PomBase – the scientific resource for fission yeast

4 authors, including:

Valerie Wood
University of Cambridge
78 PUBLICATIONS   **8,427** CITATIONS

Antonia Lock
University College London
16 PUBLICATIONS   **1,044** CITATIONS

Midori Harris
University of Cambridge
46 PUBLICATIONS   **30,712** CITATIONS

Some of the authors of this publication are also working on these related projects:

Gene Ontology View project

Gene Ontology View project

Chapter 4

PomBase: The Scientific Resource for Fission Yeast

Antonia Lock, Kim Rutherford, Midori A. Harris, and Valerie Wood

Abstract

The fission yeast *Schizosaccharomyces pombe* has become well established as a model species for studying conserved cell-level biological processes, especially the mechanics and regulation of cell division. PomBase integrates the *S. pombe* genome sequence with traditional genetic, molecular, and cell biological experimental data as well as the growing body of large datasets generated by emerging high-throughput methods. This chapter provides insight into the curation philosophy and data organization at PomBase, and provides a guide to using PomBase for infrequent visitors and anyone considering exploring *S. pombe* in their research.

Key words *Schizosaccharomyces pombe*, Fission yeast, Biological database, Model organism, Gene ontology, GO slim, Annotation extensions

1 Introduction

PomBase (http://www.pombase.org/), funded by the Wellcome Trust, is the model organism database (MOD) for the fission yeast *Schizosaccharomyces pombe*. Its primary goals are:

- To support the exploratory and hypothesis-driven research needs of those using the model eukaryote fission yeast as an experimental system.

- To provide an integrated model of a eukaryotic cell.

- To promote and support the use of fission yeast as a model eukaryotic system, with particular relevance to human biology.

- To provide a community hub, and support for in-depth community led curation.

To accomplish these goals, PomBase integrates the *S. pombe* genome sequence and features with genome-wide datasets and detailed, comprehensive gene-oriented manual curation of published literature, and provides tools to interrogate these data [1, 2].

As an experimental organism, fission yeast is inexpensive to grow, proliferates rapidly and is amenable to genetic manipulation. Researchers typically study isogenic strains of *S. pombe* derived from a single isolate known as 968 h90. This facilitates the comparison of results across different laboratories. Aside from the PomBase database, several organism specific resources are available to fission yeast researchers, including the genome-wide deletion mutant collection [3, 4] and the Orfeome localization collection [5].

The sequence of the reference strain 972 h–(a naturally occurring 968 h90 derivative) was published in 2002 [6]. Fission yeast has a compact genome, 14 Mb in size, that consists of three chromosomes of 3.5–5.6 Mb plus a 19 kb mitochondrial genome. 5071 protein-coding genes, of which more than two thirds are conserved in human, are annotated in the reference genome, along with rRNAs, tRNAs, snRNAs, snoRNAs, and other noncoding RNAs.

The fission yeast community comprises ~2000 researchers working primarily or exclusively with *S. pombe* or other *Schizosaccharomyces* species. In addition, fission yeast is used extensively as a complementary organism by those studying conserved cellular mechanisms in vertebrate systems, including the cell cycle, cytokinesis, chromosome segregation, epigenetic mechanisms, DNA metabolism, and drug responses [7–10]. PomBase thus serves a large (~15,000 unique users per month) and varied user base with diverse experience and requirements.

2   Data Curation in PomBase

The most precise and reliable molecular data in PomBase are generated by manual curation of the fission yeast literature. Automated methods, such as annotation transfer based on sequence orthology, and high-throughput datasets supplement the body of manually curated data.

To enable the fission yeast community to contribute directly to PomBase, we have developed Canto [11], an intuitive web-based literature curation tool. Canto allows both professional curators and community researchers to use state-of-the-art annotation techniques to build complex connections among genes, ontology terms, and supporting metadata. Notably, the use of ontology terms and "annotation extensions" described below underlies the generation of comprehensive curated networks representing biological processes. By combining the topic-specific expertise of biological experts with PomBase curators' familiarity with ontologies and annotation practices, Canto usage yields literature curation of a particularly high standard of accuracy and specificity [12]. To date (August 2017) approximately 10,000 annotations have been submitted by community curators for over 500 publications.

3 PomBase Gene Page Organization

Like other model organism databases, PomBase organizes data into pages summarizing genes, publications, ontology terms, and others, of which the most intensively used are gene pages. Each gene page gathers all data relevant to the gene into one place, with a menu that shows available data types at a glance and facilitates navigation within the page (Fig. 1). Gene pages can be accessed directly by typing a gene name in the search field at the top right corner of each PomBase page, and selecting it from the drop-down list (e.g., *clp1*, *cdc2*, *cdc25*, *mde4*).

Curated data types include ontology-based annotations for gene function (Gene Ontology; GO), phenotypes, and modifications, genetic and physical interactions, qualitative and quantitative gene expression data, protein features, complementation,

Fig. 1 The quick-links menu. The menu displayed on the left-hand side of gene pages provides an overview of the different data types available for specific genes, and enables rapid navigation between the different sections of the gene page

orthologs and taxonomic conservation. Gene pages also provide gene and protein sequences, and links to gene-specific entries in external databases, and a collection of literature relevant to each gene. We discuss several of these in depth below.

3.1 Gene Ontology Data

The Gene Ontology (GO) section of the gene pages shows a table of annotations using each of the main branches of GO: molecular function, biological process, and cellular component [13, 14]. By default, the tables display a nonredundant summary of annotated terms and extensions. Figure 2A shows a selection of molecular function and biological process annotations for the protein phosphatase *clp1*. An expanded view shows all annotations as well as supporting metadata such as references, evidence, term IDs, and qualifiers. Ontology terms, genes, and references in the annotation views link to additional PomBase pages. The biological process section also lists any GO slim terms (*see* below) applicable to the gene.

The *clp1* GO molecular function annotation shown in Fig. 2A also illustrates the usage of GO annotation extensions. PomBase was a pioneer in the implementation of annotation extensions [15], which allow curation of effector–target relationships (such as protein kinase substrates) or spatiotemporal information (such as where and when a function is executed). Extensions on the *clp1* "serine/threonine phosphatase activity" molecular function annotation indicate that Clp1 dephosphorylates different substrates to contribute to different regulatory processes (e.g., Clp1 dephosphorylates Mde4 to positively regulate spindle elongation during anaphase).

Figure 2B shows a summary of the relationships used at PomBase to curate annotation extensions, and then, as described below, to build networks using the resulting connections among gene products.

