

RESEARCH ARTICLE

YARG: A repository for arsenic-related genes in yeast

Jagat Rathod¹, Hao-Ping Tu², Yung-I Chang², Yu-Han Chu², Yan-Yuan Tseng³, Jiin-Shuh Jean¹, Wei-Sheng Wu^{2*}

1 Department of Earth Sciences, National Cheng Kung University, Tainan, Taiwan, **2** Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan, **3** Center for Molecular Medicine and Genetics, Wayne State University School of Medicine, Detroit, MI, United States of America

* wessonwu@gmail.com



OPEN ACCESS

Citation: Rathod J, Tu H-P, Chang Y-I, Chu Y-H, Tseng Y-Y, Jean J-S, et al. (2018) YARG: A repository for arsenic-related genes in yeast. PLoS ONE 13(7): e0201204. <https://doi.org/10.1371/journal.pone.0201204>

Editor: Alvaro Galli, CNR, ITALY

Received: March 31, 2018

Accepted: July 10, 2018

Published: July 26, 2018

Copyright: © 2018 Rathod et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper.

Funding: This work was supported by National Cheng Kung University and Ministry of Science and Technology of Taiwan [MOST-105-2221-E-006-203-MY2 and MOST-106-2628-E-006-006-MY2]. Funding for open access charge: Ministry of Science and Technology of Taiwan. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

Abstract

Arsenic is a toxic metalloid. Moderate levels of arsenic exposure from drinking water can cause various human health problems such as skin lesions, circulatory disorders and cancers. Thus, arsenic toxicity is a key focus area for environmental and toxicological investigations. Many arsenic-related genes in yeast have been identified by experimental strategies such as phenotypic screening and transcriptional profiling. These identified arsenic-related genes are valuable information for studying arsenic toxicity. However, the literature about these identified arsenic-related genes is widely dispersed and cannot be easily acquired by researchers. This prompts us to develop YARG (Yeast Arsenic-Related Genes) database, which comprehensively collects 3396 arsenic-related genes in the literature. For each arsenic-related gene, the number and types of experimental evidence (phenotypic screening and/or transcriptional profiling) are provided. Users can use both search and browse modes to query arsenic-related genes in YARG. We used two case studies to show that YARG can return biologically meaningful arsenic-related information for the query gene(s). We believe that YARG is a useful resource for arsenic toxicity research. YARG is available at <http://cosbi4.ee.ncku.edu.tw/YARG/>.

Introduction

Arsenic (As), the 20th most abundant element on earth, is a toxic metalloid. In nature, arsenic is found in two chemical forms: inorganic species [arsenite (As³⁺) and arsenate (As⁵⁺)] and organic species [monomethylarsonic acid (MMA) and dimethylarsinic acid (DMA)]. Inorganic forms of arsenic are more toxic than organic forms. It is well reported that arsenic affects almost all cellular processes and organ functions that manifest due to cellular stress, mitochondrial and oxidative damage, genetic mutations and epigenetic dysregulation [1]. Low to moderate levels of arsenic exposure (10–300 µg L⁻¹) from drinking water can cause health problems such as skin lesions, circulatory disorders, neurological complications, diabetes, respiratory complications, hepatic and renal dysfunction [2]. Thus, arsenic toxicity is a major concern for environmental and toxicological investigations worldwide.

The availability of complete genome sequence and mutant libraries of *Saccharomyces cerevisiae* provides multi-directional opportunities to design insightful experiments to understand

arsenic metabolism, detoxification and tolerance acquisition mechanisms, which are mostly conserved from yeast to human [3–5]. Majority of yeast genes involved in arsenic response have homolog in humans that could potentially modulate toxicity in a similar manner as their yeast counterparts. Therefore, the budding yeast *S. cerevisiae* is a useful eukaryotic model organism for studying arsenic toxicity in human health and diseases [6].

To understand how yeast cells respond to arsenic exposure, it is crucial to identify arsenic-related genes. Two main experimental strategies are used to identify arsenic-related genes systematically. The first is phenotypic screening of the *S. cerevisiae* mutant libraries under arsenic exposure [7–14]. Comparing growth kinetics of wild-type (WT) and a homozygous deletion mutant of a specific gene, both with and without arsenic exposure, identifies the potential role of that gene in arsenic sensitivity or resistance if it significantly affects the growth rate. The phenotypic screening is done for all ~4870 non-essential genes in the yeast genome [14]. Although most of the phenotypic screens are performed in nutrient rich media, different studies used different strains and exposed cells to various arsenic concentrations for different durations. Therefore, diverse phenotypic screenings have identified different sets of arsenic-related genes [7–14]. Phenotypic screenings supported with additional physiological and biochemical characterizations have highlighted the genetic determinants of arsenic susceptibility and resistance along with their designated functions in yeast. On a larger scale, it is known that arsenic (specifically arsenite) binds to α -helices affecting secondary structures of proteins, which could be corroborated by phenotypic screens suggesting functional inhibition of the chaperonin complex [13]. These details are critical to understand the arsenic molecular response mechanism in yeast.

The second strategy is genome-wide transcriptional profiling [7, 9, 15]. By comparing the genome-wide gene expression patterns between WT cells with and without arsenic exposure, the genes whose expressions are significantly affected by arsenic can be identified. By knowing differentially expressed genes under arsenic exposure, multiple inferences could be derived on the functioning of genetic networks, their response directions and diverse pathways involved. Systematically, yeast has evolved various defence strategies to tolerate and detoxify arsenic by reduction of metal uptake, enhanced extrusion, sequestration within vacuoles and chelation by metal-binding proteins and polypeptides [9, 15, 16]. Considering the need for a systematic and prompt response in yeast, signal transduction pathways undergo a rapid re-programming of cellular transcriptome that eventually modulates proteome and metabolome profiles. The functionality of differentially expressed genes in the defense against the arsenic-toxicity could identify a novel responsive candidate that might be critical for arsenic resistance and cellular processing. Thus, it is essential to compile and compare all global transcriptome profiling data obtained under arsenic exposure.

Further, experimental evidence shows that arsenic can directly bind to and activate transcription factor Arr1 [17]. Other transcription factors such as Yap1 (a key regulator of oxidative stress response) and Rpn4 (a key regulator of proteotoxic stress response) also play a critical role in arsenic detoxification. It is known that transcription factors Arr1, Rpn4 and Yap1 confer resistance to arsenic via regulating the expression levels of diverse genes [7]. Therefore, arsenic-related genes can also be identified by comparing the genome-wide gene expression differences between WT and transcription factor mutants (*arr1Δ*, *rpn4Δ* or *yap1Δ*) both under arsenic exposure [7, 15]. Transcriptional profiling also helps to understand the role of Arr1, Rpn4 and Yap1 in arsenic resistance and their effects on genome-wide gene expression patterns.

