

Citation: Wu W-S, Hsieh Y-C, Lai F-J (2016) YCRD: Yeast Combinatorial Regulation Database. PLoS ONE 11(7): e0159213. doi:10.1371/journal. pone.0159213

Editor: Robertus A M de Bruin, University College London, UNITED KINGDOM

Received: April 12, 2016

Accepted: June 28, 2016

Published: July 8, 2016

Copyright: © 2016 Wu et al. This is an open access article distributed under the terms of the <u>Creative</u> <u>Commons Attribution License</u>, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

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

Funding: This work was supported by National Cheng Kung University and Ministry of Science and Technology of Taiwan [MOST-103-2221-E-006-174-MY2]. Funding for open access charge: National Cheng Kung University and 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.

RESEARCH ARTICLE

YCRD: Yeast Combinatorial Regulation Database

Wei-Sheng Wu*, Yen-Chen Hsieh, Fu-Jou Lai

Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan

* wessonwu@gmail.com

Abstract

In eukaryotes, the precise transcriptional control of gene expression is typically achieved through combinatorial regulation using cooperative transcription factors (TFs). Therefore, a database which provides regulatory associations between cooperative TFs and their target genes is helpful for biologists to study the molecular mechanisms of transcriptional regulation of gene expression. Because there is no such kind of databases in the public domain, this prompts us to construct a database, called Yeast Combinatorial Regulation Database (YCRD), which deposits 434,197 regulatory associations between 2535 cooperative TF pairs and 6243 genes. The comprehensive collection of more than 2500 cooperative TF pairs was retrieved from 17 existing algorithms in the literature. The target genes of a cooperative TF pair (e.g. TF1-TF2) are defined as the common target genes of TF1 and TF2, where a TF's experimentally validated target genes were downloaded from YEASTRACT database. In YCRD, users can (i) search the target genes of a cooperative TF pair of interest, (ii) search the cooperative TF pairs which regulate a gene of interest and (iii) identify important cooperative TF pairs which regulate a given set of genes. We believe that YCRD will be a valuable resource for yeast biologists to study combinatorial regulation of gene expression. YCRD is available at http://cosbi.ee.ncku.edu.tw/YCRD/ or http://cosbi2.ee. ncku.edu.tw/YCRD/.

Introduction

Transcriptional regulation is one of the major mechanisms for cells to control the timing, location, and amount of gene expression. The precise transcriptional control of gene expression is typically achieved through combinatorial regulation using cooperative transcription factors (TFs) [1-3]. Therefore, to understand how a gene of interest is transcriptionally regulated, it is crucial to know the cooperative TFs which function together to regulate the gene.

YEASTRACT [4] database provides up-to-date information on experimentally validated regulatory associations between a TF and its target genes. By querying YEASTRACT, users can know the TFs which regulate a specific gene. However, one key information is missing. Users cannot know whether these TFs function cooperatively or independently in regulating the expression of the specific gene. Therefore, it would be helpful to have a database which provides regulatory associations between cooperative TFs and their target genes. Because there is

no such kind of databases in the public domain, this prompts us to construct the Yeast Combinatorial Regulation Database (YCRD).

In YCRD, we collected more than 2500 cooperative TF pairs predicted by 17 existing algorithms in the literature [5–21]. As far as we know, this is the most comprehensive collection of predicted cooperative TF pairs in yeast (see Table 1 for details). Moreover, we retrieved a TF's experimentally validated target genes from YEASTRACT database [4]. Then for each cooperative TF pair (e.g. TF1-TF2), we define its target genes as the common target genes of TF1 and TF2. Therefore, the regulatory associations between cooperative TF pairs and their target genes are of biological significance because they are experimentally validated. In YCRD, users can search the target genes of a cooperative TF pair of interest validated by different types of experimental evidence (TF binding evidence or/and TF regulation evidence). Users can also search the cooperative TF pairs which regulate a gene of interest validated by different types of experimental evidence. Moreover, for a given set of genes, YCRD provides a tool for identifying important cooperative TF pairs which regulate these genes. We believe that YCRD will be a valuable resource for yeast biologists to study combinatorial regulation of gene expression. YCRD is freely available at http://cosbi.ee.ncku.edu.tw/YCRD/ or http://cosbi2.ee.ncku.edu.tw/YCRD/.