3.2 Phenotype Data

Phenotypes are curated by PomBase using the Fission Yeast Phenotype Ontology (FYPO), an ontology of over 6000 precomposed phenotype terms [16]. Fission yeast researchers typically study isogenic strains, making it possible to define "normal" and "abnormal" phenotypes in mutants compared to the behavior of the "wild type" reference strain.

PomBase curates single mutant allele and multiallele genotypes, which are displayed in separate gene page sections. The phenotype view is further split into population and cell level phenotypes and users can toggle between a summary view (Fig. 3A) and a detailed view (Fig. 3B). Gene deletion viability is indicated at the top of the single mutant phenotype section. The displayed phenotypes can be filtered by broad phenotypic categories (viability, cell cycle, morphology, etc.), improving the usability of the very long phenotype lists now present for many genes (green box, Fig. 3A).

A

GO molecular function

Show details ...

[+] protein serine/threonine phosphatase activity

    has substrate cdc11

    has substrate clp1 **involved in** positive regulation of exit from mitosis **during** mitotic telophase

    has substrate mde4 **involved in** positive regulation of mitotic spindle elongation **during** mitotic anaphase

    has substrate nsk1 **involved in** positive regulation of attachment of mitotic spindle microtubules to kinetochore

GO biological process

Slim Terms

| microtubule cytoskeleton organization | mitotic cytokinesis | mitotic sister chromatid segregation |

| regulation of mitotic cell cycle phase transition | signaling |

Show details ...

[+] positive regulation of attachment of mitotic spindle microtubules to kinetochore

[+] positive regulation of exit from mitosis

[+] positive regulation of mitotic spindle elongation

B

bir1 nuclear chromatin *at* centromere outer repeat region *during* metaphase

Cellular Component — *During / At* → Biological Process — *During Regulates*

At / Affects / Regulates

At

sid2 negative regulation of protein localization to microtubule *regulates* klp2 *during* mitotic telophase

Sequence feature Gene product

Involved in During

Has substrate

Molecular Function

cut1 endopeptidase activity *has substrate* rec8 *involved in* meiotic chromosome separation
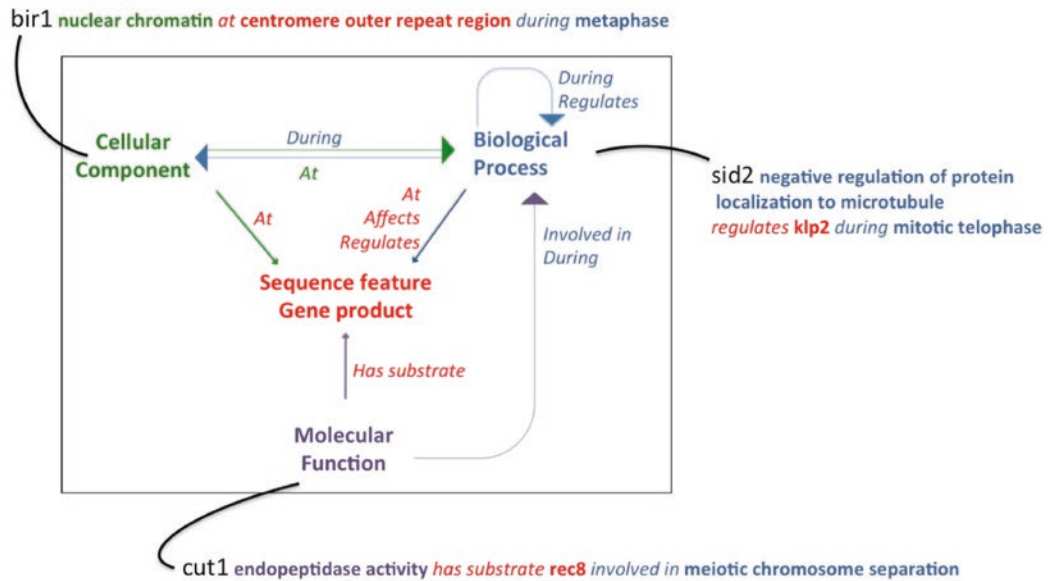
Fig. 2 GO annotations and extensions. (A) Summary view of selected annotations on the *clp1* gene page. The orange boxes highlight annotations representing Clp1's roles: Clp1 dephosphorylates the Nsk1 protein to positively regulate spindle attachment to the kinetochore. During anaphase, it dephosphorylates Mde4 to positively regulate spindle elongation. Clp1 also directly targets itself during telophase to promote mitotic exit. Processes linked to molecular functions are also shown in the biological process section. Biological process annotations that map to the PomBase GO slim are shown at the top of the biological process section Fig. 2 (continued) (purple box). (B) Relations used in GO annotation extensions, showing how each links one gene to other genes or additional ontology terms, with examples for each GO branch
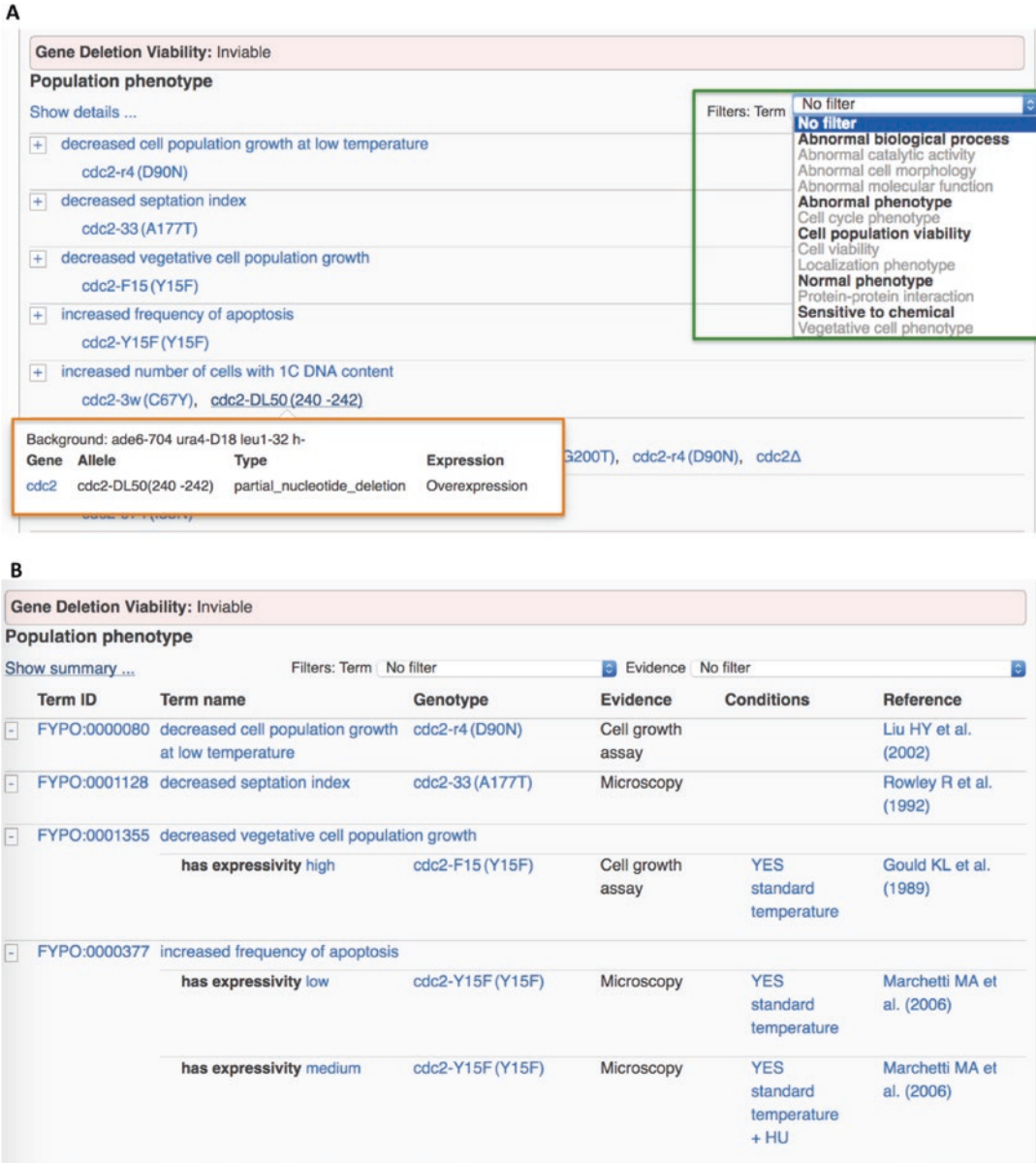
A