Additionally, Haugen *et al.* [7] utilized both phenotypic screening and gene expression analysis to draw conclusion that arsenic might channel sulfur into glutathione for detoxification, lead to indirect oxidative stress by depleting glutathione pools, and alter protein turnover

via arsenation of sulfhydryl groups on proteins. They also highlighted that phenotypically sensitive pathways are upstream of differentially expressed genes, suggesting that transcriptional and phenotypic profiling implicate distinct but functionally related pathways in the yeast system. Therefore, to comprehensively characterize the arsenic-related genes, we need to integrate both phenotypic screening and transcriptional profiling datasets from all available studies in the literature.

Thousands of arsenic-related genes in yeast have been identified by many experimental studies using either phenotypic screening or transcriptional profiling, or both [7–15]. However, the literature about these identified arsenic-related genes is widely dispersed and cannot be easily acquired by researchers. Therefore, there is a need of a database which comprehensively collects the arsenic-related genes from the literature and provide a dynamic interface to all information. To meet this need, we have constructed YARG (Yeast Arsenic-Related Genes) database, which collects 3396 arsenic-related genes and their experimental evidence. Users can search YARG by gene names and get the information about their arsenic-related correlation with collective evidence from phenotypic screening and/or transcriptional profiling. Besides, users can browse YARG to retrieve 20 different lists of arsenic-related genes from nine experimental studies [7–15]. The experimental strategy, experimental strain and experimental condition of these studies are also provided. In summary, YARG is a useful resource for the scientific community to investigate arsenic toxicity in yeast.

Construction and contents

The configuration of YARG database

Fig 1 illustrates the configuration of YARG database. Python with Django MTV framework was used to construct YARG website. Python was also used for raw data processing. The processed data was stored in MySQL. The tables were produced by Data Tables (a table plug-in for jQuery). The graphics were generated by vis.js (a browser based graphic drawing library).

Collection of 3396 arsenic-related genes

We collected 20 gene lists in nine existing studies [7–15] which experimentally identified arsenic-related genes by phenotypic screening (PS) or transcriptional profiling (TP). Among the 20 collected gene lists, 13 were generated by PS (Table 1) and 7 were generated by TP (Table 2). We then retrieved 3396 arsenic-related genes from these 20 collected gene lists. Among the 3396 arsenic-related genes, 535 are supported by both PS and TP, 737 are supported only by PS, and 2124 are supported only by TP (Fig 2A). The distribution of these 3396 arsenic-related genes on different chromosomes is shown in Fig 2B.

Testing the enrichment of arsenic-related genes in input genes

For users' input genes, YARG tests whether they are enriched with arsenic-related genes. The *p-value* is calculated using hypergeometric test [18] as follows

$$p_value = \sum_{x \geq K}^{\min(A, I)} \frac{\binom{A}{x} \binom{G-A}{I-x}}{\binom{G}{I}}$$

where $G = 6572$ is the number of genes in yeast genome, $A = 3396$ is the number of arsenic-

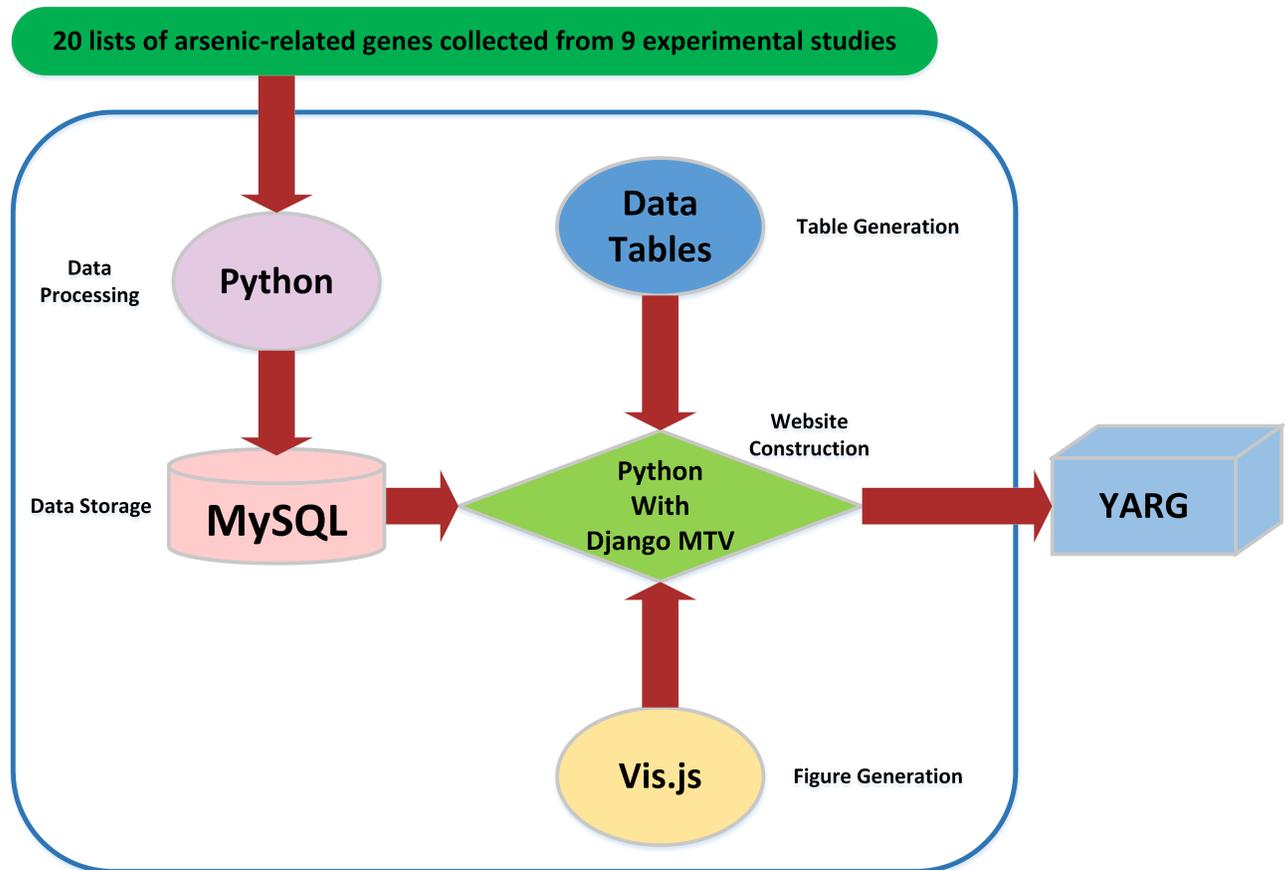


Fig 1. The configuration of YARG database.

<https://doi.org/10.1371/journal.pone.0201204.g001>

related genes in YARG, I is the number of users' input genes, and K is the number of input genes which are also arsenic-related genes.