Construction and Contents

The comprehensive collection of predicted cooperative TF pairs from 17 existing algorithms in the literature

Many existing algorithms have been developed to predict cooperative TF pairs in yeast. Each algorithm utilized distinct biological rationales, reported its own set of predicted cooperative

Existing algorithms	# of PCTFPs / # of TF pairs under study	Existing algorithms	# of PCTFPs / # of TF pairs under study
Banerjee and Zhang (2003) [5]	$31/\binom{113}{2}$	Elati et al. (2007) [14]	$20/\binom{466}{2}$
Harbison et al. (2004) [6]	$94/\binom{203}{2}$	Datta and Zhao (2008) [15]	$25/\binom{203}{2}$
Nagamine et al. (2005) [7]	$24 / \binom{113}{2}$	Chuang et al. (2009) [16]	$13 / \binom{203}{2}$
Tsai et al. (2005) [8]	$18 / \binom{203}{2}$	Wang et al. (2009) [<u>17</u>]	$159/\binom{184}{2}$
Balaji et al. (2006) [9]	$3459 / \binom{157}{2}$	Yang et al. (2010) [18]	$186 / \binom{178}{2}$
Chang et al. (2006) [<u>10</u>]	$55 / \binom{203}{2}$	Chen et al. (2012) [<u>19</u>]	$221/\binom{203}{2}$
He et al. (2006) [11]	$30/\binom{113}{2}$	Lai et al. (2014) [20]	$27/\binom{186}{2}$
Wang (2006) [<u>12</u>]	$14/\binom{9}{2}$	Wu and Lai (2015) [21]	$50 / \binom{151}{2}$
Yu et al. (2006) [13]	$300/\binom{113}{2}$		

Table 1. The ratio of the number of predicted cooperative TF pairs (PCTFPs) to the number of TF pairs under study for each of the 17 existing algorithms in the literature.

 $\binom{N}{2} = \frac{N!}{2!(N-2)!}$ means all possible pairs formed by N TFs.

doi:10.1371/journal.pone.0159213.t001

TF pairs (PCTFPs) and performed differently under different evaluation indices [22,23]. We comprehensively collected 3755 distinct cooperative TF pairs from 17 existing algorithms (see Table 1 for details). Lai et al. [22] had good background information on 14 existing algorithms. Here we briefly introduce the other three algorithms. Baljai et al. [9] regarded a TF pair as a PCTFP if the observed number of shared target genes is higher than random expectation. Lai et al. [20] regarded a TF pair as a PCTFP if (i) the two TFs have a significantly higher number of common target genes than random expectation and (ii) their binding sites tend to be co-depleted of nucleosomes. Wu and Lai [21] regarded a TF pair as a PCTFP if the overlap of the targets of these two TFs is higher than random expectation. As far as we know, our collection of 3755 distinct cooperative TF pairs is the most comprehensive collection of predicted cooperative TF pairs in the literature. Note that we collected predicted rather than experimentally verified cooperative TF pairs in the literature is too small to construct a useful database of combinatorial regulation of gene expression.

Among the 3755 collected PCTFPs, 1133 PCTFPs which contained non-TF names were removed. A protein name is regarded as a TF name only if it is annotated as a TF (activator/ repressor) or a transcription co-factor in the regulation page of SGD [24]. After this curation, we obtained 2622 PCTFPs among 143 TFs. We then removed 87 PCTFPs which have no target genes (see the next subsection for how to define the target genes of a PCTFP). Finally, 2535 cooperative TF pairs were used in this database. In order to help users judge the biological plausibility of a PCTFP, we provide three types of validation. The first type, called Algorithm Evidence, tells users the ratio of the number of existing algorithms which predict this PCTFP (e.g. TF1-TF2) to the number of existing algorithms which study this TF pair (e.g. TF1-TF2). The higher the ratio is, the higher the confidence of this PCTFP is. The second/third type, called Physical/Genetic Evidence, tells users the number of publications which experimentally show that this PCTFP has physical/genetic interactions. Having physical or genetic interactions strengthens the confidence of the biological plausibility of this PCTFP. The physical and genetic interaction data were retrieved from BioGRID database [25].