B

Fig. 3 The PomBase phenotype display. (A) Summary view (B) Detailed view of a subset of phenotypes associated with alleles of *cdc2*. In the summary view, redundant annotations (including annotation to the same phenotype term, but with different extensions or metadata) and metadata are hidden. The detailed view shows all annotations, plus the cited references, evidence, extensions indicating penetrance, expressivity, or affected gene products, and any curated experimental, conditions. Genotype details, including the type of mutation for each allele, expression level of the gene products, and any background genotype information, are provided in a mouse-over (shown in A, orange box). A drop-down menu enables filtering for subsets of phenotypes (shown in A, green box)

cdc25-22(C532Y) wee1-50(unknown)

Background pap1::kanr bfr1::hygr pmd1::natr caf5::kanr mfs1::natr

Gene	Allele	Type	Expression
cdc25	cdc25-22(C532Y)	amino_acid_mutation	Knockdown
wee1	wee1-50(unknown)	unknown	Knockdown

Annotations for this genotype

Population phenotype

Show details ...   Filters: Term   No filter

[+]   resistance to Cutin-1

Cell phenotype

Show details ...   Filters: Term   No filter

[+]   abnormal cell size

[+]   abnormal chromosome segregation

[+]   abnormal nucleus

Fig. 4 Genotype pages. Each genotype page displays allele and expression details and all annotations associated with the genotype. In this example, the double mutant comprising *cdc25-22* (C532Y) and *wee1-50,* both at reduced expression levels, in the background *pap1::kanr bfr1::hygr pmd1::natr caf5::kanr mfs1::natr* has been associated with four different phenotype terms

Each phenotype annotation also links to a page dedicated to the genotype, which displays details (name, description, expression level) for the alleles that make up the genotype, any background alleles, and all annotated phenotypes (Fig. 4).

3.3   Targets

The "Target of" section provides information about gene products or mutations that affect the gene of interest, derived from the reciprocal of annotations specifying targeted genes, such as the substrates of molecular functions. Target annotations indicate the connecting ontology term and the specific relationship between the two genes. For example, Clp1 dephosphorylates (Fig. 2A) and Cdc2 phosphorylates (Fig. 6) the Mde4 protein. Since Mde4 is targeted by these proteins, *clp1* and *cdc2* are listed in the *mde4* "Target of" section (Fig. 5). Users can thus navigate entire biological pathways; downstream by a gene product's GO molecular function substrates, and upstream by effectors in the "target of" section. Reciprocal annotations are also generated from phenotype and protein modification annotations.

3.4   Taxonomic Conservation, Orthologs, and Disease Curation

To support the growing cohort of researchers using both fission yeast and other species, PomBase maintains manually curated inventories of orthologous proteins for fission yeast vs. human and fission yeast vs. budding yeast (*Saccharomyces cerevisiae*). Both are

Fig. 5 The *mde4* "Target of" section. Because *cdc2* is annotated to a protein kinase molecular function term, with Mde4 specified as a substrate, *cdc2* is listed in the "target of" section for *mde4*. Reciprocal annotations are also generated from phenotype and protein modification annotations. For example, a mutation in *cdc2* has an effect on *mde4*, with phenotypic details available on the *cdc2* gene page, and a "target of" annotation using the "affected by mutation in" relationship on the *mde4* gene page

compiled by integrating published data and in-house analyses with the consensus from numerous orthology resources [17]. The human–fission yeast curated orthology dataset now identifies human orthologs for 69% of the fission yeast proteome.

Gene pages show any manually curated orthologous genes in human and budding yeast, and the basic gene search will retrieve available *S. pombe* orthologs using human standard gene names or budding yeast systematic (ORF) names. Where a fission yeast gene has a human ortholog that has been implicated in a disease, the PomBase gene page notes the disease and a source publication.

The taxonomic conservation section shows a broad domain kingdom or phylum level conservation for protein-coding genes. Taxon restrictions are also recorded where applicable. Other terms may also be assigned, such as whether the gene is conserved one-to-one. Classifiers are assigned manually from a small in-house controlled vocabulary (Table 1).

Taxonomic conservation can be used to retrieve high quality broad taxon classification specific datasets for analyses, or to provide functional clues for unstudied proteins based on presence or absence in particular kingdom or phyla.