Utility and discussion

Database interface

YARG provides two search modes. In the first search mode, users can input a gene name (Fig 3A). After submission, YARG returns a page showing the basic information of the input gene and links to YeastMine [19] to see the homology information such as human homologs, fungal homologs, non-fungal homologs, functional complementation and the paralogs (Fig 3B). If the input gene is an arsenic-related gene, details (experimental strain, experimental condition and reference) of the experimental evidence (phenotypic screening and/or transcriptional profiling) are provided (Fig 3C). In the second search mode, users can input a list of genes (Fig 4A). After submission, YARG uses the hypergeometric test [18] to analyze whether the input genes are enriched with arsenic-related genes (Fig 4B). YARG also provides a figure and a table to show which input genes are arsenic-related genes and total number of supporting evidences (Fig 4C). Details (experimental strain, experimental condition and reference) of the supporting evidence are also shown in Fig 4D.

YARG provides three browse modes (Fig 5A). In the first browse mode, users can browse 3396 arsenic-related genes. For each gene, YARG provides a systematic name, a standard

Table 1. Total thirteen collected lists of arsenic-related genes identified by phenotypic screening.

Source	Identified gene list	Experimental strain	Experimental growth condition (arsenic exposure)
Haugen <i>et al.</i> 2004 [7]	213 gene mutants of arsenite-sensitive phenotypes	Homozygous diploid <i>S. cerevisiae</i> BY4741 (<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0</i>) mutants	100 μM and 1 mM sodium arsenite for 0.5, 2 or 4 h
Vujcic <i>et al.</i> 2007 [8]	72 gene mutants of arsenite-sensitive phenotypes	<i>S. cerevisiae</i> BY4741 (<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0</i>) mutants	YPD plates containing sodium arsenite (2.5 or 5 mmol/L) as well as on YPD plates without any arsenic (control plates). Plates were incubated for 4 to 15 days at 30°C and phenotype of each mutant was scored as sensitive or resistant compared with control plate and internal control (BY4741 on each plate).
Jin <i>et al.</i> 2008 [9]	65 gene mutants of arsenite-sensitive phenotypes	Homozygous diploid <i>S. cerevisiae</i> BY4743 (<i>MATa/α his3Δ1/his3Δ1 leu2Δ0 /leu2Δ0 lys2Δ0/LYS2 MET15/ met15Δ0 ura3Δ0 /ura3Δ0</i>) mutants	1.25 mM sodium arsenite for 2 h
Jo <i>et al.</i> 2009 [10]	647 gene mutants of arsenite-sensitive phenotypes	Homozygous diploid <i>S. cerevisiae</i> BY4743 (<i>MATa/α his3Δ1/his3Δ1 leu2Δ0 /leu2Δ0 lys2Δ0/LYS2 MET15/ met15Δ0 ura3Δ0 /ura3Δ0</i>) mutants	75, 150 and 300 μM sodium arsenite for 5 and 15 generations and subsequently analysed by TAG4 arrays.
Thorsen <i>et al.</i> 2009 [11]	305 gene mutants of arsenite-sensitive phenotypes	Haploid <i>S. cerevisiae</i> BY4741 and homozygous diploid <i>S. cerevisiae</i> BY4743 mutants	0.5, 1.0 and 1.5 mM sodium arsenite for 24, 48 and 72 h
Zhou <i>et al.</i> 2009 [12]	245 gene mutants of arsenite-sensitive phenotypes	Haploid <i>S. cerevisiae</i> BY4741 mutants	0, 0.75 and 1 mM sodium arsenite for 60 h
Pan <i>et al.</i> 2010 [13]	191 gene mutants of arsenite-sensitive phenotypes	Heterozygous diploid <i>S. cerevisiae</i> BY4741 mutants	1 mM sodium arsenite for 1 h
Pan <i>et al.</i> 2010 [13]	33 gene mutants of arsenite-sensitive phenotypes	Heterozygous diploid <i>S. cerevisiae</i> BY4741 mutants	450 mM sodium arsenite for 10 generations
Johnson <i>et al.</i> 2016 [14]	75 gene mutants of arsenite-sensitive phenotypes	Homozygous diploid <i>S. cerevisiae</i> BY4743 (<i>MATa/α his3Δ1/his3Δ1 leu2Δ0 /leu2Δ0 lys2Δ0/LYS2 MET15/ met15Δ0 ura3Δ0 /ura3Δ0</i>) mutants	0.2 and 0.4 mM sodium arsenite for 16 and 20 h
Zhou <i>et al.</i> 2009 [12]	5 gene mutants of arsenite-resistant phenotypes	Haploid <i>S. cerevisiae</i> BY4741 mutants	0, 0.75 and 1mM sodium arsenite for 60 h
Pan <i>et al.</i> 2010 [13]	109 gene mutants of arsenite-resistant phenotypes	Heterozygous diploid <i>S. cerevisiae</i> BY4741 mutants	1 mM sodium arsenite for 1 h
Johnson <i>et al.</i> 2016 [14]	39 gene mutants of arsenite-resistant phenotypes	Homozygous diploid <i>S. cerevisiae</i> BY4743 (<i>MATa/α his3Δ1/his3Δ1 leu2Δ0 /leu2Δ0 lys2Δ0/LYS2 MET15/ met15Δ0 ura3Δ0 /ura3Δ0</i>) mutants	0.2 and 0.4 mM sodium arsenite for 16 and 20 h
Vujcic <i>et al.</i> 2007 [8]	81 gene mutants of arsenate-sensitive phenotypes	<i>S. cerevisiae</i> BY4741 and haploid <i>MATa</i> deletion mutant derived from parental strain BY4741	YPD plates containing sodium arsenate (15 or 30 mmol/L) as well as on YPD plates without any arsenic (control plates). Plates were incubated for 4 to 15 days at 30°C and phenotype of each mutant was scored as sensitive or resistant compared with control plate and internal control (BY4741 on each plate).

These 13 arsenic-related gene lists consist of 9 mutant gene lists of arsenite-sensitive phenotypes, 3 mutant gene lists of arsenite-resistant phenotypes and 1 mutant gene list of arsenate-sensitive phenotypes.

<https://doi.org/10.1371/journal.pone.0201204.t001>

name, name description, genomic location, total number of arsenic-related evidence from PS and TP (Fig 5B). In the second browse mode, users can browse 13 arsenic-related gene lists generated by phenotypic screening. These 13 arsenic-related gene lists consist of 1 mutant gene list of arsenate-sensitive phenotypes, 3 mutant gene lists of arsenite-resistant phenotypes and 9 mutant gene lists of arsenite-sensitive phenotypes (Fig 5C). In the third browse mode, users can browse 7 arsenic-related gene lists generated by transcriptional profiling. These 7 arsenic-related gene lists consist of (i) 3 lists of genes which are differentially expressed

Table 2. Total seven collected lists of arsenic-related genes identified by transcriptional profiling.