Defining the target genes of a cooperative TF pair

YEASTRACT [4] database uses three types of experimental evidence (TFB, TFR and TFB&TFR) to define a TF's experimentally validated target genes. TFB (i.e. TF binding) means the experimental evidence (from band-shift, foot-printing or ChIP assay) showing that a TF binds to the promoters of its target genes. TFR (i.e. TF regulation) means the experimental evidence (from detailed gene by gene analysis or genome-wide expression analysis) showing that a TF perturbation (knockout or over-expression) causes a significant change in the expression of its target genes. TFB and TFR evidence exist to support the regulatory associations between a TF and its target genes.

In YCRD, the target genes of a cooperative TF pair (e.g. TF1-TF2) are defined as the common target genes of TF1 and TF2, where a TF's experimentally validated target genes were retrieved from the YEASTRACT database [4]. For example, we say that X is a target gene of the cooperative TF pair (TF1-TF2) validated by TFB evidence if both the regulatory association between TF1 and X and the association between TF2 and X are validated by TFB evidence. However, in the above definition, the TFB evidences of TF1-X and TF2-X may not be in the same biological condition. If we further require that the TFB evidences of TF1-X and TF2-X must be in the same biological condition, then, on average, the number of the target genes of a cooperative TF pair reduces to 53% (33346/63418) under this stringent definition (see <u>S1 File</u> for details). Since many biological conditions have not been tested for TFB evidence in the



Table 2. The detailed statistics of YCRD.

	# of regulatory associations	# of cooperative TF pairs which can be assigned at least one target gene	# of genes which can be assigned at least one cooperative TF pair
Regulatory associations validated by TFB evidence	63,418	2,441	4,119
Regulatory associations validated by TFR evidence	364,360	2,519	6,165
Regulatory associations validated by TFB&TFR evidence	6,419	1,013	1,380

doi:10.1371/journal.pone.0159213.t002

literature, it may be too conservative to use the stringent way to define the target genes of a cooperative TF pair. Therefore, we use the less stringent way to define the target genes of a cooperative TF pair.

After collecting cooperative TF pairs and defining the target genes of each cooperative TF pair, a web interface is then constructed for users to query the regulatory associations (validated by TFB, TFR or TFB&TFR) between cooperative TF pairs and their target genes. The detailed statistics of YCRD could be found in <u>Table 2</u>.

Identifying important cooperative TF pairs which regulate a given set of genes

When researchers have a set of genes (e.g. upregulated genes under a specific biological condition), they probably want to know the cooperative TF pairs which play important roles in regulating these genes. To meet this need, YCRD provides a tool for identifying important cooperative TF pairs which regulate a given set of genes. In YCRD, a cooperative TF pair is regarded as an important regulator if its target genes are enriched in the given set of genes. The hypergeometric distribution is used to test the statistical significance of enrichment [26]. The procedure for checking whether a cooperative TF pair is an important regulator for a given set of genes is as follows. Let *S* be the set of target genes of a cooperative TF pair of interest, *G* be the given set of genes, $T = S \cap G$ be the set of the cooperative TF pair's target genes which are also in the given set of genes, and *F* be the set of all genes in the yeast genome. Then the *p*-value for rejecting the null hypothesis (H₀: the cooperative TF pair's target genes are not enriched in the given set of genes) is calculated as

$$p_{value} = P(x \ge |T|) = \sum_{x \ge |T|} \frac{\binom{|S|}{x} \binom{|F| - |S|}{|G| - x}}{\binom{|F|}{|G|}}$$
(1)

where |G| means the number of genes in set *G*. This *p*-value is then corrected by the Bonferroni correction to represent the true alpha level in the multiple hypotheses testing. A cooperative TF pair of interest is called an important regulator for the given set of genes if the Bonferroni-corrected *p*-value is less than the threshold determined by users. We also allow users to specify the false discovery rate (FDR) when identifying important cooperative TF pairs which regulate a given set of genes.