4   Building Networks

The growing body of GO annotations with annotation extensions in PomBase creates connections between gene products, and provides rich biological context to their interactions. These connections can be exploited to reconstruct biological pathways. For
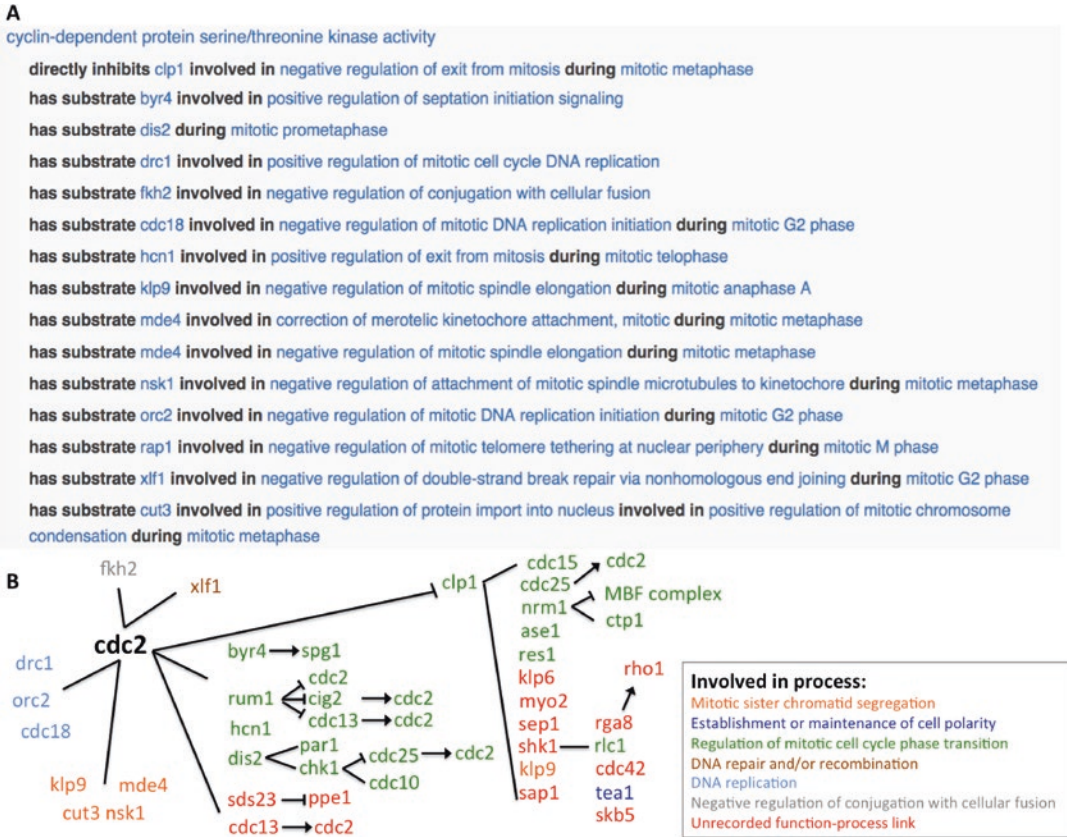
Fig. 6 *cdc2* function–process links and downstream signaling cascades. (A) Showing the subset of Cdc2 phosphorylation targets with function–process links. Biological processes that the enzyme–substrate interaction is part of, or happens during, are indicated using the "*involved in*" and "*during*" annotation extension relationships. (B) Targets downstream of Cdc2 can be accessed via the hyperlinked annotation extension substrates, enabling users to follow biological pathways. The capturing of targets makes it possible to reconstruct pathways for a systems level representation of gene networks

example, the highly conserved cyclin-dependent serine/threonine kinase Cdc2 (homolog of the mammalian *CDK1*) is known to directly phosphorylate over 140 different proteins. A number of these *cdc2*–substrate connections are linked to the biological processes that the interaction is regulating (Fig. 6A). Annotated substrates can be followed, in order to move down biological pathways (Fig. 6B).

PomBase will use the connections curated between gene products (enzyme–substrate links, and high confidence physical interaction data), and the links to the biological processes they are involved in, to automatically construct networks for biological processes. This approach will ultimately create a detailed and reliable curation-based network for a eukaryotic cell.

Table 1
Taxonomic conservation groups. Taxonomic conservation groups are assigned manually from a controlled set of terms at the kingdom/domain level. A gene may be annotated to multiple different orthologous groups. Taxon restrictions are recorded for where orthologs have not been identified outside of the noted taxa (fungi or eukaryotes). The absence of an ortholog in the *S. cerevisiae* reference sequence is also recorded. Copy number conservation is also documented, for example whether the gene is conserved one-to-one or whether orthologs cannot be distinguished. In some cases, faster evolving duplicates may be observed—this is where a copy of a gene appears to evolve faster than another copy of the gene

Orthologous groups	Conserved in archaea
	Conserved in bacteria
	Conserved in eukaryotes
	Conserved in fungi
	Conserved in metazoa
	Conserved in vertebrates
	Schizosaccharomyces specific
	Schizosaccharomyces pombe specific
Taxon restrictions	Conserved in fungi only
	Conserved in eukaryotes only
Others	No apparent *S.cerevisiae* ortholog
	Predominantly single copy (one-to-one)
	Orthologs cannot be distinguished
	Faster evolving duplicate

5   GO Slim Summary

PomBase maintains the fission yeast GO slim, a set of broad, high level GO biological process terms providing coverage for most gene products with process annotations (http://www.pombase. org/browse-curation/fission-yeast-go-slim-terms). Like other GO slim sets (*see* http://geneontology.org/page/go-slim-and-subset-guide), the fission yeast GO slim supports genome-level overview of GO annotation coverage, and can be used to summarize large-scale experimental results.

