with an average selection of <-0.3 relative to wild-type (fitness effects larger than -0.3 are not measurable in this assay and so we consider these genotypes to have fitnesses <-0.3, see Methods), or cells with wild-type phenotype (WT-like, with an average selection coefficient of 0.0026 relative to wild-type) (Fig 4B). Although we have not explored the fitness landscape exhaustively, this sample of genotypes suggests that it is made up of three distinct plateaus.

To assay possible paths through the fitness landscape that lead to sub-functionalization, we focused on genotypes that differed from each other by one 'evolutionary step'. For example, we consider the initial duplication as one evolutionary step (such that we compared the fitness difference of a strain containing one vs two SCP), as well as individual losses of functional motifs. Although in principle it is possible for a non-functional genotype to revert, we consider these events to be rare and therefore have not considered them for simplicity. The results of our analysis are displayed in Fig 5, as an evolutionary landscape with multiple evolutionary paths that 'travel' through possible intermediate genotypes. Interestingly, we were able to find at least one path without a detectable fitness defect consisting of at least three degenerations (Fig 5, path through white nodes). If we only consider paths that do not go through nodes of fitness defects (i.e. we do not allow crossing of fitness valleys), then the analysis suggests that the extant network in *S. cerevisiae* is an absorbing state (see Discussion).

Although we cannot be certain that evolution has taken any of the paths studied here, that we can find at least one seemingly neutral evolutionary path, at the resolution of our assay, strongly supports the DDC hypothesis that the reorganization of protein domains in the Bub1/Mad3 paralogs can be explained by neutral degenerative mutations (Fig 6 and see Discussion).

This model also predicts that the initial neutral step in the process (the duplication) is reversible, and that the mutations must be relatively common to explain the frequency at which the reorganization occurs during evolution (Fig 1). Consistent with this, we have identified at least one phylogenetic clade where the gene duplication in the ancestor of Saccharomyces reverted to the single-copy functional homolog: *Vanderwaltozyma polysporus* retains a single-copy protein with all the functional elements of Bub1 and Mad3 even though it diverged after the whole-genome duplication (S4 Fig). It is estimated that the ancestor to the lineages leading to *V. polysporus* and *S. cerevisiae* had already non-functionalized ~20% of the duplicates [21] suggesting that neither sub-functionalization of Bub1/Mad3 and non-functionalization were rapid.

Discussion

Neutral processes, such as described by the DDC model [7], have been shown to be important in increasing genomic complexity (see for example [22] for a study on non-adaptive increase in interactome complexity). At the limit of the resolution of our assay (discussed below), we find no evidence that the Bub1/Mad3 protein reorganization after duplication in budding yeast provides a route for adaptive conflict resolution and therefore, further work is necessary to find a fitness advantage over a single-copy protein in the spindle checkpoint pathway.

Consistent with the DDC model, none of the evolutionary intermediates that we consider functional through the comparative genomics analysis showed a fitness defect when tested