Utility and Discussion

Database interface

YCRD provides a search mode and a browse mode. In the search mode, users have two possible ways to search YCRD. First, users can (i) select the experimental evidence (TFB, TFR or

TFB&TFR) of the regulatory associations and (ii) select a cooperative TF pair of interest (see Fig 1A). Then YCRD returns the target genes of the selected cooperative TF pair shown in a table and a figure (see Fig 1B). The publications of the experimental evidence of the regulatory associations are also provided (see Fig 1C). Second, users can (i) select the experimental evidence (TFB, TFR or TFB&TFR) and (ii) type in the name of a gene of interest (see Fig 2A). Then YCRD returns the cooperative TF pairs which regulate the input gene shown in a table and a figure (see Fig 2B). The publications of the experimental evidence are also provided (see Fig 2C). Moreover, three types of validation (Algorithm Evidence, Physical Interaction Evidence and Genetic Interaction Evidence) for each cooperative TF pair are given (see Fig 2D).

In the browse mode, users have two possible ways to browse the regulatory associations (validated by TFB, TFR or TFB&TFR) between cooperative TF pairs and their target genes (see Fig <u>3A</u>). First, when users browse YCRD by the name of a cooperative TF pair, users will be given the target genes of this cooperative TF pair (see Fig <u>3B</u>). Second, when users browse YCRD by a gene name, users will be given the cooperative TF pairs that regulate this gene (see Fig <u>3C</u>).

YCRD also provides a tool for identifying important cooperative TF pairs which regulate a given set of genes. To use this tool, users have to (i) input a set of genes of interest, (ii) choose the experimental evidence (TFB, TFR or TFB&TFR) of the regulatory associations, and (iii) set the threshold of the Bonferroni-corrected p-value, where p-value is calculated using Eq (1) (see Fig 4A). YCRD then returns the important cooperative TF pairs which regulate the set of genes of interest (see Fig 4B).

Three case studies

We use three case studies to show that YCRD is likely to return biologically meaningful results. In the first case study, we would like to search the target genes of the CTFP (Gat1-Gln3) with the regulatory associations validated by TFB&TFR. After submission, YCRD returns six target genes (*DAL5, ENA1, GDH2, MEP2, PUT1, UGA4*) (see Fig 1B). TFB evidence shows that each target gene's promoter is bound by Gat1 and Gln3. TFR evidence shows that the expression of each target gene is significantly changed when perturbing the expression of the gene encoding Gat1 or Gln3 (see Fig 1C). We validate the biological relevance of the six identified target genes as follows. First, it is known that Gat1 and Gln3 are two TFs which cooperatively activate a set of genes involved in nitrogen catabolite pathways [27]. Strikingly, four (*DAL5, GDH2, PUT1* and *UGA4*) of the six identified target genes are known genes involved in nitrogen catabolite pathways [27]. Second, the expression of *ENA1*, a gene encoding a lithium and sodium ion transporter essential for salt tolerance in yeast, is known to cooperatively regulate *MEP2*, a gene encoding an ammonium permease, whose expression is induced under limiting nitrogen conditions [29].

In the second case study, we would like to search the CTFPs which regulate the gene *DAL5* with the regulatory associations validated by TFB&TFR. After submission, YCRD returns three CTFPs (Gat1-Gln3, Gcn4-Gln3 and Gat1-Gcn4) (see Fig 2B). TFB evidence shows that the promoter of *DAL5* is bound by Gat1, Gcn4 and Gln3. TFR evidence shows that the expression of *DAL5* is significantly changed when perturbing the expression of the gene encoding Gat1, Gcn4 or Gln3 (see Fig 2C). We validate the biological relevance of the three identified CTFPs as follows. First, it is known that *DAL5*, a gene encoding an allantoate permease, is involved in nitrogen catabolite pathways [27]. The identified CTFP Gat1-Gln3 is known to form a transcriptional complex to regulate *DAL5* in response to nitrogen limitation [27]. Second, the identified CTFP Gcn4-Gln3 is known to form a transcriptional complex to regulate *DAL5* when cells are grown under nitrogen derepressive conditions and amino acid deprivation [30].