Source	# of identified differentially expressed genes ^a	Differentially expressed when comparing	Experimental strain	Experimental growth condition (arsenic exposure)
Thorsen <i>et al.</i> 2007 [15]	756	WT 0.2 mM As(III) vs. WT	<i>S. cerevisiae</i> W303-1A	0.2 mM sodium arsenite for 1 h
Thorsen <i>et al.</i> 2007 [15]	1066	WT 1.0 mM As(III) vs. WT	<i>S. cerevisiae</i> W303-1A	1.0 mM sodium arsenite for 1 h
Jin <i>et al.</i> 2008 [9]	1194	WT 0.4 mM As(III) vs. WT	<i>S. cerevisiae</i> BY4742 (<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i>)	0.4 mM sodium arsenite for 2 h
Thorsen <i>et al.</i> 2007 [15]	76	<i>yap1Δ</i> 1.0 mM As(III) vs. WT 1.0 mM As(III)	<i>S. cerevisiae</i> RW124-W303-1A <i>yap1Δ::loxP</i>	1.0 mM sodium arsenite for 1 h
Haugen <i>et al.</i> 2004 [7]	179	<i>yap1Δ</i> 100 μM As(III) vs. WT 100 μM As(III)	<i>yap1Δ</i> of <i>S. cerevisiae</i> BY4741 (<i>MATα, his3Δ, leu2Δ0, met15Δ0, uraΔ0</i>)	100 μM sodium arsenite for 2 h
Haugen <i>et al.</i> 2004 [7]	415	<i>rpn4Δ</i> 100 μM As(III) vs. WT 100 μM As(III)	<i>rpn4Δ</i> <i>S. cerevisiae</i> BY4741 (<i>MATα, his3Δ, leu2Δ0, met15Δ0, uraΔ0</i>)	100 μM sodium arsenite for 2 h
Haugen <i>et al.</i> 2004 [7]	875	<i>arr1Δ</i> 100 μM As(III) vs. WT 100 μM As(III)	<i>arr1Δ</i> of <i>S. cerevisiae</i> BY4741 (<i>MATα, his3Δ, leu2Δ0, met15Δ0, uraΔ0</i>)	100 μM sodium arsenite for 2 h

These 7 arsenic-related gene lists consist of (i) 3 lists of genes which are differentially expressed between WT and WT under arsenic exposure and (ii) 4 lists of genes which are differentially expressed between WT and the transcription factor mutants (*arr1Δ*, *rpn4Δ* or *yap1Δ*) both under arsenic exposure.

^aA gene was considered differentially expressed if the fold-change value was greater than or equal to twofold and if the *p*-value was less than 0.001 (in [7]) or 0.01 (in [9] and [15]).

<https://doi.org/10.1371/journal.pone.0201204.t002>

between WT and WT under arsenic exposure and (ii) 4 lists of genes which are differentially expressed between WT and transcription factor mutants (*arr1Δ*, *rpn4Δ* or *yap1Δ*) both under arsenic exposure (Fig 5D).

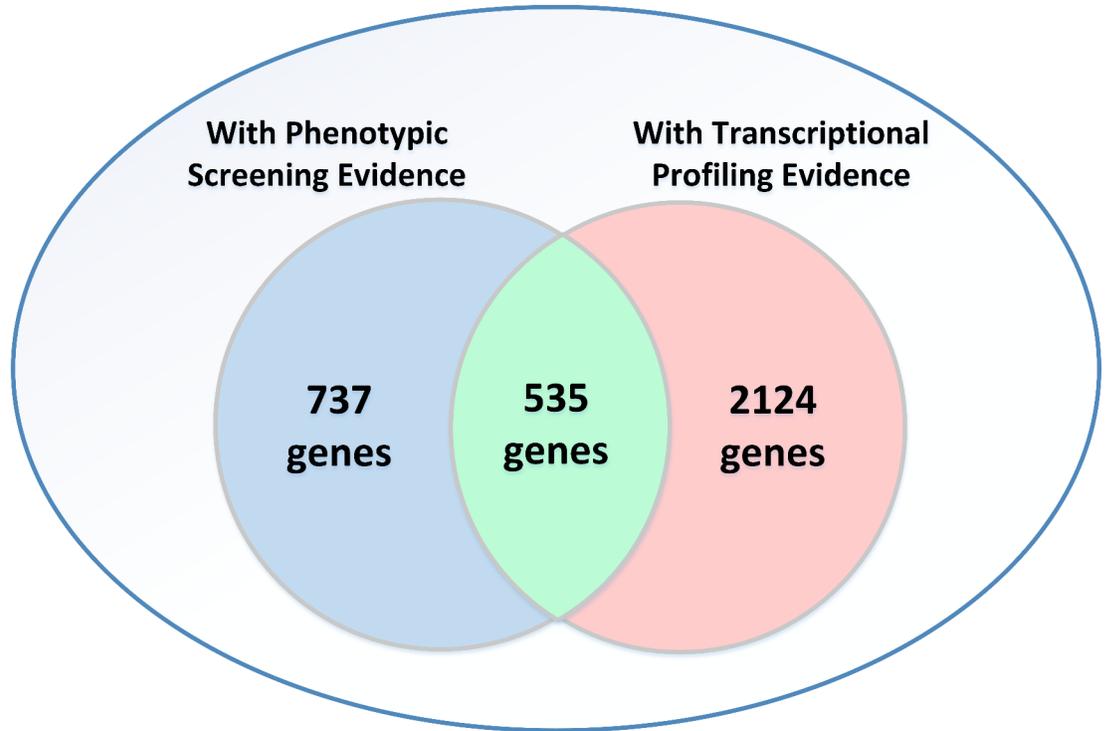
Two case studies

Here we give two case studies to show that the search modes of YARG can return biologically meaningful arsenic-related information for the users' query gene(s). The first case illustrates a scenario of a single gene name submission. Yap1 is a transcription activator known to be involved in arsenic adaptation process via regulation of expression of ACR (arsenic compounds resistance) genes [17, 20, 21]. When we input a single gene name *YAPI* (Fig 3A), YARG successfully identified *YAPI* as an arsenic-related gene and provided seven arsenic-related existing experimental evidences (Fig 3C). Specifically, five phenotypic screening studies [7, 10, 11, 13, 14] utilized different experimental yeast host strains that collectively elucidate that *YAPI* is an arsenic-sensitive gene, signifying that *yap1Δ* mutant has decreased fitness and transforms host strains into arsenic-sensitive phenotype. For example, Haugen *et al.* [7] identified deletion mutants with increased sensitivity to growth inhibition utilizing an available deletion mutant library of nonessential genes (4,650 homozygous diploid strains) pinpointing that *yap1Δ* is present in the first 50 arsenic-sensitive deletion strains. Moreover, Thorsen *et al.*'s transcriptional profiling experiment [15] showed that *YAPI* is differentially expressed between wild-type strains with and without arsenic exposure. Haugen *et al.*'s transcriptional profiling experiment [7] identified 50 differentially expressed genes between wild-type and *yap1Δ* strain both under arsenic exposure. Strikingly, 20 of these 50 genes are known to play a critical role in protection against arsenic exposure, suggesting that the transcription factor Yap1 might strongly mediate arsenic-induced stress adaptation [7].

