SOFTWARE TOOL ARTICLE

Oncoyeast: a web-based application to translate data obtained from *Saccharomyces cerevisiae* high-throughput drug screens into cancer therapeutics [version 1; peer review: 1 approved with reservations, 2 not approved]

Ruby Gupta^{1*}, Samir Cayenne^{2*}, Madhu Dyavaiah^{3*}, Pragnya Srinivas^{4*}, David Otohinoi^{5*}, Debjyoti Talukdar^{5*}, Moheem Halari^{5*}, Chidambra Halari^{5*}, Ashok Ramani^{6*}, Joshua Yusuf⁵, Khushdeep Chahal⁷, Rupinder Kaur⁷, Ankit Patel^{id}⁸, Avaniben Patel^{id}⁸, Ravindrasingh Rajput⁸, Harish Siddaiah⁹, Shilpadevi Patil^{9*}, Ashish Patil^{1,10*}, Nikhilesh Anand^{id}^{5*}

¹Providence Hospital, Washington, D.C., USA

²Clear Springs High School, League City, Texas, USA

³Pondicherry University, Puducherry, India

⁴Narayana College, Bangalore, India

⁵All Saints University School of Medicine, Roseau, Dominica

⁶American University of Antigua, Coolidge, Antigua and Barbuda

⁷School of Medicine, University of Alabama, Huntsville, Alabama, USA

⁸Spartan Health Sciences University, Vieux Fort, Saint Lucia

⁹Department of Anesthesiology, LSU Health Sciences Center Shreveport, Shreveport, Louisiana, USA

¹⁰Department of Pathology and Translational Pathobiology, Louisiana State University Health Sciences Center, Shreveport, LA, USA

* Equal contributors

V1 First published: 18 Jun 2018, 7:757
<https://doi.org/10.12688/f1000research.15168.1>
 Latest published: 18 Jun 2018, 7:757
<https://doi.org/10.12688/f1000research.15168.1>

Abstract

The budding yeast (*Saccharomyces cerevisiae*) gene deletion library consists of a collection of more than 6,000 gene-deletion mutants and is useful for high-throughput screening of anti-cancer drugs. Because of the shorter doubling time and the significant homology the budding yeast shares with human cells, using a high-throughput chemical screen of budding yeast gene deletion library, one can rapidly identify various genetic targets of anti-cancer drugs. But analyzing the data derived from a yeast library screen to identify corresponding human homologs and their status in various cancers is a cumbersome process. We have developed a web-based app, **Oncoyeast**, which enables the researcher to automatically identify the corresponding human homologs of *S. cerevisiae* and the status of these homologs genes in tumor samples from The Cancer Genome

Open Peer Review

Approval Status ? X X

	1	2	3
version 1	?	X	X
18 Jun 2018	view	view	view
1. Ayaz Najafov ^{id} , Harvard Medical School , Boston, USA			
2. Alvaro Galli ^{id} , Institute of Clinical Physiology, National Research Council (CNR), Pisa, Italy			
3. Robert J. D. Reid , Columbia University			

Atlas Database and cell line samples from the Cancer Cell Line Encyclopedia. This would enable the scientists to identify the tumors and choose cell lines to work on and thus serve as an indispensable tool to translate their research into human cancers.

Keywords

saccharomyces cerevisiae, Cancer, drug screening, toxicity screen, human homologs, cancer cell lines, TCGA, CCLE, translational research

Medical Center, New York, USA

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding authors: Shilpadevi Patil (spati1@lsuhsc.edu), Ashish Patil (aravin@lsuhsc.edu)

Author roles: Gupta R: Writing – Original Draft Preparation; Cayenne S: Data Curation; Dyavaiah M: Project Administration; Srinivas P: Data Curation; Othinoi D: Software; Talukdar D: Software; Halari M: Writing – Review & Editing; Halari C: Writing – Review & Editing; Ramani A: Resources, Writing – Review & Editing; Yusuf J: Resources; Chahal K: Resources; Kaur R: Resources; Patel A: Writing – Review & Editing; Patel A: Writing – Review & Editing; Rajput R: Writing – Review & Editing; Siddaiah H: Formal Analysis; Patil S: Supervision; Patil A: Conceptualization; Anand N: Investigation, Supervision

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Copyright: © 2018 Gupta R *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 work is properly cited.

How to cite this article: Gupta R, Cayenne S, Dyavaiah M *et al.* **Oncoyeast: a web-based application to translate data obtained from *Saccharomyces cerevisiae* high-throughput drug screens into cancer therapeutics [version 1; peer review: 1 approved with reservations, 2 not approved]** F1000Research 2018, 7:757 <https://doi.org/10.12688/f1000research.15168.1>

First published: 18 Jun 2018, 7:757 <https://doi.org/10.12688/f1000research.15168.1>

Introduction

Cancer chemotherapies, the backbone of cancer, treatment have multiple shortcomings including significant toxicities, narrow therapeutic indexes, and the emergence of resistance. Now, a better understanding of cancer pathogenesis and genomics has given rise to new treatment options in terms of targeted therapies^{1,2}. Targeted approaches aim to modulate the molecular pathways that are crucial for tumor growth and maintenance, without causing significant systemic toxicities. This leads to the selective killing of cancer cells with more precision without affecting normal cells, thus generating fewer side effects². For example, imatinib is commonly used for the treatment of the chronic myeloid leukemia (CML), targeting fusion protein BCR-ABL, and in the treatment of gastrointestinal stromal tumors (GIST) that harbor cKIT and PDGFRA mutations³⁻⁵. In addition, there are case reports that have documented the regression of cutaneous squamous cell cancer in patients with CML treated with imatinib⁶. Another example is rituximab, one of the earliest targeted therapies which are antibodies against cell surface marker CD20 used in patients with leukemia and lymphoma^{7,8}.