doi:10.1371/journal.pone.0159213.g001

(a)	Regulatory Associ TFB evidence	ations Validated by	Search Re	guiatory Association	ons		
	TFB&TFR eviden	ee • Search th	e CTEPs which regulat	e the input gene :	DAL5		
• •			SUBMIT				
b)							
he queried gene: DALS	Description	_			a of CTE	Ps which regulate the to	rget gene
Gene Name Notes ML5 YJR152W allantoate permease(U	REP1 Allantoate per and ureidosuc	mease; ureidosuccinate permease; als cinate; expression is constitutive but se	transports dipeptides, th nsitive to nitrogen catabol	ough with lower affinity ite repression	than for allantoate 3	i by TFB&TFR evidence	
ne 3 cooperative TP pairs w	thm Evidencelul	1 Dhumical Evidence1+[11	Conotio Evidence	21			
(1/16 Algorithm E	vidence]	T Physical Evidence (11	Genetic Evidenci	2)			
The Cooperative TF Pair(TF1-TF2) TF1 TF2	Evidence of Cor	Ø of evidence operativity IFB	of TF1-DAL5 regulator	y association	# of evidence of TF2-	DALS regulatory asso	iation
at1 Gcn4	A(1/16)	1	4	_	1	2	
at1 Gin3	A(415)+P1+G11	1	4		3	7	
ine Ono	A(418)112103	!'	ŕ		,	· · ·	
		4			6	***	
:)					Layout: Force	Directed Circle	Radia
The 1 <u>TFB evidence</u> of TF-t Reternce Geors I et al., Mol Cell Biol, 2011.3	carget gone (Gat1- 11(1):92-104	-DAL5) pair C Evidenc Dret: Ca	s Code :	Association Type Not applicable	Layout: Force	Directed Circle	Radial
The 1 <u>TFB evidence</u> of TF-4 Reference Genera Let <i>a</i> , Mai Cell Del, 2011.3 The 4 <u>TFR evidence</u> of TF-4	arget gene (Gat1- 1(1) 52-104 arget gene (Gat1-	-DAL5) pair Dried: Ca -DAL5) pair	s Cods ¢ P	Association Type Not applicable	Layout: Force Environ Rapamyca	Directed Circle	Radia
The 1 <u>TEB evidence</u> of TF-4 Reference Games <i>tet</i> at , Mil CeB Biol 20113 The 4 <u>TER evidence</u> of TF-4 Reference	arget gene (Gat1 1(1) 52-104 arget gene (Gat1	-DALD) pair Dract CR -DALD) pair -DALD) pair Extense Cole	e Cods e pp Association 7,00	Association Type Not applicable	Layout: Earce Rearry Control Rearry	Directed Circle mental Condition n treatment monrail median with 0.	Radia
The 1 TEB evidence of TF4 Reference Corons of an Mol Cell Box 2013 The 1 TER evidence of TF4 Reference Continue An et al., Mol Cell Box 19 Continue An et al., Mol Cell Box 19 Cell Box	arget gene (Gat1- 1(1) 52-104 arget gene (Gat1- 06, 16(1) 847-58 cal Chemisty,	DAL5) par Desc D Desc Desc D Desc Desc D Desc Desc Desc Desc Desc Desc Desc Desc	e Code C p Association e Postive Postive	Association Type Not applicable Expresented CC Expresented CC Statuments glutament	Layout: Force Repanyon Repanyon cells cultured a Wowthern's	Directed Circle mental Constitue mmannal medium with 0	Radia
The 1 TEB evidence of TF-4 Prime Const Let A Link College 12 11 1 Const Let A Link College 12 11 1 Const Let A Link College 11	arget gene (Gatt 1(1) 52-104 arget gene (Gatt 90; 16(2) 847-58 04 Chenshy, 3) 591-74	DALS) pair Control COALS) pair COALS) pair Coalest Co	р - Сек с р - С	Association Type Net application Environmental Ci- Environmental Ci- Strangenerical phase of the perform glutterine Cells calculation	Layout: Ecroe Reparyor estimation cashs callevaled a Wowmenny synthetic growth medium supplie	Directed Circle mental Controls mental Controls mental Controls mental meduan with 0 mental without 02 mentals	Radia