In addition to showing arsenic-related evidences of *YAPI*, YARG also provides links to YeastMine [19] for users to find out homology information of *YAPI*. For example, the human homolog(s) link to YeastMine (Fig 3B) reveals that arsenic-related gene *YAPI* has a human

(a)

3396 Arsenic-Related Genes



(b)

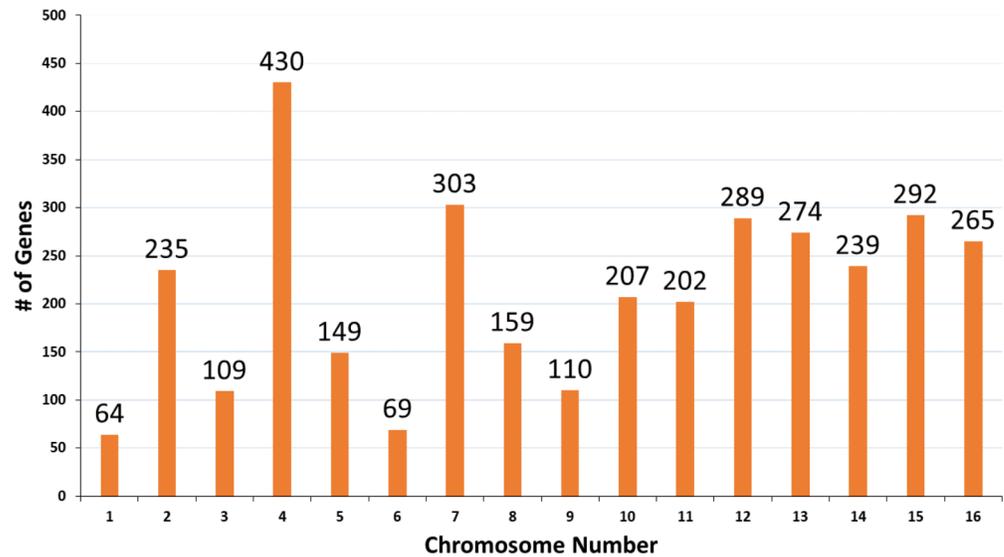


Fig 2. Total 3396 arsenic-related genes in YARG. (a) The 3396 arsenic-related genes are supported by phenotypic screening, transcriptional profiling, or both. (b) The distribution of the 3396 arsenic-related genes on 16 yeast chromosomes.

<https://doi.org/10.1371/journal.pone.0201204.g002>

homolog *TUSC1* which is known to play a role in various kinds of tumorigenesis [22–25], providing a possible explanation as to why arsenic is a potential carcinogen for various kinds of

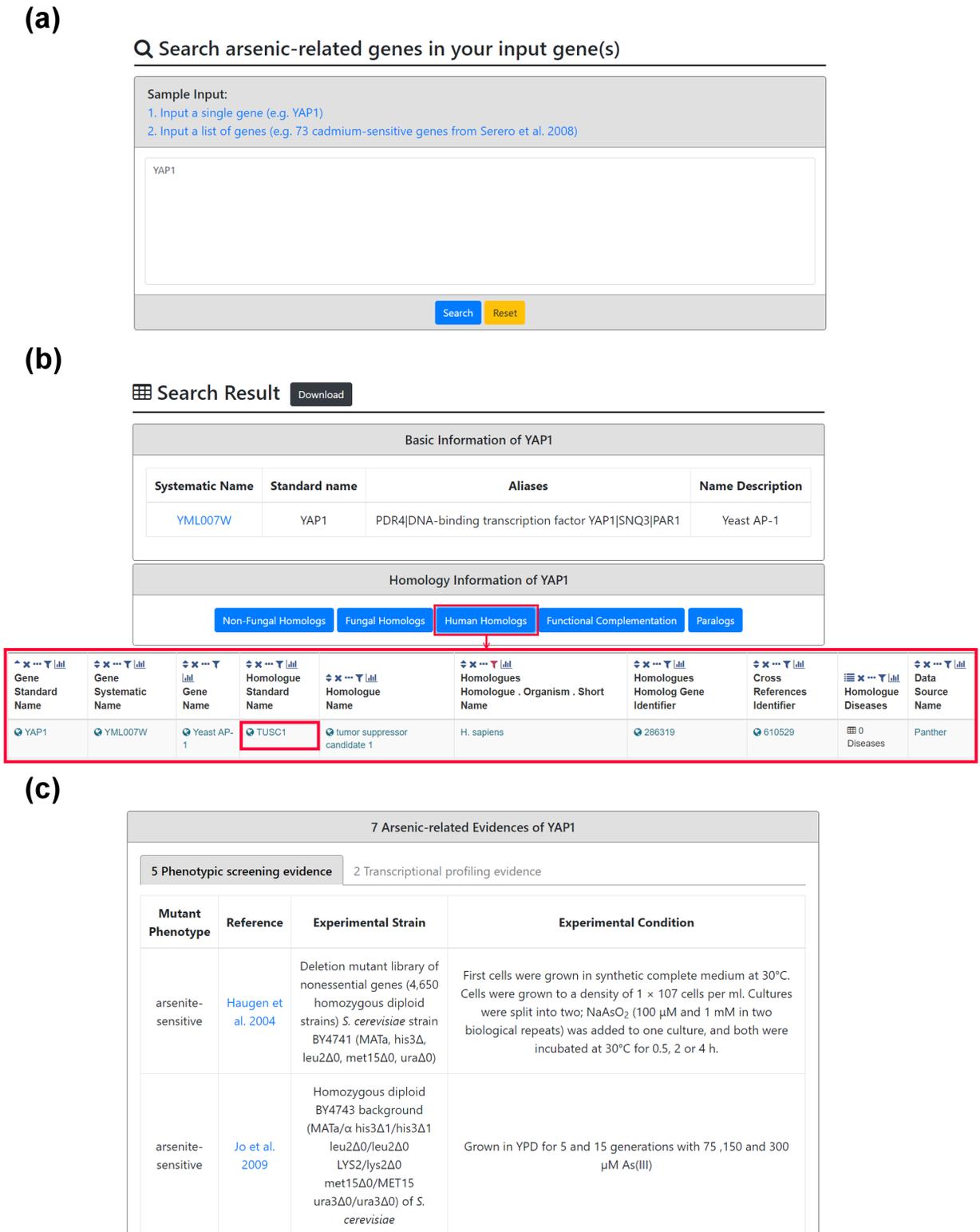


Fig 3. The first search mode. (a) Input a single gene name *YAP1*. (b) The basic information of *YAP1* and homology links to YeastMine. (c) Details of the arsenic-related evidence (phenotypic screening or transcriptional profiling) of *YAP1*.

<https://doi.org/10.1371/journal.pone.0201204.g003>

(a) **Q Search arsenic-related genes in your input gene(s)**

Sample Input:
 1. Input a single gene (e.g. YAP1)
 2. Input a list of genes (e.g. 73 cadmium-sensitive genes from Serero et al. 2008)

FZO1
 YDL020C
 YCF1
 REF2
 YGR155W
 YHR088C
 YIL101C
 RPS21B

Search Reset

(b) **Search Result** Download

Your 73 genes are enriched with arsenic-related genes.