Although this type of molecular profiling led to tailored therapy for cancer patients, it did not take a long time to discover that it has its own shortcomings. It was found that although these therapeutic agents produce dramatic effects in most patients, some did not respond at all and others showed only a partial response to the treatment^{1,9,10}. One of the mechanisms attributed to the development of resistance is intra-tumor heterogeneity⁹⁻¹¹.

Intra-tumor heterogeneity

Diversity within cells of a single neoplastic lesion is called intra-tumor heterogeneity. This is due to genome instability attributed by the pool of endogenous and exogenous mutations in a given microenvironmental context (selection pressure), which is known to alter the evolutionary trajectory of a tumor leading to clonal and subclonal expansions of cancer cells. This ultimately results in a genetically diversified tumor¹²⁻¹⁴. For example, in a study of a patient with malignant melanoma who initially responded to the targeted BRAF inhibitor vemurafenib, but later progressed and developed five metastatic lesions, it was found that resistance to vemurafenib was developed via several different, independent mechanisms. BRAFV600E amplification was found in three resistant metastases, which likely occurred as independent events^{15,16}. The fourth drug-resistant metastatic lesion harbored an aberrant form of BRAF, whereas, in yet another metastasis, a new activating in-frame insertion in MEK was identified, thus highlighting the extent of inter- and intra-tumoral heterogeneity and its impact on clinical outcomes¹⁷.

Studies have shown that intra-tumoral heterogeneity is the key mechanism underlying tumor progression and the frequent lack of therapeutic responses¹⁷. Intra-tumor variability has not only posed a significant challenge in predicting the behavior of the neoplasm, but has also resulted in the emergence of neoplastic clone's resistance to a given therapy^{12,17}. It is critical to explore novel methodologies to gain a better insight into the biological and therapeutic impact of intra-tumor genetic heterogeneity for the improvement of existing therapies in cancer management.

Inter-tumoral heterogeneity

Studies have shown that two tumors that are histologically similar may behave and respond differently to the same treatment due to being genetically different—a phenomenon that can now be explained by the concept of inter-tumor heterogeneity¹⁸. Recent studies that sequenced the genomes of breast and colorectal cancers have shown us that the tumor evolution is a complicated multi-step process caused by a combination of series of mutations intermixed in a vast genomic landscape, confirming tumor heterogeneity^{19,20}. Inter-tumor heterogeneity has significant clinical implications. For example, among patients with colon cancer there has been an attempt to classify the disease with respect to genetic defects. For example, activating mutations of the *BRAF* oncogene are associated with microsatellite instability and mutations in *TP53* or *KRAS* are associated with chromosomal instability²¹. This understanding of inter-tumor heterogeneity in colorectal cancers has significant clinical implications, not only for predicting the prognosis of the patient but also for tailoring the targeted therapy to be used for individual tumor type. Approximately 60% of colon cancer patients have *KRAS* wild-type mutations and are being treated with EGFR inhibitors, but only 30% benefit from the treatment. Non-response to EGFR inhibitors in the remaining 30% of patients has been attributed to other genetic biomarkers, such as *BRAF* mutations (8–10%), *PTEN* null mutations (10–15%) and *PI3KC* mutations (10–12%)^{18,19}. It has been proposed that these patients with different genetic profiles may benefit from therapies targeted towards individual biomarker thus reminding us the clinical and therapeutic implications of inter-tumoral heterogeneity^{15,16,18}.

The Cancer Genome Atlas (TCGA) is an ever-expanding catalog of genetic mutations identified in tumors in different patients that are responsible for causing cancer. So far it has cataloged and characterized genomic changes including copy number variations, mutations, and RNA expression in around 33 cancer types in more than 45,000 patients, including for 10 rare cancers^{22,23}. TCGA Data Portal enables the researcher to download copy number variations, mutations and mRNA expression profiles of all the genes in a particular type of tumor^{24,25}. However, if the researcher is interested in studying a particular type of genetic alteration across a set of different cancer types and across a set of patients then they will have to use interactive exploratory tools like the cBio Portal to perform data analysis across TCGA Database²⁶.

TCGA also contains the dataset from the Cancer Cell Line Encyclopedia (CCLE) project, which provides information concerning DNA copy number variation, mRNA expression profile and, mutational data of genes in more than 1000 cancer cell lines, providing an indispensable tool for scientists conducting pharmacogenomics and pharmacotherapeutics studies in oncology. The CCLE has been successfully used in the genetic prediction of genomic profile-based drug responses in the preclinical setting and its incorporation into cancer clinical trial design has helped in the emergence of “personalized” therapeutic regimens²⁷.

Using the cBioPortal to interactively explore cancer genomics datasets, the researcher can access data from more than 45,000

tumor samples from cancer studies in TCGA database. This enables the scientists to choose cancer cell lines and type of tumor models to pursue their research of targeted cancer therapy²⁶.

Model organism

The principal objective of cancer genomics research has been to identify alterations in gene expression, provide information on mutations and copy number variations, the sequences of cancer genes and data on the mutational processes that occur are functioning in tumors, and to develop genetic profile-based targeted therapies against these tumors during cancer evolution. One of the challenges in doing so is finding a method wherein genetic studies can be performed conveniently and rapidly²⁸. Many practical and ethical obstacles severely limit the scope of experiments using human cells in genetic studies. One of the biggest limitations is that the human cell lines are diploid and take a longer time to divide than yeast or bacterial cells, making it difficult to perform manipulation of genes in them. In addition, the human cell lines take a long time to divide, thus limiting the ability to use human cell lines in high-throughput screening of drugs and gene manipulation studies in humans. To overcome this limitation, model organisms have been used, in which most of the fundamental biological mechanisms and pathways that control development and survival have been evolutionarily conserved between species. These organisms were found to have genes that have structural and functional homology with human genes²⁹.

Budding yeast (*Saccharomyces cerevisiae*) is one of the model organism used to analyze phenomena that involves important basic eukaryotic cell functions, such as metabolism, regulation of the cell cycle, membrane targeting and dynamics, protein folding, and DNA repair. The *S. cerevisiae* genome was first fully sequenced in 1996, which opened the gateway to various researchers as it provided a treasure of information on genome organization and evolution³⁰.