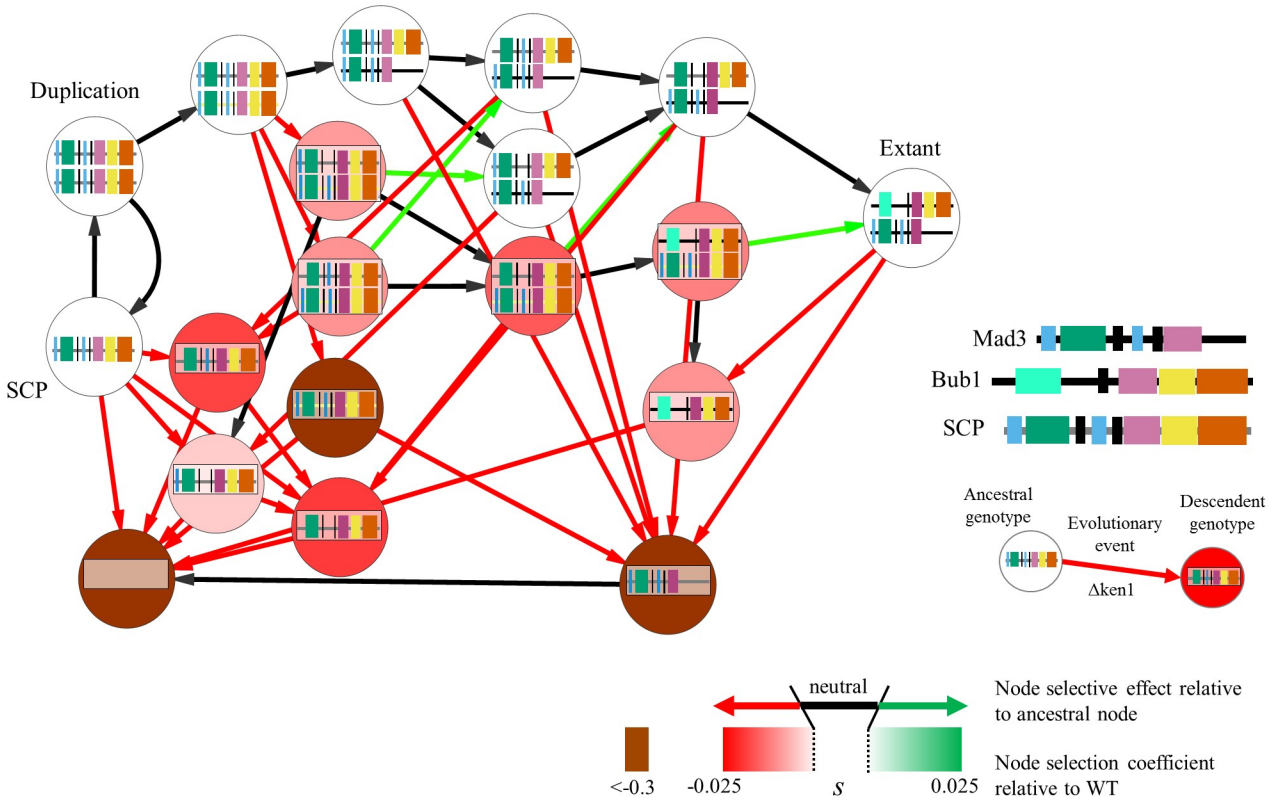

Fig 5. Evolutionary paths during network rewiring. Schematic of the intermediates we have assayed in our experiment if we assume changes in the MAD3 promoter occurred first following the gene duplication event. When the compared genotypes (nodes) are separated by a single mutational event (such as a degeneration of a functional element, which are shown as colored boxes inside the nodes), these nodes are connected by directed edges in the graph (see the example given for a loss of the first KEN box). The color of the node represents the fitness of the genotype when compared to wild-type. The color of the edge corresponds to the fitness of the descendent genotype when compared to the ancestral genotype, where red indicates a deleterious selection coefficient, black a selection coefficient with undetectable fitness effect, and green a positive selection coefficient. The color of the line representing the protein is grey for the SCP under the Bub1 promoter, yellow for the SCP under the Mad3 promoter, and black for the extant *S. cerevisiae* Mad3 and Bub1. Coloured boxes represent functional elements within the proteins (See Fig 1 and text for details).

https://doi.org/10.1371/journal.pgen.1006735.g005

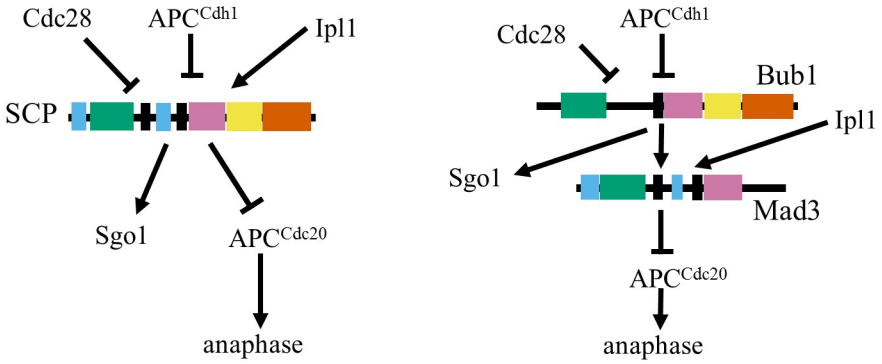

Fig 6. Model for network rewiring. On the left is a schematic of the inferred ancestral spindle checkpoint with the single-copy protein (SCP is the genotype marked as *L. kluyveri* in Fig 1), while on the right is a schematic of the extant *S. cerevisiae* network (genotype same as marked in Fig 1).

https://doi.org/10.1371/journal.pgen.1006735.g006

under laboratory conditions requiring functional spindle checkpoint and when driven by the *BUB1* promoter. Nevertheless, there are several caveats to our experiment. First, the resolution of our assay meant that we could only detect fitness effects in the range of $s = 0.005$. Because of the population size of budding yeasts in nature [23] and our estimated effective population size during the experiment (see Methods), selection could be efficient even on undetected differences in fitness (given a high enough recombination rate). Therefore, it remains possible that the sub-functionalization to Bub1/Mad3 is truly adaptive. However, even if we assume a very small beneficial effect that was not detected, because we performed our assay in the presence of high concentration of benomyl (which was used to characterize all the components of the spindle checkpoint pathway [24]), we believe that under reasonable growth conditions the adaptive effect of this sub-functionalization would be even smaller. Another possibility is that Bub1 and Mad3 participate in completely orthogonal molecular processes to the spindle checkpoint such that our assayed environment would not be able to detect the real functional differences between the duplicate genes and the ancestral single-copy protein. Yet a third possibility is that there exists one context where the effects of the duplication are under much stronger selection. However, if this context exists, it must be rare due to the finding that the phylogenetic clade leading to *V. polysporus* reverted to the single-copy gene. Ultimately, it is not possible to know the environmental context, nor the genetic context of the ancestral yeast and we cannot rule out that adaptation by escape of an adaptive conflict drove the sub-functionalization of the ancestral protein. Despite these caveats, we note that the functionally relevant motifs in Bub1 and Mad3 were identified in the context of the spindle checkpoint pathway, and this pathway is activated every cell division to ensure proper chromosomal attachment prior to anaphase. We therefore believe that large effects would have been captured even in this laboratory environment.