The PomBase GO slim terms encompass 99.5% of all genes with a biological process annotation. Uninformative (very high level grouping terms) are excluded from the PomBase GO-slim set. Table 2 shows the number of gene products annotated to each fission yeast GO slim term. Of the 5071 *S. pombe* proteins, 748 do not have a biological process annotation because their biological

Table 2
Fission yeast GO slim annotations. For each term in the fission yeast GO slim, the number of annotated genes is shown. Note that a gene may be annotated to more than one slim term

GO slim term	Number of genes
Actin cytoskeleton organization	89
Apoptotic process	8
Ascospore formation	74
Autophagy	49
Carbohydrate derivative metabolic process	276
Carbohydrate metabolic process	138
Cell adhesion	20
Cell wall organization or biogenesis	104
Cellular amino acid metabolic process	190
Chromatin organization	278
Cofactor metabolic process	177
Conjugation with cellular fusion	100
Cytoplasmic translation	485
Detoxification	59
DNA recombination	122
DNA repair	177
DNA replication	118
Establishment or maintenance of cell polarity	74
Generation of precursor metabolites and energy	81
Lipid metabolic process	232
Meiotic nuclear division	112
Membrane organization	174
Microtubule cytoskeleton organization	75
Mitochondrial gene expression	167
Mitochondrion organization	146
Mitotic cytokinesis	100
Mitotic sister chromatid segregation	176
mRNA metabolic process	271
Nitrogen cycle metabolic process	16

Table 2
(continued)

GO slim term	Number of genes
Nucleobase-containing small molecule metabolic process	191
Nucleocytoplasmic transport	108
Peroxisome organization	22
Polyphosphate metabolic process	2
Protein catabolic process	212
Protein complex assembly	126
Protein folding	84
Protein glycosylation	68
Protein maturation	60
Protein modification by small protein conjugation or removal	98
Protein targeting	103
Regulation of mitotic cell cycle phase transition	165
Regulation of transcription, DNA-templated	415
Ribosome biogenesis	348
Signaling	292
snoRNA metabolic process	33
snRNA metabolic process	19
Sulfur compound metabolic process	109
Telomere organization	45
Transcription, DNA-templated	470
Transmembrane transport	355
tRNA metabolic process	170
Vesicle-mediated transport	329
Vitamin metabolic process	42
Proteins with a biological process annotation not covered by the slim	*27*
Proteins with no slim or biological process annotation	*748*

role is currently unknown in any species (i.e., neither the *S. pombe* protein nor any ortholog has been experimentally characterized in detail).

PomBase also maintains a list of "priority unstudied genes" for genes conserved across taxa, but not yet characterized in any species (http://www.pombase.org/status/priority-unstudied-genes).

6 Ontology Term Views

Each ontology term used in annotations or extensions (GO, Fission Yeast Phenotype Ontology (FYPO), the Sequence Ontology (SO) [18], the chemical ontology ChEBI [19], and the PSI-MOD protein modification ontology [20]) has a term page in PomBase. The top of the term page shows the name, ID, direct links to more general "parent" terms in the ontology, and external links to relevant resources (Fig. 7A). For ontologies used directly in annotations (GO, FYPO, PSI-MOD), genes are associated with the most specific annotated descendant term (Fig. 7B shows a subset of the genes annotated directly to GO:0023052 "signaling" or any of its descendant terms). As on gene pages, the default summary view can be expanded to display annotation extensions, the type of relationship between child and parent terms (e.g., *is_a*, *part_of* or *regulates*), and supporting metadata (Fig. 7C).

7 Publication Pages

Every paper cited in support of PomBase annotations has a publication page that displays citation details, the abstract, and all annotations curated from the publication (Fig. 8). Publication pages are connected from annotation tables and the literature section on gene pages, and from all pages that display the corresponding annotations. The page also acknowledges any community member who has contributed to the annotations derived from the publication.

8 Querying

PomBase offers simple and advanced search tools for querying genes and their annotations. The simple search, available on every page, retrieves individual genes by standard name, systematic ID or an *S. cerevisiae* or human ortholog name.

The advanced search retrieves sets of genes that match criteria specified by an assortment of filters (Fig. 9A). For example, ontology terms can be queried by name or ID to find annotated genes. All genes can be queried by criteria such as name, ID, product description, or chromosomal location. Additional filters are available for features of protein-coding genes. Queries can be combined to narrow down results to genes matching several criteria (Fig. 9B). Queries can be combined using the Boolean operators AND (intersect), NOT (subtract), and OR (union), and saved for reuse (Fig. 9C). For genes matching a query, IDs, names, product

A

GO:0023052 - signaling

Definition

The entirety of a process in which information is transmitted within a biological system. This process begins with an active signal and ends when a cellular response has been triggered.

This term is part of the biological process overview (GO slim) - View the esyN network

External links: AmiGO I QuickGO I BioPortal

View genes annotated with this term ...

Parents

is_a biological_process

B

GO biological process annotations for GO:0023052 and its descendants

Show details ...

[+]   adenylate cyclase-activating glucose-activated G-protein coupled receptor signaling pathway

      cyr1, git1, git11, git3, git5, gpa2, pka1

[+]   negative regulation of adenylate cyclase-activating glucose-activated G-protein coupled receptor signaling pathway

      cgs1, sck1

      atf1 **involved in** positive regulation of mitotic G1 cell cycle arrest in response to nitrogen starvation

      cgs2 **involved in** positive regulation of mitotic G1 cell cycle arrest in response to nitrogen starvation

      pcr1 **involved in** positive regulation of mitotic G1 cell cycle arrest in response to nitrogen starvation

C

GO biological process annotations for GO:0023052 and its descendants

Show summary ...

Gene	Term ID	Term name	Evidence	Qualifiers	Reference	Count
[-] cyr1	↑ part_of GO:0010619	adenylate cyclase-activating glucose-activated G-protein coupled receptor signaling pathway	IMP		Higuchi T et al. (2002)	12
cyr1			IMP		Landry S et al. (2001)	
cyr1			IMP		Ivey FD et al. (2005)	
cyr1			IMP		Demirbas D et al. (2011)	
git1			NAS		GO_REF:0000051	
git11			IMP		Landry S et al. (2001)	
git3			IMP		Mudge DK et al. (2014)	
git3			IGI with gpa2		Welton RM et al. (2000)	

Fig. 7 Ontology term pages. (A) The top of the page shows the term name, ID, and definition, along with immediate parent terms. Links to external resources are provided. (B) The summary view shows genes annotated directly to the term or to any of its child terms, and includes extensions. (C) The detailed view provides additional information such as the relationship of child terms to the parent term, evidence codes and references

The Msd1-Wdr8-Pkl1 complex anchors microtubule minus ends to fission yeast spindle pole bodies.

Contact curators ...