Now, *S. cerevisiae* has also been considered a principal model organism for conducting research owing to the ease with which gene manipulation can be done, short doubling time harbored by these cells in and significant genotypic and phenotypic homology shared with human cells. In addition, the complete mapping of

the yeast genome and creation of the yeast gene deletion collection (**library of deletion mutants lacking all non-essential genes**) has resulted in a shift in research focus from individual genes to a more global view of genetic networks. This has resulted in the development of high-throughput screens to identify multiple genetic targets of a drug using the *S. cerevisiae* gene deletion library (**Figure 1**)³¹. Data from these screens helps researchers to identify novel drug candidates

High-throughput screening using the yeast gene deletion library enables researchers to study small molecules in chemical libraries that induce a phenotype in the clones that are engineered to reflect specific genetic changes in human cancers or obtain essential information for further drug design^{31,32}. This eukaryotic microorganism shares a number of fundamental cellular and molecular properties with human cells. This homology has been useful in the study of gene structure and function in *S. cerevisiae* and the findings have been applied to human biology^{30,33}. Although the entire genetic concentration in a yeast packs into 16 chromosomes, containing only 10% of the DNA of human chromosomes, the packaging of the genetic material in this micro-organism is much more compressed and dense than the human genome, making it fit for use in studies related to human biology³⁴.

Once a researcher identifies the sensitive *S. cerevisiae* target genes from the high-throughput drug screen, bioinformatics tools can be used to identify corresponding human homologs and then browse through TCGA using gateways like cBioPortal to identify the status of these homologs in different tumors and cancer cell lines. This process is time-consuming and there is an increased chance of human errors, especially when a researcher gets a list of many genes from the screen.

In our study, we have developed a web-based application, Oncoyeasti, that matches genes *S. cerevisiae* genes with corresponding human homologs and browses through the TCGA and cBioPortal database to show the status of these matched human homolog genes in various cancers. Our web app not only saves time and resources the researcher needs to commit in order to analyze the data obtained from the drug screen, it also reduces chances of human errors because the entire data-analysis is streamlined.

Figure 1. The mutant clones are plated on a yeast extract peptone dextrose (YPD) agar plate containing the cancer drug of interest and incubated for 2 to 3 days at 30°C. The clones with gene deletions that lead to drug sensitivity will not grow or grow slowly compared to the wild-type or non-sensitive clones after 3 days of incubation and will not form colonies. The YPD-agar plates are scanned to identify genes deletion leading to sensitivity for a given anti-cancer drug tested.

Methods

Core functionality and architecture

The main objective of Oncoyeasti is to enable the oncologist to use *S. cerevisiae* as a screening tool to rapidly identify genetic targets of anti-cancer drugs and to establish the status of the human homologs of the identified targets in the tumor database of TCGA. It is a cumbersome task to translate the results obtained from a *S. cerevisiae* high-throughput drug screen into application-oriented cancer research into cancerous tumors and cancerous cell lines, especially when the drug screen results in the identification of multiple genetic targets. We have developed Oncoyeasti to serve as a platform that analyzes the data obtained from the yeast screen, to identify the corresponding human homologs and status of these homologs in different cancers and cancer cell lines data deposited in TCGA (Figure 2). Oncoyeasti has been programmed to browse TCGA database by automatically using the interactive data exploration and visualization.

cBioPortal provides a visual output of the status of human homologs of *S. cerevisiae* genes in various cancers. At the time of writing, the TCGA database had a genetic profile of more than 45,000 tumor samples and more than a 1000 cancer cell lines. The sample size is expected to continuously increase with time by updating the database with the sequencing of newer tumors.

System implementation

In order to have a platform that provides scientist the human homologs of yeast genes for easy anticancer drug screening, we used the server-side scripting language PHP for web development platform along with MYSQL interface for matching the input *S. cerevisiae* genes with the corresponding human homologs. The URL for the cross cancer summary of the matched human homologs in tumors present in TCGA database, and navigated using cBioPortal interface, was generated using the MSQl interface. Using the MSQl interface and cBioPortal the URL for the status of the matched human homologs in the tumor cell lines data present in the CCLE was generated. We used a simple approach to

build the algorithm for the search program which has the capability to analyze up to 100 *S. cerevisiae* genes at once. The algorithm functions as follows. First, it takes the input in form of *S. cerevisiae* standard or systematic gene names and starts searching the human homologs dataset for a possible match by comparing the input text with *S. cerevisiae* standard and systematic gene names. Once the matched human homologs are identified, the cBioPortal URL for the status of the corresponding homolog in various cancer datasets is automatically generated. Using this technique, we are able to display multiple search results for corresponding human homolog gene data stored in the database table along with URL links to the cBioPortal/TCGA database. The design of the web based application **Oncoyeasti** is unique in terms of identification of human homologue *S. cerevisiae* genes. It uses feature rich, fast JQuery API which works across multitude of browsers. The project features improved page start-up performance with modified properties and methods which can load JavaScript events and data from the server and return the matched element. Apart from that, the w3 container class uses bootstrap grid system that allows us to organize the data in grid system where classes can be combined in a more responsive, flexible and dynamic layouts.

Oncoyeasti is implemented in **Bluehost**, a highly versatile and reliable web hosting company. Their platform can efficiently support large volumes of traffic at remarkable speed. As a result, the time lapse from the dataset search of Oncoyeasti database to the automatic retrieval of corresponding links from the cBioPortal/TCGA database is estimated to be 761 milliseconds, depending on the internet service provider.

Operating systems and browsers

The usability of oncoyeasti is observed to have efficient function among old and new operating system and browsers. Oncoyeasti has a minimal requirement of Internet Explorer 9, Firefox v60, Safari (all versions), Opera v51, Chrome v66, iOS (all versions), and Android (all versions).