Although in our study all functional evolutionary paths lead to the same genotype through the same number of degenerations, the probability that a path is taken is dependent on the rate of mutation, the selection coefficient of the intermediates and the effective population size [25]. We here discuss only the scenario of large population size because small population sizes would allow all genotypes, including any potential neo-functionalization, to be effectively neutral. When population size is large and mutation rate is low, then crossing a fitness valley (as measured by our quantitative fitness assay) requires the population to fix each intermediate genotype (this is the deleterious sequential fixation regime discussed in [25]). In this large population size and low mutation rate regime, paths through deeper fitness valleys are essentially never taken because selection is very effective. In this regime, the relative frequencies of the effectively neutral states are entirely dependent on the mutational rate between these states [7]. Because the mutational rate to degenerate is likely to be higher than the mutational rate to recreate a functional element, the evolutionary outcome of Bub1/Mad3 homologs after sub-functionalization could be nearly deterministic. Therefore, the relative probability of observing Bub1/Mad3 functional homologs after duplication is equal to the rate of sub-functionalization divided by the rate of non-functionalization [7]. We propose that the rate of sub-functionalization can be high due to the very small number and mechanistically simple degenerations: we showed here that sub-functionalization and degeneration of the Bub1/Mad3 protein can occur within only two mutations following the gene duplication (generation of stop codon to remove the kinase function in Mad3, and a shift in start position to remove the first KEN box in Bub1), both of which have been observed frequently over evolution in other genes [26,27]. If this sub-functionalization is truly neutral, we propose that the repeated reorganization of the Bub1/Mad3 homologs may be due to the fact that no other possible outcome of the duplication can be easily observed (it is an absorbing state), such that the observed genotype is surrounded by fitness valleys or reversion to the single-copy protein. Fixation of the re-organization may

have occurred through hitchhiking with other beneficial mutations, a scenario that often occurs with large population size and a high mutation rate [28], or due to changes in gene expression (discussed below).

Previous studies have considered the functional implications of evolution of individual enzymes and protein complexes through gene duplication and divergence (e.g. [29,30]). However, many proteins that function within regulatory networks contain complex multi-domain architectures and disordered regions [31]. In the case of Bub1 and Mad3, numerous functional steps occurred after gene duplication, and it is not possible to reconstruct them at the resolution of single amino acid substitution. Nevertheless, we chose several key evolutionary steps during the functional reorganization and assessed the fitness of those intermediates. We believe that our study represents a practical way forward for studies of the evolution of complex eukaryotic proteins. Unlike other previous studies (e.g. [29]), we have not performed ancestral gene resurrection via gene synthesis, but instead we have chosen a gene from another species which we believe is representative of the ancestral allele. Our approach has several advantages. First, it is more likely that the gene is functional in at least one genetic environment. Second, the reconstruction of the ancestral gene might not be accurate in proteins with highly diverged disordered regions. Finally, the gene has evolved for the same period of time as the duplicate genes, providing a direct test of whether adaptation from the ancestral allele is due to the resolution of an adaptive conflict. Similar experimental designs have been used to assay functional differences in whole transcriptional regulatory networks that happen on relatively shorter divergence times [32].

Transcriptional evolution has been shown to be an important aspect of functional divergence after gene duplication [33–35]. An important caveat of our study is that we have not tested the evolution at the promoter, due to the difficulty in finding the functional elements of the promoter region [36]. However, we do have evidence that the Mad3 promoter and the Bub1 promoter are not functionally equivalent as the single-copy protein does not fully rescue the spindle-checkpoint defects when placed at the *MAD3* locus (S5 Fig). Those promoter changes, however, are still consistent with a model of degeneration where the Bub1 promoter is similar to the ancestral SCP promoter. On the other hand, it is also possible that after duplication the Bub1 promoter acquires beneficial mutations relative to the ancestral promoter, and these changes drive fixation of the gene duplication. Thus, although there is clear evidence that the promoters of Mad3 and Bub1 have diverged, we did not test the effects of these changes.

Under our tested conditions, our study shows that there is no evidence that the reorganization of protein function was to escape an 'adaptive conflict' for the spindle checkpoint pathway in mitotic cells of *S. cerevisiae*. Our data is consistent with the DDC model and our study suggests that parallel evolution through degenerative processes does not have to be rare or adaptive. Further work will be required to see if this is also the case in other organisms where this reorganization has been observed. This situation is reminiscent to the convergent evolution of holocentric chromosomes across the tree of life [37], and it is still unclear whether it provides an advantage during growth, especially considering the more complex meiotic segregation of chromosomes [38].

Interestingly, although the core spindle checkpoint pathway is conserved in all eukaryotic life, several other differences exist in this pathway [39]. These differences include non-conserved proteins important for spindle checkpoint function (such as p31) or different copy number of paralogous proteins (such as Cdc20 in human). Our study provides experimental techniques to test the step-wise effects of evolutionary changes that have been detected through comparative genomics on multiple loci (such as the ones in the spindle checkpoint). We anticipate that these sensitive quantitative fitness measurements will be useful in the computational

modeling of sequence and protein evolution within the context of a complete regulatory network [40,41], as has been performed on other important regulatory networks [42,43].

Materials and methods

Yeast strains and culturing

All strains were derived from either BY4741 or BY4742 using standard yeast genetic techniques or synthetic gene arrays (see next method subsection). The single-copy gene was PCR amplified from purified genomic DNA (Fermentas, #K0512) of *L. kluyveri* (NRRL Y-12651). All integrations were verified by PCR, and key strains containing the single-copy gene were verified for absence of Bub1 or Mad3 when relevant.

Strain construction for the library of alleles was performed using the same method as described in [44]. SGA query strains were created by transferring the Ste3pr_LEU2 marker from Y8205 into the *CAN1* locus of BY4742. Fluorophores with the Ste2pr_LkHIS3 were cloned into the pAN200a plasmid (based on pFA6a [45]) using standard cloning techniques and transformed into the *CAN1pr* locus using delitto perfetto [46].

Benomyl (10mg/mL DMSO stock) is used at outlined concentrations and added to boiling-hot media until completely dissolved. 5-fluoroorotic acid (5-FOA, 100mg/mL DMSO stock) was used to select against uracil biosynthesis prototrophs [47] and plates were poured at 1g/L 5-FOA final concentration (supplemented with all amino acids, including 72ug/mL uracil). Geneticin (G418) was used at 200ug/mL to select for geneticin resistance.