Authors   Yukawa M, Ikebe C, Toda T

Citation   J. Cell Biol. 2015 May 25;209(4):549-62

ID   PMID:25987607

Links   Europe PMC I PubMed

[θ] Community curation provided by Takashi Toda

Abstract   The minus ends of spindle microtubules are anchored to a microtubule-organizing center. The conserved Msd1/SSX2IP proteins are localized to the spindle pole body (SPB) and the centrosome in fission yeast and humans, respectively, and play a critical role in microtubule anchoring. In this paper, we show that fission yeast Msd1 forms a ternary complex with another conserved protein, Wdr8, and the minus end-directed Pkl1/kinesin-14. Individual deletion mutants displayed the identical spindle-protrusion phenotypes. Msd1 and Wdr8 were delivered by Pkl1 to mitotic SPBs, where Pkl1 was tethered through Msd1-Wdr8. The spindle-anchoring defect imposed by msd1/wdr8/pkl1 deletions was suppressed by a mutation of the plus end-directed Cut7/kinesin-5, which was shown to be mutual. Intriguingly, Pkl1 motor activity was not required for its anchoring role once targeted to the SPB. Therefore, spindle anchoring through Msd1-Wdr8-Pkl1 is crucial for balancing the Cut7/kinesin-5-mediated outward force at the SPB. Our analysis provides mechanistic insight into the spatiotemporal regulation of two opposing kinesins to ensure mitotic spindle bipolarity.

Annotations from this publication:

GO molecular function

Show details ...

± ATP-dependent microtubule motor activity, minus-end-directed

    pkl1 **has substrate** msd1, wdr8 **involved in** protein transport along microtubule to spindle pole body

GO biological process

Show details ...

± microtubule anchoring at mitotic spindle pole body

    msd1, pkl1, wdr8

± mitotic sister chromatid segregation

    msd1, wdr8

± protein localization to mitotic spindle pole body

    msd1 **localizes** pkl1

GO cellular component

Show details ...

± mitotic spindle

    msd1, pkl1, wdr8

± mitotic spindle pole body

    msd1 **during** mitotic M phase

    pkl1 **during** mitotic M phase

    wdr8 **during** mitotic M phase

Single allele phenotype

Population phenotype

Show details ...   Filters: Term No filter

± sensitive to thiabendazole

    msd1Δ, pkl1Δ, wdr8Δ

Cell phenotype

Show details ...   Filters: Term No filter

± abnormal protein localization to microtubule minus-end

    msd1Δ **affecting** cut7

    wdr8Δ **affecting** cut7

Physical interaction

Gene	Product		Interacting gene	Interacting product	Evidence
pkl1	minus-end directed microtubule motor, kinesin Pkl1	affinity captures	msd1	microtubule-anchoring factor Msd1	Affinity Capture-Western

Fig. 8 Publication pages. The PMID:25987607 page shows publication details and a community curator acknowledgement at the top, and annotations derived from the paper. GO and FYPO annotations have summary and detailed views as on gene and ontology term pages
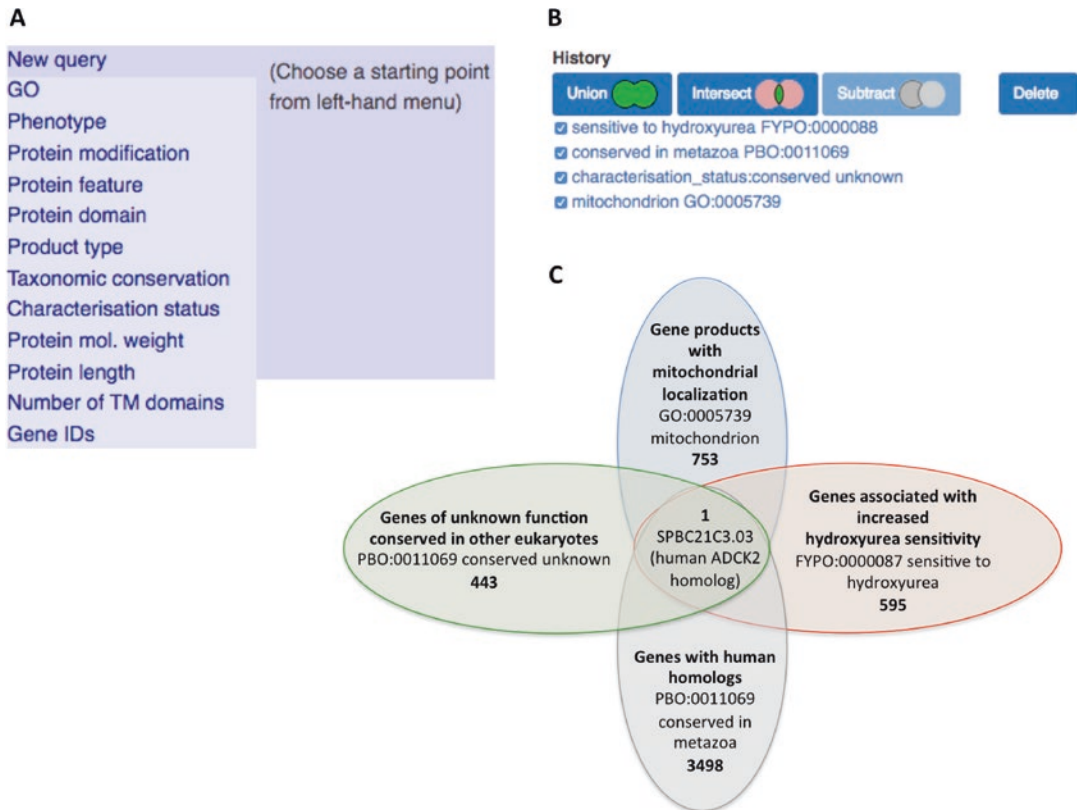
Fig. 9 Query builder filtering. (A) A list of the different filters available to identify genes of interest. (B) The history section can be used to review previously run queries. Queries can be combined using the union, intersect and subtract operators. (C) An example of the results of running and combining queries. 753 genes (August 2017) are annotated to the GO term mitochondrion. Of these, 3498 are conserved in metazoa, 595 genes where any type of allele has been associated with hydroxyuruea sensitivity and 411 genes with the characterization status "conserved unknown," i.e., of unknown function but conserved in other organisms. The union of these four queries identifies one gene

descriptions, and sequences can be downloaded. More flexible download options for query results are slated for addition to the advanced search.

An additional stand-alone motif search tool searches all protein coding sequences to identify genes containing a specified amino acid pattern of interest.

9   Genome Browser and Datasets

The PomBase genome browser supports sequence viewing based on coordinates or feature identifiers. Data tracks are available for sequence-based datasets submitted by the fission yeast community

from a variety of high-throughput experiments, including transcriptomic data [21–23], nucleosome positioning [23], replication profiling [24], polyadenylation sites [25, 26], and chromatin binding [27]. (Note: at the time of writing, PomBase is in the process of transitioning from a legacy browser to a JBrowse [28] implementation.)