Figure 2. Flow chart describing the algorithm of Oncoyeasti. Once the researcher input the *Saccharomyces cerevisiae* genes, Oncoyeasti matches the *S. cerevisiae* genes with the corresponding human homologs and generates a link for those homologs to the corresponding cBioPortal webpage describing the status of these homologs in tumors sequence data in The Cancer Genome Atlas (TCGA) database and cancer cell line genomic alteration data in the Cancer Cell Line Encyclopedia (CCLE) database.

Use cases

Data presentation and visualization

Once you visit the portal (<http://www.oncoyeasti.org/>), you will be prompted to enter *S. cerevisiae* gene names (Figure 3). You can enter up to 100 gene names, with 1 gene per line. The gene names can be either in standard or systematic gene name format. For example, the *RAD9* gene has the standard name *RAD9* and systematic name *YDR217C*, you can either enter *RAD9* or *YDR217C* and click submit.

Once you click submit, Oncoyeasti will process your entry and match your *S. cerevisiae* gene entry with the corresponding human homolog. In the case of *RAD9* it is *TP53BP1* (Figure 4). It will also automatically generate a link to <http://www.cbioportal.org/> that will provide information of the status of *TP53BP1* in more than 45,000 tumor samples, in the TCGA database.

This link is provided under the section cross cancer summary of the human homolog in tumors present in TCGA database. In addition, generate a link to check for the status of the *TP53BP1* in the cancer cell lines sequence data present in the CCLE database.

In order to check the status of *TP53BP1* in TCGA database, one needs to click on the Cross Cancer Summary Tab. Upon clicking on the corresponding cross cancer summary tab for *RAD9*, we will reach the cBioPortal page visually depicting the cross-cancer alteration summary for *TP53BP1* in tumor profiles present in 147 different cancer studies (Figure 5). The visual depiction of the data is in a histogram form, showing a percentage of various types of chromosomal instabilities likes homozygous deletion indicated by blue, homozygous amplification indicated by red and mutations indicated by green.

Figure 3. On visiting Oncoyeasti.org the researchers can input up to 100 *Saccharomyces cerevisiae* genes in standard gene name or systematic gene name format. In this case, *RAD9* is the input.

Results					
S/N	S Cerevisiae Standard Gene Name	S Cerevisiae Systematic Gene Name	Homologous Homo Sapien Gene/Genes	Cross Cancer Summary in Tumors present in TCGA Database	CCLE Cancer Cell Lines Summary
1	RAD9	YDR217C	TP53BP1	Cross Cancer Summary	CCLE Cell Lines Summary

Figure 4. Oncoyeasti identifies the human homolog of the corresponding *Saccharomyces cerevisiae* genes, and generates a link to the cBioPortal visualization portal to browse the status of the corresponding human homologs in tumor samples from The Cancer Genome Atlas (TCGA) Database and Cancer Cell Lines Encyclopedia (CCLE).

Figure 5. Oncocyte generates a link to cross cancer summary for corresponding human homolog gene (in this case, *TP53BP1* for *RAD9*). The link directs the researcher to the cBioPortal page for the cross cancer summary of the corresponding human homolog (*TP53BP1* in this case), where in one gets a histogram depicting cross-cancer alteration summary for *TP53BP1* in approximately 147 studies in The Cancer Genome Atlas database. Image obtained from cBioPortal^{26,35}.

You can click on the histogram bars for the respective cancer study to get an “oncoprint”, which is the visualization of distinct genomic alterations in different patients in a given study, such as copy number alterations, somatic mutations, and mRNA expression. Individual genes are represented as rows, and individual cases or patients are represented as columns. For example, if we click on the uterine corpus endometrial carcinoma (provisional TCGA database of 242 samples), we will be directed to the oncoprint of *TP53BP1* gene status in 242 uterine corpus endometrial carcinoma samples (Figure 6). The tumors with homozygous amplification of the given gene in an oncoprint (*TP53BP1* in this case) is indicated by the red bar and homozygous deletion is indicated by the blue bar.

Data export from cBioPortal

One can download the copy number alterations by clicking on the download tab and export it into an excel form to perform a copy number analysis for the gene. Once exported into the excel, -2 indicates homozygous deletion, -1 heterozygous deletion, 0 is diploid, 1 indicates a heterozygous amplification and 2 a homozygous amplification

Similarly, if the researcher wants to pick a cancer cell line, depending on the yeast phenotypic screen data, he can click link

generated by Oncocyte under the CCLE column. The link will direct the researcher to cBioPortal CCLE section and provide him with an oncoprint for the gene of interest (*TP53BP1*, a *RAD9* homolog in this case). This will enable the researcher to visualize the copy number alterations of the gene of interest in more than 1000 different types of cancer cell lines, perform copy number alteration analysis in excel as described previously and thus select appropriate cancer cells lines depending on their *S. cerevisiae* genetic screen results.

Discussion

S. cerevisiae is one of the most widely used eukaryotic model organisms³⁶. It has been used to study biological processes including cell cycle, gene regulation, signal transduction, cell cycle, apoptosis, aging and neurodegenerative disorders^{29,36}. Up to 30% of genes implicated in human disease may have orthologs in the yeast³⁷. Out of an estimated 2,271 known disease-associated genes in humans, 526 (~23%) have a close ortholog in the yeast genome. Similarities in the gene sequence and function have made this organism useful in elucidating biological pathways in humans. Yeast genomic methodologies have been applied in drug development studies including drug target identification, identification of targeted genes and pathways of therapeutic drugs³⁸. The genomic methodologies include use of deletion mutants for

Figure 6. To look at the genetic alterations of tumors in an individual study, in this case (alterations of *TP53BP1* in uterine corpus endometrial carcinoma), one clicks on Uterine Corpus Endometrial Carcinoma Bar in Figure 5 and it cBioPortal directs us to oncoprint, which is a graphical visualization of genetic alterations for a gene of interest in different individuals. In an oncoprint, each tumor is indicated by a bar. Red color indicates homozygous amplification for the given gene and blue color indicates homozygous deletion for the gene of interest. Image obtained from cBioPortal^[26,35].