Synthetic genetic arrays

Cells are grown according to slight modifications to the protocols outlined in [48] as shown as a schematic in S3 Fig. We modified our query strains to have the following cassette integrated at the *CAN1* locus: RPL39pr_fluorophore_Ste2pr_LkHIS3_Ste3pr_LEU2. Fluorophores used for our study were yeast-enhanced monomeric green fluorescent protein (ymEGFP) and yeast-mCherry (ymCherry). For the general construction of our strains, a query strain is first crossed with all the desired alleles at a particular locus. Diploid selection is performed by selecting for complementary auxotrophies. Overnight diploid cells from plate patches are then scraped into liquid sporulation media (1% Potassium acetate, 0.005% Zinc acetate) and supplemented with amino acid requirements for diploid strains at 25% of the normal usage and incubated on a roller wheel for three days at room temperature. Usually, about 30% sporulation is observed and 5ul of the mixture is spread or spotted on selection plates that select for MATα and other selection markers such as auxotrophies and drug markers.

We found that modifications to the germination and outgrowth procedure were necessary in our hands to obtain colonies after the SGA procedure. In our hands, addition of lysine to the media greatly enhances the initial outgrowth of spores for all strains that were constructed using our query strains (even the ones that did not express a fluorophore or were lysine prototrophs). Spore outgrowth was normal for standard SGA query strains, indicating that strain specific variation, or heterozygous *lys2* deletion strains or homozygous *LYP1* affected the outgrowth of our strains (*LYS2*/*Δlys2 LYP1*/*LYP1* compared with *LYS2*/*LYS2 LYP1*/*Δlyp1*). Therefore, to select for lysine prototrophs, the colonies are replicated to media lacking lysine only after the initial growth. To select for lysine auxotrophy, replica plating was used to isolate colonies that did not grow on media lacking lysine, however alpha-aminoadipate could be used instead [49].

The final strains used in the competitive fitness assay all had the following genotypes: MATα, can1::RPL39pr_fluorophore_Ste2pr_LkHIS3_Ste3pr_LEU2, bub1::allele::CaUR-A3MX, mad3::allele::KanMX, Δlys2, Δhis3, Δura3, Δleu2.

Genomic sequences and comparative analyses

Protein sequences used for the comparative analyses were from the Yeast Gene Order Browser [50]. These proteins were then aligned with MAFFT [51]. Genomic sequences for BY4741 and BY4742 were obtained from the Saccharomyces Genome Database [52].

To perform our comparative analyses, protein sequences were analyzed using methods described in [9] and by visual inspection. To test for changes in constraints in the duplicate protein, we first predict regions of conservation in the proteins pre-duplication using a phylogenetic hidden Markov model that detects regions that have significantly lower evolutionary rates relative to their flanking region [53]. Having predicted these conserved regions, we can then map them to the duplicate protein and ask whether two rates of evolution (one for the pre-duplication clade, and one post-duplication) better explain the evolution of the selected region as opposed to a single rate. Because the rate of evolution of predicted conserved regions in the ancestral protein is very low, when two rates of evolution better explain the data, it typically implies that the region under purifying selection prior to the whole-genome duplication is now under relaxed constraints after the whole-genome duplication. Partitioning of these losses in selection constraints in the two paralogs is an indication of sub-functionalization. To identify potential new motifs in the post-duplication proteins, the phylogenetic hidden Markov model can be used on the post-duplicate protein and the same analysis for changes in constraints can be performed.

X-ray crystallography files (3ESL: Bub1 [54] and 4AEZ: Mad3 [11]) were obtained from the Protein Data Bank [55] and analyzed using PyMOL [56].

Chromosome loss assay

Chromosome loss rate was measured in the classical strain carrying a linear non-essential chromosome in the W303 background [16]. Briefly, strains carrying the ochre allele *ade2-101* mutation exhibit a visible red pigment when growing in media with low adenine supplementation. This phenotype is suppressed in strains with a single-copy functional *SUP11*, which is present in the artificial chromosome. Therefore, chromosome loss rate can be measured by counting the number of red-sectored colonies in agar plates containing low adenine. This assay is performed by plating about 200 colonies, and we report the number of sectored colonies over the number of white colonies after two days of growth.

Quantitative fitness assay

Quantitative fitness assays were performed using the MACSQuant VYB (Miltenyi Biotec Inc.). Briefly, strains are grown for 48 hours in 5mL of cultures on a rolling wheel. The competitive fitness experiment is started by mixing relatively equal proportion of green cells and red cells in deep 96-well blocks (100ul of a single ymCherry expressing strain and 100ul of a single ymEGFP expressing strain into 600ul distilled water) at a 4-fold dilution. To obtain further dilutions, 20ul of the competition is diluted into 300ul distilled water to form a 16 fold dilution, then 20ul of this dilution is diluted into 300ul of defined media supplemented with amino acids and 100ug/mL ampicillin to form a final 16-fold dilution. The cells are therefore diluted 1024-fold and this operation is performed every 24 hours. Given a conservative estimate of $2^{*}10^{8}$ yeast cells per mL at saturation, we estimate an effective population size (Ne) of approximately $3.44 * 10^{5}$. Each screen competes 8 genotypes against the others (64 wells), with an additional 16 wells used as contamination control. We use the diagonal of the competitions as an additional negative control as these were competing genetically identical strains constructed independently with different fluorophores. Competitive fitness assays were performed in synthetic media with 10ug/mL benomyl, which is a condition at which wild-type cells still

undergo ten divisions per day but seriously impairs growth of cells lacking spindle checkpoint function.