C) The 4 322 residence of 17-4 Commentary and the commentary of 17-4 Commentary and the commentary of 17-4 Commentary and the commentary of the commentary o	arget gene (Gatt http://datgene gene (Gatt of No(3) M7-49 of Okensky, 3) 801-74 0(23) 4407-012	DAL5) par CAL5) par CAL5) par CAL5) par CAL50 par	Conv to the second seco	Association Type Not applicable O the profession O the pr	Layout: Force Reserve entering software	Directed Circle meeted Constitue meeted Constitue meeted without vith o meeted without 0.2 me	Radia
C) The 4 322 avalance of 174 Manual Const of at And Cab Bas 2013 The 4 The Statematic Cab Bas 2013 The 4 A The Statematic Cab Bas 1 Constant of at More that the Statematic Cab Bas 1 Constant of at More that the Statematic Cab Bas 1 Constant of at More that the Statematic Cab Bas 1 Constant of at More that the Statematic Cab Bas 1 Constant of at Data Caba 2013 The Statematic Cab Bas 2013 The Statematic Cab Bas 1 Constant of at Cab Caba 2013 The Statematic Cab Bas 1 Constant of at Cab Caba 2013 The Statematic Cab Bas 1 Constant of at Cab Caba 2013 The Statematic Cab Bas 1 Constant of at Cab Caba 2013 The Statematic Cab Bas 1 Constant of Advancements The Statematic Cab Bas 1 Constant of Cab Bas 1 Const	arget gene (Gatt 111) 52-164 Arget gene (Gatt 06, 1603) 847-60 06, 1603) 847-60 06, 1603) 847-60 06, 1603, 1604 06, 160400000000000000000000000000000000000	DAL5y par DAL5y par	Color p Color Association Postore Postore Postore Postore	Association Type Not applicable Caponetical phone 0.1% period guidannie Caponetical phone 0.1% period	Layou: Force Reserve Reserve settem setter collusion in Wowmhams without growth methum supple of	Directed Circle mental Caustion mental Caustion mental medium with 0 mented without 0.2 men	Radia
c) The 4.32B andemony of 17-4 Method Goose is at a fund call bits (2013) The 4.32B excidence of 17-4 Method Colonics Act at a fund call bits (2013) Colonics Act at a fund call bits (2013) Colonics (at al. fund call bits (2013) Colonics (2013)	arget gene (Gatf 1(1)22-194 arget gene (Gatf 9), 1(2)) 817-58 old Chemistry, 3),861-74 90(22) 4407-912	OALDy pair Deck Dy pair Deck Dy pair OALDy pair Deck Dy pair Deck D	e Cole p 2 2 2 2 2 2 2 2 2 2 2 2 2	Accession Type Ret application 0.1% paymented place galaxies Regenerated places and accession of the Regeneration destination of the Regneration destination destination of the Regneration destination	Layou: Force Reserve Reserve addition a	Directed Circle	Radia
b) The 1 THB weddence of TF-4 Warren Gross tell, "And Call Box 2011 3 The 4 3 THR veddence of TF-4 Variance Colonge of a THA second Box 201 Colonge of a THA second Box 2012 Colong tell, a Second Box 2012 Colong tell, a Second Box 2012 Colong tell, a Second Box 2012 Colonge of A Se	arget gene (Gatt 1(1) 120 104 80 16(2),847 30 80 16(2),847 30 2001-74 20(22) 4487 412 2002 2001-74 20(22) 4487 412	CALS) pair Course Course Course Course Course C	Police Po	Accession of the second	Layou: Force Reserve Reserve assessment asse	Directed Circle	Radia
C) The 1 TEB variations of TF4 Constant of the	arget gene (Gat1 t(†12-104 arget gene (Gat1 d) (101) 847-84 d) (2010) 847-84 B((2) 4487-612 C) (2) (2) 4487-612 C) (2) (2) (2) (2) (2) (2) (2) (2) (2) (2	CALS) pair CALS) pair CALS par CALS par CA	e Celeir p 2 4 2 4 2 4 2 4 2 4 2 4 2 4 2 4 2 4 2 4 2 4 2 4 2 4 2 4 2 4 2 4 2 4 2 4 4 4 4 4 4 4 4 4 4 4 4 4	An account on type Not applicable Composition of the Composition of th	Layou: Force Reserve R	Directed Circle	Radia