PomBase also provides a set of static pages describing various aspects of genome-level curation status and links to external community resources. The genome sequence and several annotation datasets (GO, phenotype, and modification data, orthologs, interactions, protein features, etc.) can be downloaded from PomBase's FTP site.

10   Community and Outreach

PomBase makes community engagement a high priority, welcoming data submissions and feedback on the resources we provide.

In addition to using Canto community curation as the primary mechanism for data collection from newly published papers, PomBase invites researchers to submit large-scale datasets for phenotype, expression, and other annotations in spreadsheet-compatible formats as well as datasets suitable to appear on genome browser tracks. The most recent community curation submissions are showcased on the PomBase front page, and PomBase is exploring mechanisms for curation attribution via ORCIDs (https://orcid.org/).

We communicate with fission yeast researchers directly via our 1200-member community mailing list (pombelist) and at workshops and conferences, notably the biennial international *S. pombe* conference. To support PomBase usage, we run a helpdesk and maintain extensive documentation covering PomBase pages, annotation conventions, and Canto features. Advice on data usage and analysis disseminated via the helpdesk becomes incorporated into the extensive FAQ. Documentation and FAQs are available at http://www.pombase.org/help. We run periodic surveys to determine community priorities for new PomBase features and improvements to existing resources, and actively encourage corrections, improvements and suggestions to existing content of PomBase at all times.

Acknowledgments

References

1. McDowall MD, Harris MA, Lock A, Rutherford K, Staines DM, Bahler J, Kersey PJ, Oliver SG, Wood V (2015) PomBase 2015: updates to the fission yeast database. Nucleic Acids Res 43(Database issue):D656–D661. https://doi.org/10.1093/nar/gku1040

2. Wood V, Harris MA, McDowall MD, Rutherford K, Vaughan BW, Staines DM, Aslett M, Lock A, Bahler J, Kersey PJ, Oliver SG (2012) PomBase: a comprehensive online resource for fission yeast. Nucleic Acids Res 40(Database issue):D695–D699. https://doi.org/10.1093/nar/gkr853

3. Kim DU, Hayles J, Kim D, Wood V, Park HO, Won M, Yoo HS, Duhig T, Nam M, Palmer G, Han S, Jeffery L, Baek ST, Lee H, Shim YS, Lee M, Kim L, Heo KS, Noh EJ, Lee AR, Jang YJ, Chung KS, Choi SJ, Park JY, Park Y, Kim HM, Park SK, Park HJ, Kang EJ, Kim HB, Kang HS, Park HM, Kim K, Song K, Song KB, Nurse P, Hoe KL (2010) Analysis of a genome-wide set of gene deletions in the fission yeast *Schizosaccharomyces pombe*. Nat Biotechnol 28(6):617–623. https://doi.org/10.1038/nbt.1628

4. Spirek M, Benko Z, Carnecka M, Rumpf C, Cipak L, Batova M, Marova I, Nam M, Kim DU, Park HO, Hayles J, Hoe KL, Nurse P, Gregan J (2010) S. pombe genome deletion project: an update. Cell Cycle 9(12):2399–2402. https://doi.org/10.4161/cc.9.12.11914

5. Matsuyama A, Arai R, Yashiroda Y, Shirai A, Kamata A, Sekido S, Kobayashi Y, Hashimoto A, Hamamoto M, Hiraoka Y, Horinouchi S, Yoshida M (2006) ORFeome cloning and global analysis of protein localization in the fission yeast Schizosaccharomyces pombe. Nat Biotechnol 24(7):841–847. https://doi.org/10.1038/nbt1222

6. Wood V, Gwilliam R, Rajandream MA, Lyne M, Lyne R, Stewart A, Sgouros J, Peat N, Hayles J, Baker S, Basham D, Bowman S, Brooks K, Brown D, Brown S, Chillingworth T, Churcher C, Collins M, Connor R, Cronin A, Davis P, Feltwell T, Fraser A, Gentles S, Goble A, Hamlin N, Harris D, Hidalgo J, Hodgson G, Holroyd S, Hornsby T, Howarth S, Huckle EJ, Hunt S, Jagels K, James K, Jones L, Jones M, Leather S, McDonald S, McLean J, Mooney P, Moule S, Mungall K, Murphy L, Niblett D, Odell C, Oliver K, O'Neil S, Pearson D, Quail MA, Rabbinowitsch E, Rutherford K, Rutter S, Saunders D, Seeger K, Sharp S, Skelton J, Simmonds M, Squares R, Squares S, Stevens K, Taylor K, Taylor RG, Tivey A, Walsh S, Warren T, Whitehead S, Woodward J, Volckaert G, Aert R, Robben J, Grymonprez B, Weltjens I, Vanstreels E, Rieger M, Schafer M, Muller-Auer S, Gabel C, Fuchs M, Dusterhoft A, Fritzc C, Holzer E, Moestl D, Hilbert H, Borzym K, Langer I, Beck A, Lehrach H, Reinhardt R, Pohl TM, Eger P, Zimmermann W, Wedler H, Wambutt R, Purnelle B, Goffeau A, Cadieu E, Dreano S, Gloux S, Lelaure V, Mottier S, Galibert F, Aves SJ, Xiang Z, Hunt C, Moore K, Hurst SM, Lucas M, Rochet M, Gaillardin C, Tallada VA, Garzon A, Thode G, Daga RR, Cruzado L, Jimenez J, Sanchez M, del Rey F, Benito J, Dominguez A, Revuelta JL, Moreno S, Armstrong J, Forsburg SL, Cerutti L, Lowe T, McCombie WR, Paulsen I, Potashkin J, Shpakovski GV, Ussery D, Barrell BG, Nurse P (2002) The genome sequence of *Schizosaccharomyces pombe*. Nature 415(6874):871–880. https://doi.org/10.1038/nature724

7. Hoffman CS, Wood V, Fantes PA (2015) An ancient yeast for young geneticists: a primer on the *Schizosaccharomyces pombe* model system. Genetics 201(2):403–423. https://doi.org/10.1534/genetics.115.181503

8. Nguyen TT, Chua JK, Seah KS, Koo SH, Yee JY, Yang EG, Lim KK, Pang SY, Yuen A, Zhang L, Ang WH, Dymock B, Lee EJ, Chen ES (2016) Predicting chemotherapeutic drug combinations through gene network profiling. Sci Rep 6:18658. https://doi.org/10.1038/srep18658