We defined the relative selection coefficient (*s*) on the basis of the deterministic continuous time model of logistic growth of an allele against another: dR/dt = sRG, where R and G are the frequencies of red and green cells in the population, t the number of generations, and s the selection coefficient [57]. To simplify the parameter inference, we ignore drift (because the timescale of the experiment is very short compared to Ne), mutational processes that happen during the fitness assay (i.e. we ignore mutations occurring in red or green cells that may alter the lineage trajectories), recombination, and any non-transitive effects. Constraining on G = 1-R, this model describes the logistic equation and has solution [57]:

$$R(t) = \frac{R(0)e^{st}}{(1 - R(0)) + R(0)e^{st}}$$

Because we do not observe the frequencies directly, we estimate the parameters s (and R(0) if needed) by counting each cell sampled by the flow-cytometer as a random variable, taking values of red or green (as it is described for a binomial logistic regression). In total, 50 000 cells are analyzed at the 20th and the 40th generation for each competition experiment and well-defined proportions of green and red cells are gated to remove doublets and relative cell counts are obtained from each competition [58,59]. The likelihood function for this counting process is simply:

$$\mathrm{L} = \prod_{t=1}^{T} \frac{n_t!}{r_t!(n_t - r_t)!} R(t)^{r_t} (1 - R(t))^{n_t - r_t}$$

Where t are the time points, r the number of red cells counted at that time point, n the total number of cells counted at that time point and R(t) the expected frequency of red cells at time point t. We estimate the initial frequency R(0) and s (due to their relationship with R(t)) by maximizing the likelihood, and the maximum likelihood estimate for the selection coefficient can be obtained by performing a log-linear regression on the ratio of red and green cells in the population:

$$\log\left(\frac{R(t)}{G(t)}\right) = \log\left(\frac{R(0)}{G(0)}\right) + st$$

For more than two time points, the parameters can be estimated by an iterative reweighted least-squares linear regression or with Newton's method. For two time points, the solution is equivalent to the simple linear regression and we use the same formula as in [59]:

$$\frac{\log\left(\frac{r(t2)}{g(t2)}\right) - \log\left(\frac{r(t1)}{g(t1)}\right)}{t2 - t1} = s$$

Clearly, if *s* is positive, then the ratio of red cells to green cells increases at the next generation and the value of the selection coefficient is therefore the selective effect of a beneficial allele as in [60].

The upper limit of detection occurs when we expect fewer than 1 out of 50000 cells of the worst genotype at the 40th generation, which occurs at approximately s = -0.3. In practice, some genotypes can no longer be detected at the 20th generation, and we simply report these as $s <$ -0.3. In some other cases, the number of red or green cells is fewer than 50, which leads to highly inaccurate estimates of the selection coefficient and we also report these as $s <$ -0.3.

When the total number of cells counted for both strains in a competition was fewer than 50, we reported $s = 0$ to mean equally lethal.

All genotypes were made in strains expressing red or green fluorescent proteins, and therefore assayed in two biological replicates. Key genotypes with undetectable fitness effects were further assayed with at least two technical replicates (where the same strains were assayed more than once on different days).

Localization analysis

Cells were grown in low-fluorescence media with appropriate auxotrophic requirements to log-phase and imaged using a Leica SP8 confocal microscope. Proteins of interest were tagged with EGFP [61], while the cytoplasm of cells were marked with cytoplasmic mCherry or mTagBFP2 [62] under a constitutive ribosomal promoter. We imaged the green fluorescence first, followed by the cytoplasmic marker to prevent bleaching. This setup enables highly controlled and quantitative analysis of localization because strains with differently tagged proteins can be imaged on the same field of view under identical conditions.

Double-blind quantification of the localization pattern was performed by manually inspecting the green fluorescence localization pattern first and scoring for whole nucleus, kinetochore (punctae), or a mixture of both. The scored cells were then assigned their proper genotype by looking at the red or blue fluorescence channels.

Supporting information

S1 Fig. Amino acid resolving power of changes in constraints are biologically relevant. A) Alignment of an example change in constraint in the Bub1 lineage. Numbers indicate residue position within the *S. cerevisiae* Bub1 protein. Alignment within the blue rectangle represents in the first N-terminal KEN box, while in green represents a subset of the TPR domain. B) Structural alignment of the region shown in A). The TPR domain of Bub1 (green) and Mad3 (cyan) are shown, along with the N-terminal tail of Mad3. Regions highlighted in red, orange and yellow correspond to regions boxed in A). C) Schematic representation of the binding interaction of Mad3 with Cdc20 can help elucidate why the regions shown in A) have correlated evolution. D) Alignment of a change in constraint in the TPR of Bub1 in independent duplication events.
(TIF)

S2 Fig. Effect of benomyl on the growth of other yeast species. Growth of *S. cerevisiae* cells of different genotypes is impaired at high benomyl concentration, while other yeast species are relatively unaffected as assessed by a 10-fold spot dilution assay. YPD plate was imaged after 2 days. YPD plate containing 30ug/mL benomyl was imaged after 3 days.
(TIF)

S3 Fig. Schematic of the Synthetic Genetic Array (SGA) methodology. Query strains containing SGA mating type reporters and fluorescent proteins are sequentially crossed to an ordered array of yeast strains containing marked alleles at different loci. The initial cross introduces alleles belonging to one locus, and the product of the haploid selection process after sporulation produces new query strains that differ by the allele introduced at one locus. The alleles at the second locus are then introduced in a second cross. Pink yeast outlines represent the MATα mating type, while blue yeast outlines represent the MATa mating type.
(TIF)