The p-value is calculated by hypergeometric test.

$$P_{value} = \sum_{x=T}^{\min(S,G)} \frac{\binom{S}{x} \binom{P-S}{G-x}}{\binom{P}{G}} = 1.024e^{-8}$$

- F=6572 is the number of genes in the yeast genome.
- S=3396 is the number of arsenic-related genes in YARG.
- G=73 is the number of users' input genes.
- T=61 is the number of input genes which are also arsenic-related genes.

(c)

61 out of 73 input genes are arsenic-related genes

— Arsenic-related Evidence from Phenotypic Screening (PS)
 — Arsenic-related Evidence from Transcriptional Profiling (TP)
 ● Gene Encoding Transcription Factor

Show 10 entries Search:

Systematic Name	Standard Name	Name Description	Genomic Location	#of Arsenic-related Evidence from PS	#of Arsenic-related Evidence from TP
YAL021C	CCR4	Carbon Catabolite Repression	chr01:113359-110846	1	0
YBLOS8W	SHP1	Suppressor of High-copy PP1	chr02:111437-112708	0	1

(d)

CCR4 is an arsenic-related gene.

1 Phenotypic Screening Evidence

Mutant Phenotype	Reference	Experimental Strain	Experimental Condition
arsenite-sensitive	Thorsen et al. 2009	Haploid strain BY4741 (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0) and the homozygous diploid strain BY4743 (MATa/MAT his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 ura3Δ0/ura3Δ0 LYS2/lys2Δ0 met15Δ0/MET15) of <i>S. cerevisiae</i>	Sodium arsenite (0.5, 1.0 and 1.5 mM) for phenotypic mutant screening

SHP1 is an arsenic-related gene.

1 Transcriptional Profiling Evidence

Up/Down Regulation	Differentially Expressed When Comparing	Reference	Experimental Strain	Experimental Condition
Up	WT 1.0 mM As(III) vs. WT No arsenic	Thorsen et al. 2007	<i>S. cerevisiae</i> strain W303-1A	1.0 mM sodium arsenite 1 h

Fig 4. The second search mode. (a) Input a list of 73 cadmium-sensitive genes. (b) YARG tests whether the input genes are enriched with arsenic-related genes. (c) YARG provides a figure and a table to show which input genes are arsenic-related genes and the total number of supporting evidence. (d) Details of arsenic-related evidence (phenotypic screening or transcriptional profiling).

<https://doi.org/10.1371/journal.pone.0201204.g004>

cancers in human. Thus, YARG enables investigators to expose the various novel possible nexus between arsenic toxicity and human disease manifestations utilizing existing yeast toxicity studies. In summary, this case study successfully demonstrates that YARG can provide arsenic-related information and homology information for the user's queried gene.

The second case study illustrates a scenario of a gene list submission. It is known that cadmium induces unfolded protein response, endoplasmic reticulum (ER) and oxidative stress, and hampers energy metabolism in yeast [26]. Several experimental studies have shown that *S. cerevisiae* uses similar detoxification mechanisms against cadmium and arsenic as other higher eukaryotic systems [9, 15, 16]. It is also well documented that the genes required for cadmium resistance have significant overlap with the genes required for arsenic resistance [27, 28]. When we input a list of 73 cadmium-sensitive genes identified by phenotypic screening from Serero *et al.* [29], YARG successfully identified that these 73 cadmium-sensitive genes are enriched (p -value = 1.024E-8 calculated by hypergeometric test [18]) with arsenic-related genes (Fig 4B), which are consistent with the existing knowledge [27, 28]. Specifically, 61 cadmium-sensitive genes are also arsenic-related genes with experimental evidence of phenotypic screening or/and transcriptional profiling (Fig 4C), suggesting that *S. cerevisiae* may use similar detoxification mechanisms against cadmium and arsenic [9, 16]. This case study clearly demonstrates that YARG can support users to compare the gene lists related to different metals and toxins, which may help in identifying novel candidate genes for toxicological research.

Comparison with SGD and YeastMine

YARG collected 3396 arsenic-related genes supported by phenotypic screening or/and transcriptional profiling evidence from the literature. The advantages of YARG over SGD [30] and YeastMine [19] are as follows. First, SGD only allows users to check one gene at a time whether it is an arsenic-related gene according to the phenotype annotations. Second, YeastMine only allows users to retrieve a list of genes that are annotated to "metal resistance"; therefore, users still need to extract arsenic-related genes from this gene list. Third, both SGD and YeastMine define arsenic-related genes using only the phenotype annotations; neither of them provides the arsenic-related genes supported by transcriptional profiling. In summary, YARG is a useful resource of arsenic research since it provides arsenic-related genes supported by transcriptional profiling or/and phenotypic screening evidence from the literature.

Conclusions

In this study, we present YARG, a database which is a collection of 3396 arsenic-related genes from the literature. For each arsenic-related gene, the number and types of experimental evidence (phenotypic screening and/or transcriptional profiling) are provided. Users can use both search and browse modes to query arsenic-related genes in YARG. Two case studies (a single gene *YAP1* and a list of 73 cadmium-sensitive genes) have been provided to show that YARG can retrieve biologically meaningful arsenic-related information along with experimental evidence for the users' query gene(s). In future, we will keep updating YARG as and when new arsenic-related gene lists are available from newly published papers. YARG will be maintained regularly by our laboratory personnel. Therefore, the long-term stability of YARG is guaranteed. We also provide two backup sites (<http://cosbi5.ee.ncku.edu.tw/YARG/> and

(a) **3 Browse Modes**

Total 3396 Arsenic-Related Genes

[Browse](#)

13 Arsenic-Related Gene Lists Generated by Phenotypic Screening

[Browse](#)

7 Arsenic-Related Gene Lists Generated by Transcriptional Profiling

[Browse](#)

(b) **Total 3396 Arsenic-Related Genes** [Download](#)

Show 10 entries

Systematic Name	Standard Name	Name Description	Genomic Location	# of Arsenic-related Evidence from PS	# of Arsenic-related Evidence from TP
YAL002W	VPS8	Vacuolar Protein Sorting	chr01:143707-147531	2	0
YAL003W	EFB1	Elongation Factor Beta	chr01:142174-143160	0	1
YAL004W			chr01:140760-141407	0	1
YAL005C	SSA1	Stress-Seventy subfamily A	chr01:141431-139503	0	2

(c) **13 Arsenic-Related Gene Lists Generated by Phenotypic Screening** [Download](#)

1 mutant gene list of arsenate sensitive phenotypes

# of Genes	Experimental Strain	Experimental Condition	Reference
81	WT <i>S. cerevisiae</i> strain BY4741 (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0) and a collection of nonessential haploid MATa deletion strains derived from parental strain BY4741	Cultures were spotted onto YPD plates containing varying concentrations of sodium arsenate (5, 10, 15, and 25 mmol/L) and plates were incubated for 3 to 6 days at 30 °C. Growth of the WT BY4741 strain on the plates containing arsenic was compared with its growth on the control plate (without any arsenite or arsenate).	Vujcic et al. 2007 PubMed Journal