drug-target identification, cDNA microarray analysis of drug-target action, complete synthetic screens, fitness profiling using barcoded mutants and analysis of cellular responses to external stimuli were developed in yeast^{38,39}. In addition, individual pathways such as cell cycle regulators and histone deacetylase inhibitors have been used in the identification of new anticancer drug targets³⁹. From all these yeast genomic technologies, a list of genes that respond to a given drug will be available. For example, yeast gene deletion library screening to a particular drug, we get a list of genes whose deletion resulted in cell sensitivity or resistant to a particular drug⁴⁰. In other yeast technologies, such as cDNA microarray analysis of drug-target action and fitness profiling using barcoded mutants, experiments also resulted in a list of genes up-regulated or down-regulated in response to the drug⁴¹. The ultimate aim of the yeast model in either drug discovery or the study of pathways is to apply to humans and thus finding the drug responsive yeast genes in a human would be difficult for a large number of genes. Therefore, Oncocyte, a web-based application represents a useful tool for researchers who are using yeast phenotypic or genotypic screens to rapidly identify genetic targets of anticancer drugs. It enables them to identify the corresponding human homologs of the *S. cerevisiae* genes of interest and automatically generates a link to cBioPortal to visualize the genetic alterations of the corresponding homologs in more than 45,000 different tumor sample datasets present in the TCGA database. As shown in Figure 4, Oncocyte identifies the sequential human homolog (*TP53BP1*) of the corresponding *S. cerevisiae* gene *RAD9* and the corresponding human homologs in tumor samples of TCGA database and cancer cell lines. In addition, Oncocyte generates a corresponding link to cBioPortal showing histogram depicting cross-cancer alteration summary of *TP53BP1*

in 147 studies in TCGA and CCLE (Figure 5). By clicking on the bar of CCLE dataset in the histogram, researchers can navigate CCLE datasets to identify the genetic alterations of their genes of interest (*TP53BP1* in this case) in the cancer cell lines datasets present in the CCLE database. Thus it can serve as a useful tool for the researchers to predict different cancerous tumors; the data derived from the yeast screen can be used and serve as a framework to select tumors or cancer subtypes in which their research would be useful. It also serves as a tool for researchers to select appropriate cancer cell lines for their further studies and thus translate their research from budding yeast to humans.

Oncocyte identifies human target genes and the tumors in which these genes are perturbed based on a sequential homology with *S. cerevisiae*. But homolog variants identified using sequential homology from a model organism like *S. cerevisiae* are assumed to be well tolerated in humans, but in certain cases may not perform the same function in humans and vice versa, in functional complementation studies⁴². Hamza *et al.* showed that approximately 60% of *S. cerevisiae* null mutant clones can be rescued by functional complementation by corresponding human homolog genes, indicating that utility of the cross-species platform to screen and identify targets based on sequential homology⁴³. Since 60% of budding yeast homologs show functional complementation, the human cancer cell lines with genetic targets homologous to the yeast mutants identified in the chemical screen should behave similarly to the corresponding yeast mutants identified in the high-throughput chemical screen (the chemical used in this case may be a potential therapeutic drug). Hence Oncocyte, when used in synchrony with yeast gene deletion drug screen can serve as an indispensable tool for scientists to rapidly identify target tumors

and cancer cell lines that are sensitive or resistant to a given drug and translate yeast research into human therapeutics.

Conclusion

The ultimate aim of understanding cancer biology using model organisms is to apply this knowledge to human cancer and cell lines of interest for drug discovery. Oncoyeast serves as a useful tool for scientists who are working with the yeast model to understand cancer gene functions, pathways and drug discovery.

Data availability

All data underlying the results are available as part of the article and source code, and no additional source data are required.

Software availability

Oncoyeast available from: <http://www.oncoyeasti.org/>

Source code available from: <https://github.com/oncoadm/onco-yeasti.org>.

Archived source code as at time of publication: <https://doi.org/10.5281/zenodo.1257717>³⁵.

License: MIT License.

Competing interests

No competing interests were disclosed.

Grant information

The author(s) declared that no grants were involved in supporting this work.

Acknowledgements

Special thanks to Dr. Stacia R. Engel from the *Saccharomyces* Genome Database (Department of Genetics, Stanford University) for providing us curated human homolog gene list of *Saccharomyces cerevisiae*.