S4 Fig. Reversion to single-copy protein in the *Vanderwaltozyma* clade. A) Schematic of the critical functional regions of the proteins homologous to the SCP in the *Vanderwaltozyma*

clade, which is also a post-whole-genome duplication species, suggesting a reversion to SCP after gene duplication. B) Alignment of the critical functional regions of the single *Vanderwaltozyma polysporus* homologous protein to the SCP (L. klu), to Bub1 and Mad3 from *S. cerevisiae*.
(TIF)

S5 Fig. Effect of the MAD3 promoter in the expression of the single-copy protein. Spot dilution assay showing that the SCP does not rescue the spindle checkpoint defect of cells lacking Mad3 when placed at the MAD3 locus (under the MAD3 promoter). YPD plate containing 25ug/mL benomyl was imaged after 3 days.
(TIF)

Acknowledgments

We are grateful for the *L. kluyveri* strain that were generously donated by Dr. Marc-André Lachance. We thank Brenda Andrews for access to a microscope, as well as lab members of the Andrews lab for useful help during the project. We especially thank the Peisajovich lab for helping with the flow cytometer. We also thank Moses lab members, Dr. Nicholas Provart and Dr. Philip Kim for useful feedback during this study. We thank Damien D'Amours for suggesting the chromosome loss assay experiment and for the strain reporter. We also thank Dr. Andrew Murray for suggesting the competitive fitness assay and for other comments during the study.

Author Contributions

Conceptualization: ANNB AM.

Data curation: ANNB.

Formal analysis: ANNB.

Funding acquisition: ANNB AMM.

Investigation: ANNB BS SO EAL.

Methodology: ANNB.

Project administration: ANNB AMM BS TZ.

Resources: ANNB AMM.

Supervision: ANNB AMM TZ BS.

Validation: ANNB BS.

Visualization: ANNB BS AMM.

Writing – original draft: ANNB.

Writing – review & editing: ANNB AMM TZ EAL.

References

1. Conant GC, Wolfe KH. Turning a hobby into a job: how duplicated genes find new functions. Nat Rev Genet. 2008; 9: 938–950. https://doi.org/10.1038/nrg2482 PMID: 19015656
2. Gompel N, Prud'homme B. The causes of repeated genetic evolution. Dev Biol. 2009; 332: 36–47. https://doi.org/10.1016/j.ydbio.2009.04.040 PMID: 19433086

3. Suijkerbuijk SJE, van Dam TJP, Karagöz GE, von Castelmur E, Hubner NC, Duarte AMS, et al. The vertebrate mitotic checkpoint protein BUBR1 is an unusual pseudokinase. Dev Cell. 2012; 22: 1321–1329. https://doi.org/10.1016/j.devcel.2012.03.009 PMID: 22698286
4. Murray AW. Don't make me mad, Bub! Dev Cell. 2012; 22: 1123–1125. https://doi.org/10.1016/j.devcel.2012.05.020 PMID: 22698277
5. Lynch M. The frailty of adaptive hypotheses for the origins of organismal complexity. Proc Natl Acad Sci U S A. 2007; 104 Suppl 1: 8597–8604.
6. Lynch M. The evolution of genetic networks by non-adaptive processes. Nat Rev Genet. 2007; 8: 803–813. https://doi.org/10.1038/nrg2192 PMID: 17878896
7. Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J. Preservation of duplicate genes by complementary, degenerative mutations. Genetics. 1999; 151: 1531–1545. PMID: 10101175
8. Hittinger CT, Carroll SB. Gene duplication and the adaptive evolution of a classic genetic switch. Nature. 2007; 449: 677–681. https://doi.org/10.1038/nature06151 PMID: 17928853
9. Nguyen Ba AN, Strome B, Hua JJ, Desmond J, Gagnon-Arsenault I, Weiss EL, et al. Detecting functional divergence after gene duplication through evolutionary changes in posttranslational regulatory sequences. PLoS Comput Biol. 2014; 10: e1003977. https://doi.org/10.1371/journal.pcbi.1003977 PMID: 25474245
10. King EMJ, van der Sar SJA, Hardwick KG. Mad3 KEN boxes mediate both Cdc20 and Mad3 turnover, and are critical for the spindle checkpoint. PloS One. 2007; 2: e342. https://doi.org/10.1371/journal.pone.0000342 PMID: 17406666
11. Chao WCH, Kulkarni K, Zhang Z, Kong EH, Barford D. Structure of the mitotic checkpoint complex. Nature. 2012; 484: 208–213. https://doi.org/10.1038/nature10896 PMID: 22437499
12. Krenn V, Wehenkel A, Li X, Santaguida S, Musacchio A. Structural analysis reveals features of the spindle checkpoint kinase Bub1-kinetochore subunit Knl1 interaction. J Cell Biol. 2012; 196: 451–467. https://doi.org/10.1083/jcb.201110013 PMID: 22331848
13. Di Fiore B, Davey NE, Hagting A, Izawa D, Mansfeld J, Gibson TJ, et al. The ABBA motif binds APC/C activators and is shared by APC/C substrates and regulators. Dev Cell. 2015; 32: 358–372. https://doi.org/10.1016/j.devcel.2015.01.003 PMID: 25669885
14. London N, Ceto S, Ranish JA, Biggins S. Phosphoregulation of Spc105 by Mps1 and PP1 regulates Bub1 localization to kinetochores. Curr Biol CB. 2012; 22: 900–906. https://doi.org/10.1016/j.cub.2012.03.052 PMID: 22521787
15. Hardwick KG, Johnston RC, Smith DL, Murray AW. MAD3 encodes a novel component of the spindle checkpoint which interacts with Bub3p, Cdc20p, and Mad2p. J Cell Biol. 2000; 148: 871–882. PMID: 10704439
16. Connelly C, Hieter P. Budding yeast SKP1 encodes an evolutionarily conserved kinetochore protein required for cell cycle progression. Cell. 1996; 86: 275–285. PMID: 8706132
17. Warren CD, Brady DM, Johnston RC, Hanna JS, Hardwick KG, Spencer FA. Distinct chromosome segregation roles for spindle checkpoint proteins. Mol Biol Cell. 2002; 13: 3029–3041. https://doi.org/10.1091/mbc.E02-04-0203 PMID: 12221113
18. Fernius J, Hardwick KG. Bub1 kinase targets Sgo1 to ensure efficient chromosome biorientation in budding yeast mitosis. PLoS Genet. 2007; 3: e213. https://doi.org/10.1371/journal.pgen.0030213 PMID: 18081426
19. Breslow DK, Cameron DM, Collins SR, Schuldiner M, Stewart-Ornstein J, Newman HW, et al. A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. Nat Methods. 2008; 5: 711–718. https://doi.org/10.1038/nmeth.1234 PMID: 18622397
20. Tong AH, Evangelista M, Parsons AB, Xu H, Bader GD, Pagé N, et al. Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science. 2001; 294: 2364–2368. https://doi.org/10.1126/science.1065810 PMID: 11743205
21. Scannell DR, Frank AC, Conant GC, Byrne KP, Woolfit M, Wolfe KH. Independent sorting-out of thousands of duplicated gene pairs in two yeast species descended from a whole-genome duplication. Proc Natl Acad Sci U S A. 2007; 104: 8397–8402. https://doi.org/10.1073/pnas.0608218104 PMID: 17494770
22. Fernández A, Lynch M. Non-adaptive origins of interactome complexity. Nature. 2011; 474: 502–505. https://doi.org/10.1038/nature09992 PMID: 21593762
23. Lynch M, Conery JS. The origins of genome complexity. Science. 2003; 302: 1401–1404. https://doi.org/10.1126/science.1089370 PMID: 14631042
24. Straight AF, Murray AW. The spindle assembly checkpoint in budding yeast. Methods Enzymol. 1997; 283: 425–440. PMID: 9251039