9. Rhind N, Russell P (2012) Signaling pathways that regulate cell division. Cold Spring Harb Perspect Biol 4(10). https://doi.org/10.1101/cshperspect.a005942

10. Rosas-Murrieta NH, Rojas-Sánchez G, Reyes-Carmona SR, Martínez-Contreras RD, Martínez-Montiel N, Millán-Pérez-Peña L, Herrera-Camacho IP (2015) Study of cellular processes in higher eukaryotes using the yeast *Schizosaccharomyces pombe* as a model. In: Shah MM (ed) Microbiology in agriculture and human health. https://doi.org/10.5772/60720

11. Rutherford KM, Harris MA, Lock A, Oliver SG, Wood V (2014) Canto: an online tool for community literature curation. Bioinformatics 30(12):1791–1792. https://doi.org/10.1093/bioinformatics/btu103

12. Oliver SG, Lock A, Harris MA, Nurse P, Wood V (2016) Model organism databases: essential resources that need the support of both funders and users. BMC Biol 14:49. https://doi.org/10.1186/s12915-016-0276-z

13. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. The gene ontology consortium. Nat Genet 25(1):25–29. https://doi.org/10.1038/75556

14. Consortium GO (2015) Gene ontology consortium: going forward. Nucleic Acids Res 43(Database issue):D1049–D1056. https://doi.org/10.1093/nar/gku1179

15. Huntley RP, Harris MA, Alam-Faruque Y, Blake JA, Carbon S, Dietze H, Dimmer EC, Foulger RE, Hill DP, Khodiyar VK, Lock A, Lomax J, Lovering RC, Mutowo-Meullenet P, Sawford T, Van Auken K, Wood V, Mungall CJ (2014) A method for increasing expressivity of gene ontology annotations using a compositional approach. BMC Bioinformatics 15:155. https://doi.org/10.1186/1471-2105-15-155

16. Harris MA, Lock A, Bahler J, Oliver SG, Wood V (2013) FYPO: the fission yeast phenotype ontology. Bioinformatics 29(13):1671–1678. https://doi.org/10.1093/bioinformatics/btt266

17. Wood V (2005) *Schizosaccharomyces pombe* comparative genomics; from sequence to systems. In: Sunnerhagen P, Piskur J (eds) Topics in current genetics, vol 15. Springer, Berlin, pp 233–285. https://doi.org/10.1007/4735_97

18. Eilbeck K, Lewis SE, Mungall CJ, Yandell M, Stein L, Durbin R, Ashburner M (2005) The sequence ontology: a tool for the unification of genome annotations. Genome Biol 6(5):R44. https://doi.org/10.1186/gb-2005-6-5-r44

19. Hastings J, de Matos P, Dekker A, Ennis M, Harsha B, Kale N, Muthukrishnan V, Owen G, Turner S, Williams M, Steinbeck C (2013) The ChEBI reference database and ontology for biologically relevant chemistry: enhancements for 2013. Nucleic Acids Res 41(Database issue):D456–D463. https://doi.org/10.1093/nar/gks1146

20. Montecchi-Palazzi L, Beavis R, Binz PA, Chalkley RJ, Cottrell J, Creasy D, Shofstahl J, Seymour SL, Garavelli JS (2008) The PSI-MOD community standard for representation of protein modification data. Nat Biotechnol 26(8):864–866. https://doi.org/10.1038/nbt0808-864

21. Marguerat S, Schmidt A, Codlin S, Chen W, Aebersold R, Bahler J (2012) Quantitative analysis of fission yeast transcriptomes and proteomes in proliferating and quiescent cells. Cell 151(3):671–683. https://doi.org/10.1016/j.cell.2012.09.019

22. Rhind N, Chen Z, Yassour M, Thompson DA, Haas BJ, Habib N, Wapinski I, Roy S, Lin MF, Heiman DI, Young SK, Furuya K, Guo Y, Pidoux A, Chen HM, Robbertse B, Goldberg JM, Aoki K, Bayne EH, Berlin AM, Desjardins CA, Dobbs E, Dukaj L, Fan L, FitzGerald MG, French C, Gujja S, Hansen K, Keifenheim D, Levin JZ, Mosher RA, Muller CA, Pfiffner J, Priest M, Russ C, Smialowska A, Swoboda P, Sykes SM, Vaughn M, Vengrova S, Yoder R, Zeng Q, Allshire R, Baulcombe D, Birren BW, Brown W, Ekwall K, Kellis M, Leatherwood J, Levin H, Margalit H, Martienssen R, Nieduszynski CA, Spatafora JW, Friedman N, Dalgaard JZ, Baumann P, Niki H, Regev A, Nusbaum C (2011) Comparative functional genomics of the fission yeasts. Science 332(6032):930–936. https://doi.org/10.1126/science.1203357

23. Soriano I, Quintales L, Antequera F (2013) Clustered regulatory elements at nucleosome-depleted regions punctuate a constant nucleosomal landscape in Schizosaccharomyces pombe. BMC Genomics 14:813. https://doi.org/10.1186/1471-2164-14-813

24. Xu J, Yanagisawa Y, Tsankov AM, Hart C, Aoki K, Kommajosyula N, Steinmann KE, Bochicchio J, Russ C, Regev A, Rando OJ, Nusbaum C, Niki H, Milos P, Weng Z, Rhind N (2012) Genome-wide identification and characterization of replication origins by deep sequencing. Genome Biol 13(4):R27. https://doi.org/10.1186/gb-2012-13-4-r27

25. Mata J (2013) Genome-wide mapping of poly-adenylation sites in fission yeast reveals widespread alternative polyadenylation. RNA Biol 10(8):1407–1414. https://doi.org/10.4161/rna.25758

26. Schlackow M, Marguerat S, Proudfoot NJ, Bahler J, Erban R, Gullerova M (2013) Genome-wide analysis of poly(A) site selection in Schizosaccharomyces pombe. RNA 19(12):1617–1631. https://doi.org/10.1261/rna.040675.113

27. Woolcock KJ, Gaidatzis D, Punga T, Buhler M (2011) Dicer associates with chromatin to repress genome activity in *Schizosaccharomyces pombe*. Nat Struct Mol Biol 18(1):94–99. https://doi.org/10.1038/nsmb.1935

28. Skinner ME, Uzilov AV, Stein LD, Mungall CJ, Holmes IH (2009) JBrowse: a next-generation genome browser. Genome Res 19(9):1630–1638. https://doi.org/10.1101/gr.094607.109