References

- Aparicio S, Caldas C: **The implications of clonal genome evolution for cancer medicine.** *N Engl J Med.* 2013; **368**(9): 842–51.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Saijo N: **Progress in cancer chemotherapy with special stress on molecular-targeted therapy.** *Jpn J Clin Oncol.* 2010; **40**(9): 855–62.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Din OS, Woll PJ: **Treatment of gastrointestinal stromal tumor: focus on imatinib mesylate.** *Ther Clin Risk Manag.* 2008; **4**(1): 149–62.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Maleddu A, Pantaleo MA, Nannini M, *et al.*: **The role of mutational analysis of KIT and PDGFRA in gastrointestinal stromal tumors in a clinical setting.** *J Transl Med.* 2011; **9**: 75.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Hochhaus A, Larson RA, Guilhot F, *et al.*: **Long-Term Outcomes of Imatinib Treatment for Chronic Myeloid Leukemia.** *N Engl J Med.* 2017; **376**(10): 917–927.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kawakami Y, Nakamura K, Nishibu A, *et al.*: **Regression of cutaneous squamous cell carcinoma in a patient with chronic myeloid leukaemia on imatinib mesylate treatment.** *Acta Derm Venereol.* 2008; **88**(2): 185–6.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Maury S, Chevret S, Thomas X, *et al.*: **Rituximab in B-Lineage Adult Acute Lymphoblastic Leukemia.** *N Engl J Med.* 2016; **375**(11): 1044–53.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Cheson BD, Leonard JP: **Monoclonal antibody therapy for B-cell non-Hodgkin's lymphoma.** *N Engl J Med.* 2008; **359**(6): 13–26.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Rivera G, Wakelee HA: **Resistance to Therapy.** *Cancer Treat Res.* 2016; **170**: 183–202.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Suda K, Mitsudomi T: **Successes and limitations of targeted cancer therapy in lung cancer.** *Prog Tumor Res.* 2014; **41**: 62–77.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Sun XX, Yu Q: **Intra-tumor heterogeneity of cancer cells and its implications for cancer treatment.** *Acta Pharmacol Sin.* 2015; **36**(10): 1219–27.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- McGranahan N, Swanton C: **Biological and therapeutic impact of intratumor heterogeneity in cancer evolution.** *Cancer Cell.* 2015; **27**(1): 15–26.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Iwasa Y, Michor F: **Evolutionary dynamics of intratumor heterogeneity.** *PLoS One.* 2011; **6**(3): e17866.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Sottoriva A, Spiteri I, Piccirillo SG, *et al.*: **Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics.** *Proc Natl Acad Sci U S A.* 2013; **110**(10): 4009–14.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Grimaldi AM, Simeone E, Festino L, *et al.*: **Novel mechanisms and therapeutic approaches in melanoma: targeting the MAPK pathway.** *Discov Med.* 2015; **19**(107): 455–61.
[PubMed Abstract](#)
- Holderfield M, Deuker MM, McCormick F, *et al.*: **Targeting RAF kinases for cancer therapy: BRAF-mutated melanoma and beyond.** *Nat Rev Cancer.* 2014; **14**(7): 455–67.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Burrell RA, Swanton C: **Tumour heterogeneity and the evolution of polyclonal drug resistance.** *Mol Oncol.* 2014; **8**(6): 1095–111.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Cusnir M, Cavalcante L: **Inter-tumor heterogeneity.** *Hum Vaccin Immunother.* 2012; **8**(8): 1143–5.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- De Sousa EMF, Vermeulen L, Fessler E, *et al.*: **Cancer heterogeneity--a multifaceted view.** *EMBO Rep.* 2013; **14**(8): 686–95.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Polyak K: **Heterogeneity in breast cancer.** *J Clin Invest.* 2011; **121**(10): 3786–8.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Markowitz SD, Bertagnoli MM: **Molecular origins of cancer: Molecular basis of colorectal cancer.** *N Engl J Med.* 2009; **361**(25): 2449–60.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Wang Z, Jensen MA, Zenklusen JC: **A Practical Guide to The Cancer Genome Atlas (TCGA).** *Methods Mol Biol.* 2016; **1418**: 111–41.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Lee JS: **Exploring cancer genomic data from the cancer genome atlas project.** *BMB Rep.* 2016; **49**(11): 607–611.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- McCain J: **The cancer genome atlas: new weapon in old war?** *Biotechnol Healthc.* 2006; **3**(2): 46–51B.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, *et al.*: **The Cancer Genome Atlas Pan-Cancer analysis project.** *Nat Genet.* 2013; **45**(10): 1113–20.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Gao J, Aksoy BA, Dogrusoz U, *et al.*: **Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal.** *Sci Signal.* 2013; **6**(269): pii1.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Barretina J, Caponigro G, Stransky N, *et al.*: **The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity.** *Nature.* 2012; **483**(7391): 603–7.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Hunter P: **The paradox of model organisms. The use of model organisms in research will continue despite their shortcomings.** *EMBO Rep.* 2008; **9**(8): 717–20.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

29. Duina AA, Miller ME, Keeney JB: **Budding yeast for budding geneticists: a primer on the *Saccharomyces cerevisiae* model system.** *Genetics*. 2014; **197**(1): 33–48.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
30. Engel SR, Dietrich FS, Fisk DG, *et al.*: **The reference genome sequence of *Saccharomyces cerevisiae*: then and now.** *G3 (Bethesda)*. 2014; **4**(3): 389–98.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
31. Huang S, O'Shea EK: **A systematic high-throughput screen of a yeast deletion collection for mutants defective in *PHO5* regulation.** *Genetics*. 2005; **169**(4): 1859–71.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
32. Hammonds TR, Maxwell A, Jenkins JR: **Use of a rapid throughput *in vivo* screen to investigate inhibitors of eukaryotic topoisomerase II enzymes.** *Antimicrob Agents Chemother*. 1998; **42**(4): 889–94.
[PubMed Abstract](#) | [Free Full Text](#)
33. Botstein D, Chervitz SA, Cherry JM: **Yeast as a model organism.** *Science*. 1997; **277**(5330): 1259–60.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
34. Botstein D: **Why yeast?** *Hosp Pract (Off Ed)*. 1991; **26**(10): 157–61, 164.
[PubMed Abstract](#)
35. Cerami E, Gao J, Dogrusoz U, *et al.*: **The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data.** *Cancer Discov*. 2012; **2**(5): 401–404.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
36. Karathia H, Vilaprinyo E, Sorribas A, *et al.*: ***Saccharomyces cerevisiae* as a model organism: a comparative study.** *PLoS One*. 2011; **6**(2): e16015.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
37. Galao RP, Scheller N, Alves-Rodrigues I, *et al.*: ***Saccharomyces cerevisiae*: a versatile eukaryotic system in virology.** *Microb Cell Fact*. 2007; **6**: 32.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
38. Bharucha N, Kumar A: **Yeast genomics and drug target identification.** *Comb Chem High Throughput Screen*. 2007; **10**(8): 618–34.
[PubMed Abstract](#) | [Publisher Full Text](#)
39. Alamgir M, Erukova V, Jessulat M, *et al.*: **Chemical-genetic profile analysis of five inhibitory compounds in yeast.** *BMC Chem Biol*. 2010; **10**: 6.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
40. Giaever G, Nislow C: **The yeast deletion collection: a decade of functional genomics.** *Genetics*. 2014; **197**(2): 451–65.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
41. Lum PY, Armour CD, Stepaniants SB, *et al.*: **Discovering modes of action for therapeutic compounds using a genome-wide screen of yeast heterozygotes.** *Cell*. 2004; **116**(1): 121–37.
[PubMed Abstract](#) | [Publisher Full Text](#)
42. Marini NJ, Thomas PD, Rine J: **The use of orthologous sequences to predict the impact of amino acid substitutions on protein function.** *PLoS Genet*. 2010; **6**(5): e1000968.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
43. Hamza A, Tammperre E, Kofoed M: **Complementation of Yeast Genes with Human Genes as an Experimental Platform for Functional Testing of Human Genetic Variants.** *Genetics*. 2015; **201**(3): 1263–1274.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Open Peer Review