25. Weissman DB, Desai MM, Fisher DS, Feldman MW. The rate at which asexual populations cross fitness valleys. Theor Popul Biol. 2009; 75: 286–300. https://doi.org/10.1016/j.tpb.2009.02.006 PMID: 19285994

26. Kellis M, Birren BW, Lander ES. Proof and evolutionary analysis of ancient genome duplication in the yeast Saccharomyces cerevisiae. Nature. 2004; 428: 617–624. https://doi.org/10.1038/nature02424 PMID: 15004568

27. Main BJ, Smith AD, Jang H, Nuzhdin SV. Transcription start site evolution in Drosophila. Mol Biol Evol. 2013; 30: 1966–1974. https://doi.org/10.1093/molbev/mst085 PMID: 23649539

28. Lang GI, Rice DP, Hickman MJ, Sodergren E, Weinstock GM, Botstein D, et al. Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. Nature. 2013; 500: 571–574. https://doi.org/10.1038/nature12344 PMID: 23873039

29. Finnigan GC, Hanson-Smith V, Stevens TH, Thornton JW. Evolution of increased complexity in a molecular machine. Nature. 2012; 481: 360–364. https://doi.org/10.1038/nature10724 PMID: 22230956

30. Baker CR, Hanson-Smith V, Johnson AD. Following gene duplication, paralog interference constrains transcriptional circuit evolution. Science. 2013; 342: 104–108. https://doi.org/10.1126/science.1240810 PMID: 24092741

31. Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN. Flexible nets. The roles of intrinsic disorder in protein interaction networks. FEBS J. 2005; 272: 5129–5148. https://doi.org/10.1111/j.1742-4658.2005.04948.x PMID: 16218947

32. Roop JI, Chang KC, Brem RB. Polygenic evolution of a sugar specialization trade-off in yeast. Nature. 2016; 530: 336–339. https://doi.org/10.1038/nature16938 PMID: 26863195

33. Gu Z, Rifkin SA, White KP, Li W-H. Duplicate genes increase gene expression diversity within and between species. Nat Genet. 2004; 36: 577–579. https://doi.org/10.1038/ng1355 PMID: 15122255

34. Huminiecki L, Wolfe KH. Divergence of spatial gene expression profiles following species-specific gene duplications in human and mouse. Genome Res. 2004; 14: 1870–1879. https://doi.org/10.1101/gr.2705204 PMID: 15466287

35. Gu X, Zhang Z, Huang W. Rapid evolution of expression and regulatory divergences after yeast gene duplication. Proc Natl Acad Sci U S A. 2005; 102: 707–712. https://doi.org/10.1073/pnas.0409186102 PMID: 15647348

36. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. 2005; 15: 1034–1050. https://doi.org/10.1101/gr.3715005 PMID: 16024819

37. Dernburg AF. Here, there, and everywhere: kinetochore function on holocentric chromosomes. J Cell Biol. 2001; 153: F33–38. PMID: 11402076

38. Melters DP, Paliulis LV, Korf IF, Chan SWL. Holocentric chromosomes: convergent evolution, meiotic adaptations, and genomic analysis. Chromosome Res Int J Mol Supramol Evol Asp Chromosome Biol. 2012; 20: 579–593.