C) The 1 TEU workdence of TF44 Texture Control of a texture of the definition of the	arget gans (Gatt 11) 10:100-100 arget gans (Gatt 00:10) 017-00 30:000-00 30:00-00 30:00-00 30:00-00 30:00-00 30:000-00 3	CALS) pair CALS) pair CALS) pair CALS, par CALS, pa	s Club p	Accession Type Net explosion Constraints (Constraints) additions a	Layou: Force Environment Recompose environment environ	Directed Circle	Radia
The 1 TH subdemon of TF4 The 1 TH subdemon of TF4 The 4 TH subdemon of TF4 The 4 TH subdemon of TF4 The 4 TH subdemon of TF4 The 5 TH subdemon of TF4 TH s	arget gene (Gatt 11) 52-164 arget gene (Gatt 0) 10() 97-50 30 10() 97-50	CALS) pair CALS) pair CALS) pair CALS) par CALS) par CALSS pa	Postore P	Accession of the second of the	Layou: Force Environment Recompose R	Directed Circle	Radia
The 1 THE websites of 11 TH Memory and 12 TH	arget gene (Gatt 11) 520-184 arget gene (Gatt 20) 50/3 67-50 20 20/1-44 20 20/1-44 20/1-44 20/1-44	CALS) pair CALS) pair CALS) pair CALS) pair CALS) pair CALS) pair CALSO	s Cink 2017	Reasonance Type Relations December Second Control Second Control Contr	Layou: Force Environ Recently a Moderney of Recently	Directed Circle anneal Circle anneal Circle anneal Circle anneal Circle anneal Circle anneal Structure anneal Str	Radia
The 1 THE verdence of The 4 Network of the 1 Model and 2 Model an	arget gene (Gatt 11) 520-164 arget gene (Gatt 20) 500 647-54 at Construction 20 201-74 at Construction 20 20 201-74 at Construction 20 201-74 at Construction 20 20 201-74 at Construction 20 201-74 at Construction 20 20 20 20 20 20 20 20 20 20 20 20 20	CALS) pair CALS) pair CALS) pair CALS) pair CALS) pair CALSS	s Color pr	Reasonments of Comparison of the Comparison of t	Layou: Force Environ Recently a Moderney of Recently	Directed Circle and Circles and Circles and Circ	Radia
The 1 THE websites of 11 TH websites The 1 THE websites of 11 TH Websites The 1 THE websites of 11 TH Websites Collected at 1 Most Cell Data (2017) THE ALL THE websites of 11 TH Websites Collected at 1 Most Cell Data (2017) Collected at 1 Most Cell D	arget gene (Gatt 11) 122-104 arget gene (Gatt 20) 120 147-40 at Compty 2) 201-74 at Co	CALS) pair CALS)	N Galaxy Image: Control of the second s	Accession of the second	Layou: Force Reserved	Directed Circle	Radia
The 1 THE websites of 11 TH websites and the 1 TH of 1 TH websites and the 1 TH of 1	arget gene (Gatt 11) 122-104 arget gene (Gatt 20) 120 147-40 at Compty 2010-147 at Compt 2010-147 at Compt 2010 at Compt 2010-147 at Compt 2010-147 at Compt 2010 at Com	CALS) pair CALS)	сонструкти и при и	Accession of the second	Layou: Force Reserved	Directed Circle	Radia
C) The 1 JUB endemone of 15-4 Endemone To 1 JUB endemone of 15-4 Endemone To 1 JUB endemone of 15-4 Endemone To 1 JUB endemone of 15-4 Endemone Control and a low Call Base (11) Control and a low Cal	arget gans (Gat 11) 120-104 arget gans (Gat arget gans (Gat 0) 10) 97-09 (Gat 0) 10) 10) 100 100 (G	CALLS) pair CALLS pair CALLS pair CALLS par CALLS p	Close 2 P	Accession of the second of the	Layou: Force Layou: Force Reserved Reserved	Directed Circle and Circle and Circ	Radia
C) To 1 JUL encloses of 15-4 Long To 1 JUL encloses of 15-4 Lo	arget gans (Gat Itt) 102-104 arget gans (Gat arget gans (Gat arget gans (Gat arget gans (Gat arget gan (Gat arget gan (Gat	CALLS) pair CALLS	Construction	Consention Service Consention Consentio	Layou: Force Reserves Reserves automation au	Directed Circle and Circles and Circles a	Radia