Current Peer Review Status: ? X X

Version 1

Reviewer Report 10 July 2018

<https://doi.org/10.5256/f1000research.16523.r35343>

© 2018 Reid R. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Robert J. D. Reid

Department Genetics and Development, Columbia University Medical Center, New York, NY, USA

Gupta *et al.* report a web-based application, Onco-Yeast, to identify human homologs of yeast genes and link to information from The Cancer Genome Atlas through cBio portal. While the concept of the software is reasonable, there are serious limitations to the approach. The core function of the web app is to provide links to information in cBio portal based on a homologous yeast gene. There are significant issues with how the homology with human genes is determined and it is unclear whether this direct linking to cBio portal is an advantage compared to initiating a query directly in the cBio portal interface.

The main problem with the approach taken in Oncoyeast is that gene orthology between organisms as diverged as yeast and humans is difficult and there does not seem to be a systematic approach to this core function of the software. There is not a single way to accomplish the task of identifying orthologs as evidenced by the multiple algorithms and databases available, such as InParanoid, Orthofinder, Ensemble compara, etc. There are distinctions between orthologs (same gene in a different organism), paralogs (duplicated gene either within or between organisms), and functional orthologs (different gene that plays an orthologous role in another organism). It is not at all clear how these issues are addressed by Oncoyeast. My understanding from the manuscript is that the homologs are taken from a static list provided by the *Saccharomyces Genome Database*. If so, then the details of how that list is derived and curated needs to be stated. On the other hand, based on the comments section of F1000, it seems as though homology is now acquired from ENSEMBL. Details of this approach would need to be similarly defined and cited in the manuscript. In any case, there are some strange calls of orthologs. For instance, Oncoyeast shows that the yeast genes *RAS1* and *RAS2* have the Human GTPase *REM2* as the ortholog. How was this call made? I doubt that *REM2* is a better choice of ortholog than *H-RAS*, *K-RAS*, *N-RAS*, *RRAS* or *RRAS2* - especially considering that *H-RAS* is capable of complementing the inviability of a *ras1 ras2* double mutant yeast ¹. Nevertheless, there are already good web resources for identifying orthologs. Queries of yeast genes in the Alliance of Genome Resources web site (Stanford) compares multiple orthology methods such as PANTHER, InParanoid, etc and ranks orthologs in multiple species based on an aggregate of the methods. A useful extension of this approach might include functional orthologs as well as evidence of

complementation between human and yeast genes ^{2,3} as part of the scoring metric.

The second issue with Oncoyeast is whether the links it generates to cBIO portal are more useful than querying the cBIO portal website directly. Part of this is a design issue, the row and column labels for the the genes and cancer studies scroll off the page and the site becomes difficult to use. More importantly, I am not sure that a direct link to the cBIO portal oncolinks page is the best approach. For instance, the yeast *MMS2* gene is correctly identified as human *UBE2V2* and following the link for breast invasive carcinoma calls up the cBIO portal oncoprint showing the modifications in the different breast carcinoma studies. What this does not indicate is that there is significant sample overlap in these studies and the resulting value for percent altered does not automatically remove duplicated samples possibly leading to erroneous conclusions. Making the same query in cBIO portal warns the user of overlapping studies up front. This does not happen in Oncoyeast.

In summary, since the core function of the software is identifying orthologs, and it does not seem to do this in a systematic way or offer an improvement over existing methods, I cannot approve this manuscript.

References

1. Kataoka T, Powers S, Cameron S, Fasano O, et al.: Functional homology of mammalian and yeast RAS genes. *Cell*. 1985; **40** (1): 19-26 [PubMed Abstract](#)
2. Hamza A, Tammperre E, Kofoed M, Keong C, et al.: Complementation of Yeast Genes with Human Genes as an Experimental Platform for Functional Testing of Human Genetic Variants. *Genetics*. 2015; **201** (3): 1263-74 [PubMed Abstract](#) | [Publisher Full Text](#)
3. Kachroo AH, Laurent JM, Yellman CM, Meyer AG, et al.: Evolution. Systematic humanization of yeast genes reveals conserved functions and genetic modularity. *Science*. 2015; **348** (6237): 921-5 [PubMed Abstract](#) | [Publisher Full Text](#)

Is the rationale for developing the new software tool clearly explained?

Yes

Is the description of the software tool technically sound?

Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?

No

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?

No

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?

No

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Genetics, Molecular Biology, Yeast, DNA repair

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Reviewer Report 28 June 2018

<https://doi.org/10.5256/f1000research.16523.r35158>

© 2018 Galli A. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Alvaro Galli

Laboratory of Functional Genetics and Genomics, Institute of Clinical Physiology, National Research Council (CNR), Pisa, Italy

The web-based application, described in the manuscript, could be relevant in terms of applicability in cancer therapy because has the potentiality to find new targets or new factors affecting drug response.