3 mutant gene lists of arsenite resistant phenotypes

# of Genes	Experimental Strain	Experimental Condition	Reference
39	Hemizygous diploid BY4743 background (MATa to his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 lys2Δ0/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0) and deletion mutant set of haploid strain BY4741	As (III) (0.2 and 0.4 mM) 16 h and 20 h growth	Johnson et al. 2016 PubMed Journal

9 mutant gene lists of arsenite sensitive phenotypes

# of Genes	Experimental Strain	Experimental Condition	Reference
72	WT <i>S. cerevisiae</i> strain BY4741 (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0) and a collection of nonessential haploid MATa deletion strains derived from parental strain BY4741	Cultures were spotted onto YPD plates containing varying concentrations of sodium arsenite (1.25, 2.5, 5.0, and 10 mmol/L) and plates were incubated for 3 to 6 days at 30 °C. Growth of the WT BY4741 strain on the plates containing arsenic was compared with its growth on the control plate (without any arsenite or arsenate).	Vujcic et al. 2007 PubMed Journal

(d) **7 Arsenic-Related Gene Lists Generated by Transcriptional Profiling** [Download](#)

7 lists of differentially expressed genes

# of Differentially Expressed Genes	Differentially Expressed When Comparing	Experimental Strain	Experimental Condition	Reference
76	yap1Δ 1.0 mM As(III) vs. WT 1.0 mM As(II)	<i>S. cerevisiae</i> strain RW124 ---- W303-1A yap1Δ:loxP	1.0 mM sodium arsenite 1 h	Thorsen et al. 2007 PubMed Journal
179	yap1Δ 100 μM As(III) vs. WT 100 μM As(III)	yap1Δ <i>S. cerevisiae</i> strain BY4741 (MATa, his3Δ, leu2Δ0, met15Δ0, uraΔ0)	100 μM sodium arsenite 2 h incubation at 30°C	Haugen et al. 2004 PubMed Journal

76 differentially expressed genes from Thorsen et al. 2007 [Download](#)

Differentially expressed when comparing yap1Δ 1.0 mM As(III) vs. WT 1.0 mM As(II)

Show 10 entries

Up/Down Regulation	Systematic Name	Standard Name	Name Description	Genomic Location	# of Arsenic-related Evidence	Detail
Down	YBL064C	PRX1	PeroxiRedoxin	chr02:101156-100371	5	Detail

5 Arsenic-related Evidences of PRX1

1 Phenotypic screening evidence 4 Transcriptional profiling evidence

Mutant Phenotype	Reference	Experimental Strain	Experimental Condition
arsenite-sensitive	Zhou et al. 2009	Haploid <i>S. cerevisiae</i> strain BY4741	0, 0.75, and 1 mM sodium arsenite. Inoculated plates were incubated for 60 h at 30 °C and the resulting growth was measured. 0, 0.75, and 1 mM sodium arsenite exposure. Plates were incubated for 60 h at 30 °C and the resulting growth was measured.

Fig 5. Three browse modes. (a) Three browse modes. (b) The first browse mode: users can browse 3396 arsenic-related genes in YARG. A table is given to show a systematic name, a standard name, name description, genomic location, the number of arsenic-related evidence from PS and TP. (c) The second browse mode: users can browse 13 arsenic-related gene lists generated by phenotypic screening (PS). (d) The third browse mode: users can browse 7 arsenic-related gene lists generated by transcriptional profiling (TP).

<https://doi.org/10.1371/journal.pone.0201204.g005>

<http://cosbi2.ee.ncku.edu.tw/YARG/>) just in case the main website is temporarily not available. We believe that YARG is a useful resource for arsenic toxicity research in yeast, supporting research community worldwide.

Acknowledgments

This work was supported by National Cheng Kung University and Ministry of Science and Technology of Taiwan [MOST-105-2221-E-006-203-MY2 and MOST-106-2628-E-006-006-MY2]. Funding for open access charge: Ministry of Science and Technology of Taiwan.

Author Contributions

Conceptualization: Jagat Rathod, Wei-Sheng Wu.

Data curation: Jagat Rathod, Hao-Ping Tu, Yung-I Chang, Wei-Sheng Wu.

Formal analysis: Yung-I Chang.

Funding acquisition: Wei-Sheng Wu.

Investigation: Jagat Rathod, Wei-Sheng Wu.

Methodology: Jagat Rathod, Wei-Sheng Wu.

Project administration: Wei-Sheng Wu.

Software: Hao-Ping Tu.

Supervision: Yan-Yuan Tseng, Jiin-Shuh Jean, Wei-Sheng Wu.

Validation: Jagat Rathod, Wei-Sheng Wu.

Visualization: Hao-Ping Tu, Yu-Han Chu.

Writing – original draft: Jagat Rathod, Wei-Sheng Wu.

Writing – review & editing: Jagat Rathod, Wei-Sheng Wu.

References

1. Ray PD, Yosim A, Fry RC. Incorporating epigenetic data into the risk assessment process for the toxic metals arsenic, cadmium, chromium, lead, and mercury: strategies and challenges. *Front Genet* 2014; 5:201. <https://doi.org/10.3389/fgene.2014.00201> PMID: 25076963
2. Chen Y, Parvez F, Gamble M, Islam T, Ahmed A, Argos M, et al. Arsenic exposure at low-to-moderate levels and skin lesions, arsenic metabolism, neurological functions, and biomarkers for respiratory and cardiovascular diseases: review of recent findings from the Health Effects of Arsenic Longitudinal Study (HEALS) in Bangladesh. *Toxicol Appl Pharmacol* 2009; 239(2):184–192. <https://doi.org/10.1016/j.taap.2009.01.010> PMID: 19371619
3. Dhankher OP, Li Y, Rosen BP, Shi J, Salt D, Senecoff JF, et al. Engineering tolerance and hyperaccumulation of arsenic in plants by combining arsenate reductase and gamma-glutamylcysteine synthetase expression. *Nat Biotechnol* 2002; 20:1140–1145. <https://doi.org/10.1038/nbt747> PMID: 12368812
4. van Leeuwen JS, Vermeulen NP, Chris Vos J. Yeast as a humanized model organism for biotransformation-related toxicity. *Curr Drug Metab* 2012; 13(10):1464–1475. PMID: 22935071
5. Fuchs SM, Quasem I. Budding yeast as a model to study epigenetics. *Drug Discov Today Dis Models* 2014; 12:1–6.