39. Vleugel M, Hoogendoorn E, Snel B, Kops GJPL. Evolution and function of the mitotic checkpoint. Dev Cell. 2012; 23: 239–250. https://doi.org/10.1016/j.devcel.2012.06.013 PMID: 22898774

40. Hasty J, McMillen D, Isaacs F, Collins JJ. Computational studies of gene regulatory networks: in numero molecular biology. Nat Rev Genet. 2001; 2: 268–279. https://doi.org/10.1038/35066056 PMID: 11283699

41. Wagner A. Neutralism and selectionism: a network-based reconciliation. Nat Rev Genet. 2008; 9: 965–974. https://doi.org/10.1038/nrg2473 PMID: 18957969

42. Hao N, Behar M, Parnell SC, Torres MP, Borchers CH, Elston TC, et al. A systems-biology analysis of feedback inhibition in the Sho1 osmotic-stress-response pathway. Curr Biol CB. 2007; 17: 659–667. https://doi.org/10.1016/j.cub.2007.02.044 PMID: 17363249

43. Josephides C, Moses AM. Modeling the evolution of a classic genetic switch. BMC Syst Biol. 2011; 5: 24. https://doi.org/10.1186/1752-0509-5-24 PMID: 21294912

44. Li Z, Vizeacoumar FJ, Bahr S, Li J, Warringer J, Vizeacoumar FS, et al. Systematic exploration of essential yeast gene function with temperature-sensitive mutants. Nat Biotechnol. 2011; 29: 361–367. https://doi.org/10.1038/nbt.1832 PMID: 21441928

45. Wach A, Brachat A, Pöhlmann R, Philippsen P. New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. Yeast Chichester Engl. 1994; 10: 1793–1808.

46. Storici F, Lewis LK, Resnick MA. In vivo site-directed mutagenesis using oligonucleotides. Nat Biotechnol. 2001; 19: 773–776. https://doi.org/10.1038/90837 PMID: 11479573

47. Boeke JD, Trueheart J, Natsoulis G, Fink GR. 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. Methods Enzymol. 1987; 154: 164–175. PMID: 3323810

48. Tong AHY, Boone C. Synthetic Genetic Array Analysis in Saccharomyces cerevisiae. In: Xiao W, editor. Yeast Protocol. Humana Press; 2006. pp. 171–191. http://link.springer.com.myaccess.library.utoronto.ca/protocol/10.1385/1-59259-958-3%3A171

49. Chattoo BB, Sherman F, Azubalis DA, Fjellstedt TA, Mehnert D, Ogur M. Selection of lys2 Mutants of the Yeast SACCHAROMYCES CEREVISIAE by the Utilization of alpha-AMINOADIPATE. Genetics. 1979; 93: 51–65. PMID: 17248969

50. Byrne KP, Wolfe KH. The Yeast Gene Order Browser: combining curated homology and syntenic context reveals gene fate in polyploid species. Genome Res. 2005; 15: 1456–1461. https://doi.org/10.1101/gr.3672305 PMID: 16169922

51. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002; 30: 3059–3066. PMID: 12136088

52. SGD Project. Saccharomyces Genome Database [Internet]. 18 Jan 2011 [cited 18 Jan 2011]. http://www.yeastgenome.org/

53. Nguyen Ba AN, Yeh BJ, van Dyk D, Davidson AR, Andrews BJ, Weiss EL, et al. Proteome-wide discovery of evolutionary conserved sequences in disordered regions. Sci Signal. 2012; 5: rs1. https://doi.org/10.1126/scisignal.2002515 PMID: 22416277

54. Bolanos-Garcia VM, Kiyomitsu T, D'Arcy S, Chirgadze DY, Grossmann JG, Matak-Vinkovic D, et al. The crystal structure of the N-terminal region of BUB1 provides insight into the mechanism of BUB1 recruitment to kinetochores. Struct Lond Engl 1993. 2009; 17: 105–116.

55. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucleic Acids Res. 2000; 28: 235–242. PMID: 10592235

56. Schrödinger L. The PyMOL Molecular Graphics System, Version 1.3r1. 2010.

57. Gallet R, Cooper TF, Elena SF, Lenormand T. Measuring selection coefficients below 10(-3): method, questions, and prospects. Genetics. 2012; 190: 175–186. https://doi.org/10.1534/genetics.111.133454 PMID: 22042578

58. Lang GI, Murray AW, Botstein D. The cost of gene expression underlies a fitness trade-off in yeast. Proc Natl Acad Sci U S A. 2009; 106: 5755–5760. https://doi.org/10.1073/pnas.0901620106 PMID: 19299502

59. Frenkel EM, Good BH, Desai MM. The fates of mutant lineages and the distribution of fitness effects of beneficial mutations in laboratory budding yeast populations. Genetics. 2014; 196: 1217–1226. https://doi.org/10.1534/genetics.113.160069 PMID: 24514901

60. Kimura M. On the probability of fixation of mutant genes in a population. Genetics. 1962; 47: 713–719. PMID: 14456043

61. Sheff MA, Thorn KS. Optimized cassettes for fluorescent protein tagging in Saccharomyces cerevisiae. Yeast Chichester Engl. 2004; 21: 661–670.

62. Lee S, Lim WA, Thorn KS. Improved blue, green, and red fluorescent protein tagging vectors for S. cerevisiae. PloS One. 2013; 8: e67902. https://doi.org/10.1371/journal.pone.0067902 PMID: 23844123