The starting point is homology between yeast and human. The authors only showed the RAD9 as example and reported that the human counterpart is the TP53BP1 gene. Where did they find this information? To my knowledge, yeast RAD9 has no human homolog; I have checked in the yeast genome database (<https://www.yeastgenome.org/>) and I could not find any homology. TP53BP1 appears to be RAD9 homolog only if you find homology with no filtered application (PANTHER method only giving this result, 1 out of 10) I guess that the authors have to state how and where the human homologous to the yeast gene are found and show a little more examples.

Moreover, the application only shows data from cBioPortal with no statistical evaluation on how it would be relevant to translate study from yeast to human. What I mean is that there is no functional evaluation of the gene, no data about the level of conservation between human and yeast. What I think is that a high level of conservation should correspond a high degree of reliability.

Recently, Mercatanti *et al* (FEMS Yeast Research, Dec. 2017)¹ published a web tool where yeast strains could be "humanized" and evaluated the reliability score to make the functional assay more relevant for cancer risk evaluation.

I do not think that the manuscript is not acceptable in the present form.

References

1. Mercatanti A, Lodovichi S, Cervelli T, Galli A: CRIMEtoYHU: a new web tool to develop yeast-based functional assays for characterizing cancer-associated missense variants. *FEMS Yeast Res.* 2017; **17** (8). [PubMed Abstract](#) | [Publisher Full Text](#)

Is the rationale for developing the new software tool clearly explained?

No

Is the description of the software tool technically sound?

No

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?

No

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?

No

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?

No

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Reviewer Report 19 June 2018

<https://doi.org/10.5256/f1000research.16523.r35163>

© 2018 Najafov A. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ayaz Najafov

Department of Cell Biology, Harvard Medical School , Boston, MA, USA

The authors have developed a web tool called Oncoyeasti, which bridges the gap between yeast and human ortholog gene symbols, and allows for easier submission of the queried genes to the TCGA/cBioportal database search. While this software can be useful, the following issues have to be addressed:

Major points:

1. Oncoyeasti should be able to retrieve the full list of human homologs that match to the yeast orthologs. For instance, if Sch9 is an ortholog of the Akt family, the output should be Akt1, Akt2, Akt3, rather than Akt3 only.

2. Oncoyeasti does not seem to retrieve some human orthologs accurately. For example, when TOR1 and TOR2 are queried, TOR1 matches to SMG1, while TOR2 matches to mTOR. In reality, both yeast TOR1 and yeast TOR2 should match to human mTOR.

Additionally, PKC1, YPK1, and YPK2 all match to Akt3, but they should match to PRKCE and SGK2 instead.

Thus, the Oncoyeasti database needs to be improved for accurate yeast-to-human gene symbol matching. The NCBI's Homologene database can be employed for this purpose.

3. If a certain yeast gene does not have a human ortholog, this should be reported in the output of Oncoyeasti (instead of only reporting those that have a human ortholog and skipping those that don't have a human ortholog).

Minor points:

1. The article would benefit from grammatical and typographical improvement. This is especially important where such mistakes may lead to misunderstanding of the authors' message by the readers. Some examples:
 - a) "the backbone of cancer, treatment" should be "the backbone of cancer treatment".
 - b) "BRAFFV600E" should be "BRAF-V600E" or "BRAFF^{V600E}".
 - c) "BRAFFV600E amplification" should be "BRAF-V600E mutation", since the V600E mutation is not an amplification.
 - d) "containing only 10% of the DNA of human chromosomes" should be rephrased.
 - e) "humologs" should be "homologs".
 - f) "MSQL" should be either "MySQL" or "mSQL".
 - g) "feature rich" should be "feature-rich".
 - h) "cross cancer" should be "cross-cancer".
2. The introduction part of the manuscript should be as concise and as relevant as possible. Since Oncoyeasti does not address the cancer heterogeneity directly, that part of the introduction seems unimportant for the manuscript.
3. The presentation of the output can be improved. For instance, the following sentence occupies a large area in the output, but the actual link-containing columns are squeezed to the right end of the page: "THE NEXT COLUMNS CONTAINS ONCOPRINT OF TUMOR GROUPS PRESENT IN CROSS CANCER SUMMARY STUDIES OF THE PATIENT TUMOR SETS PRESENT IN TCGA DATABASE".

Is the rationale for developing the new software tool clearly explained?

Yes

Is the description of the software tool technically sound?

Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?

Yes

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?

Yes

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?

No

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 02 Jul 2018

Ashish Patil, Providence Hospital, Washington, USA

The following were the major points Major points which had raised:

1) Oncoyeasti should be able to retrieve the full list of human homologs that match to the yeast orthologs. For instance, if Sch9 is an ortholog of the Akt family, the output should be Akt1, Akt2, Akt3, rather than Akt3 only.

2) Oncoyeasti does not seem to retrieve some human orthologs accurately. For example, when TOR1 and TOR2 are queried, TOR1 matches to SMG1, while TOR2 matches to mTOR. In reality, both yeast TOR1 and yeast TOR2 should match to human mTOR.

Additionally, PKC1, YPK1, and YPK2 all match to Akt3, but they should match to PRKCE and SGK2 instead.

Thus, the Oncoyeasti database needs to be improved for accurate yeast-to-human gene symbol matching. The NCBI's Homologene database can be employed for this purpose.

Response: Initially we had used a homolog gene list which was provided to us, by www.yeastgenome.org (Stanford database), but we did not know that this list was not complete and was missing significant number of human homologs. Thank you for bringing this to our notice.

Now we

used <http://useast.ensembl.org/biomart/martview/ba508c9593727fcc0111904d444c5296> and generated the list for human homologs for *S Cerevisiae*. This is the list we generated http://useast.ensembl.org/biomart/martresults/148?file=martquery_0621224014_815.xls.gz

Then we compared ensemble list with our stanford database list, to do corrections and also to add to our database, the additional *s cerevisiae* genes and their human homologs, which were missing in our list, that we got from yeast genome.org.

Using ensemble biomaRT, we have added 4157 additional human homologs to our existing list which we got from yeastgenome.org and thus making oncoyeasti much more comprehensive.

Competing Interests: None

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research