6. Wysocki R, Tamás MJ. *Saccharomyces cerevisiae* as a model organism for elucidating arsenic tolerance mechanisms. In: Banfalvi G, editor. Cellular Effects of Heavy Metals. Netherlands: Springer; 2011. pp. 87–112.
7. Haugen AC, Kelley R, Collins JB, Tucker CJ, Deng C, Afshari CA, et al. Integrating phenotypic and expression profiles to map arsenic-response networks. *Genome Biol* 2004; 5(12):R95. <https://doi.org/10.1186/gb-2004-5-12-r95> PMID: 15575969
8. Vujcic M, Shroff M, Singh KK. Genetic determinants of mitochondrial response to arsenic in yeast *Saccharomyces cerevisiae*. *Cancer Res* 2007; 67(20):9740–9749. <https://doi.org/10.1158/0008-5472.CAN-07-1962> PMID: 17942904
9. Jin YH, Dunlap PE, McBride SJ, Al-Refai H, Bushel PR, Freedman JH. Global transcriptome and deletion profiles of yeast exposed to transition metals. *PLoS Genet* 2008; 4(4):e1000053. <https://doi.org/10.1371/journal.pgen.1000053> PMID: 18437200
10. Jo WJ, Loguinov A, Wintz H, Chang M, Smith AH, Kalman D, et al. Comparative functional genomic analysis identifies distinct and overlapping sets of genes required for resistance to monomethylarsonous acid (MMAIII) and arsenite (AsIII) in yeast. *Toxicol Sci* 2009; 111(2):424–436. <https://doi.org/10.1093/toxsci/kfp162> PMID: 19635755
11. Thorsen M, Perrone GG, Kristiansson E, Traini M, Ye T, Dawes IW, et al. Genetic basis of arsenite and cadmium tolerance in *Saccharomyces cerevisiae*. *BMC Genomics* 2009; 10:105. <https://doi.org/10.1186/1471-2164-10-105> PMID: 19284616
12. Zhou X, Arita A, Ellen TP, Liu X, Bai J, Rooney JP, et al. A genome-wide screen in *Saccharomyces cerevisiae* reveals pathways affected by arsenic toxicity. *Genomics* 2009; 94(5):294–307. <https://doi.org/10.1016/j.ygeno.2009.07.003> PMID: 19631266
13. Pan X, Reissman S, Douglas NR, Huang Z, Yuan DS, Wang X, et al. Trivalent arsenic inhibits the functions of chaperonin complex. *Genetics* 2010; 186(2):725–734. <https://doi.org/10.1534/genetics.110.117655> PMID: 20660648
14. Johnson AJ, Veljanoski F, O'Doherty PJ, Zaman MS, Petersingham G, Bailey TD, et al. Molecular insight into arsenic toxicity via the genome-wide deletion mutant screening of *Saccharomyces cerevisiae*. *Metallomics* 2016; 8(2):228–235. <https://doi.org/10.1039/c5mt00261c> PMID: 26688044
15. Thorsen M, Lagniel G, Kristiansson E, Junot C, Nerman O, Labarre J, et al. Quantitative transcriptome, proteome, and sulfur metabolite profiling of the *Saccharomyces cerevisiae* response to arsenite. *Physiol Genomics* 2007; 30(1):35–43. <https://doi.org/10.1152/physiolgenomics.00236.2006> PMID: 17327492
16. Farrugia G, Balzan R. Oxidative stress and programmed cell death in yeast. *Front Oncol* 2012; 2:64. <https://doi.org/10.3389/fonc.2012.00064> PMID: 22737670
17. Kumar NV, Yang J, Pillai JK, Rawat S, Solano C, Kumar A, et al. Arsenic directly binds to and activates the yeast AP-1-like transcription factor Yap8. *Mol Cell Biol* 2015; 36(6):913–922. <https://doi.org/10.1128/MCB.00842-15> PMID: 26711267
18. Wu WS, Li WH. Systematic identification of yeast cell cycle transcription factors using multiple data sources. *BMC Bioinformatics* 2008; 9:522. <https://doi.org/10.1186/1471-2105-9-522> PMID: 19061501
19. Balakrishnan R, Park J, Karra K, Hitz BC, Binkley G, Hong EL, et al. YeastMine—an integrated data warehouse for *Saccharomyces cerevisiae* data as a multipurpose tool-kit. *Database* 2012; 2012:bar062. <https://doi.org/10.1093/database/bar062> PMID: 22434830
20. Menezes RA, Amaral C, Batista-Nascimento L, Santos C, Ferreira RB, Devaux F, et al. Contribution of Yap1 towards *Saccharomyces cerevisiae* adaptation to arsenic-mediated oxidative stress. *Biochemical Journal* 2008; 414(2):301–311. <https://doi.org/10.1042/BJ20071537> PMID: 18439143
21. Toone WM, Morgan BA, Jones N. Redox control of AP-1-like factors in yeast and beyond. *Oncogene* 2001; 20(19):2336–2346. <https://doi.org/10.1038/sj.onc.1204384> PMID: 11402331
22. Zhang R, Yu W, Liang G, Jia Z, Chen Z, Zhao L, et al. Tumor suppressor candidate 1 suppresses cell growth and predicts better survival in glioblastoma. *Cell Mol Neurobiol*. 2017; 37(1):37–42. <https://doi.org/10.1007/s10571-016-0339-2> PMID: 26897357
23. Kanda M, Shimizu D, Nomoto S, Hibino S, Oya H, Takami H, et al. Clinical significance of expression and epigenetic profiling of TUSC1 in gastric cancer. *J Surg Oncol*. 2014; 110(2):136–144. <https://doi.org/10.1002/jso.23614> PMID: 24700496
24. Shimizu D, Kanda M, Nomoto S, Oya H, Takami H, Hibino S, et al. Identification of intragenic methylation in the TUSC1 gene as a novel prognostic marker of hepatocellular carcinoma. *Oncol Rep*. 2014; 31(3):1305–1313. <https://doi.org/10.3892/or.2013.2939> PMID: 24366000
25. Shan Z, Parker T, Wiest JS. Identifying novel homozygous deletions by microsatellite analysis and characterization of tumor suppressor candidate 1 gene, TUSC1, on chromosome 9p in human lung cancer. *Oncogene*. 2004; 23(39):6612–6620. <https://doi.org/10.1038/sj.onc.1207857> PMID: 15208665

26. Gardarin A, Chédin S, Lagniel G, Aude JC, Godat E, Catty P, et al. Endoplasmic reticulum is a major target of cadmium toxicity in yeast. *Mol Microbiol* 2010; 76(4):1034–1048. <https://doi.org/10.1111/j.1365-2958.2010.07166.x> PMID: 20444096
27. Guo L, Ganguly A, Sun L, Suo F, Du LL, Russell P. Global fitness profiling identifies arsenic and cadmium tolerance mechanisms in fission yeast. *G3 (Bethesda)* 2016; 6(10):3317–3333.
28. Hosiner D, Gerber S, Lichtenberg-Frate H, Glaser W, Schüller C, Klipp E. Impact of acute metal stress in *Saccharomyces cerevisiae*. *PLoS One* 2014; 9(1):e83330. <https://doi.org/10.1371/journal.pone.0083330> PMID: 24416162
29. Serero A, Lopes J, Nicolas A, Boiteux S. Yeast genes involved in cadmium tolerance: identification of DNA replication as a target of cadmium toxicity. *DNA Repair (Amst)* 2008; 7(8):1262–1275.
30. Cherry JM. The *Saccharomyces* Genome Database: A Tool for Discovery. *Cold Spring Harb Protoc.* 2015; 2015(12):pdb.top083840.