e) In 1 JUI and and an of 11 In 1 JUI and an of a star of 11 In 1 In 11 In	arget gans (Gat Itt) 102-104 arget gans (Gat arget gans (Gat arget gans (Gat arget gans (Gat arget gan (Gat arget gan (Gat (Gat arget gan (Gat	CALS) pair CALS,	Close 2 P	Concentration 2012 Concentration 2012 Concentration Concentrat	Layou: Force Reserves Reserves and addition at the cultured as Weak-means at the cultured as Wea	Directed Circle and Circles and Circles mining industry mining industr	Radia
C) To 1 JUL and and on 1 JUL To 1 JUL and and an 1 JUL To 1 JUL and and an 1 JUL To 1 JUL and and an 1 JUL To 1	arget gans (Gat Itt) 120-104 arget gans (Gat arget gans (Gat arget gans (Gat arget gans) (Gat arget gan (Gat arget gan (Gat (Gat arget gan (Gat	CALLS) pair CALLS	Close 2 P	Accession Terms Terreneward Co Desperation from Constructions Terreneward Co Desperation from Constructions Constructio	Layou: Force Reserves Reserves automatical states cultured as Weaksheers at 22791-705 x	Directed Circle and Circle	Radia
C) To a 1222 audiona of 1974 To a 1222 audio	arget gans (Gat Itt) 50-104 arget gans (Gat arget gans (Gat arget gans (Gat arget gans (Gat arget gan (Gat arget gan (Gat (DAL5) part DAL5) part CAL5.9			Layou: Force Layou: Force Reserved Rese	Directed Crote and Crote a	Radia
C) To a 1272 vacious of 17-4 V	arget gene (Gett tt(112-104 arget gene (Gett ge, 14(2)-947-49 ge, 14(2)-947-49 gene (Gett gene (Gett) gene (Gett) gene (Gett) gene (Gett) gene (Gett) gene (CALS) part CALS part	Image: Section 2014 S	Looperators and a second	Layou: Error Reserved	Directed Crother among Constitution invested authors (1) 2 mice among a	Radia
C) To a 1272 residence of 17-4	arget gene (Gatt tt(192-194 ar get gene (Gatt get) 100,101,104 100,100,047,40 200,100,047,40 200,0400,0400,04 200,0400,0400,0400,0400,0400,0400,0400,			Assessment of the second secon	Layou: Eoros Reserved	Directed Crote annual countries meened metalow with of a me meened without 0 2 me annual metalow with of a 2 me annual metalow with of a 2 me annual metalow with of a 2 me annual metalow without 0 2 me annual	Radia
C) The 1 THE undersome of Tri-1 the THE Undersome of Tri-1 The AT THE ACCENT OF THE UNDERSOME OF THE UNDERSO	arget gene (Gat) tr(1102104 arget gene (Gat) arget gene (Gat) (Cat) 607-84 (Cat)		Close 2 P Second Seco	Sector and a secto	Layou: Force Reading and Anti- Reading and Anti-	Directed Crotheres Control Con	Radia
C) To I TIP undersom of Tri-I TIP undersom of TIP undersom of TIP undersom of T	arget gene (Gat) titt 20:104 arget gene (Gat) arget gene		Control	Secondary 2014	Layou: Force Reserved Re	Directed Circle	Radia
C) The 1 THE underseare of Tri-4 tores Content of a tri-back of the 2 and 2 Content of a tri-back of the 2 and 2 Content of a tri-back of the 2 and 2 Content of a tri-back of the 2 and 2 Content of a tri-back of the 2 and 2 Content of a tri-back of the 2 and 2 Content of a tri-back of the 2 and 2 Content of a tri-back of the 2 and 2 Content of a tri-back of the 2 and 2 Content of a tri-back of the 2 and 2 Content of a tri-back of the 2 and 2 Content of a tri-back of the 2 and 2 Content of a tri-back of the 2 and 2 Content of a tri-back of the 2 and 2 Content of a tri-back of the 2 and 2 Content of a tri-back of the 2 and 2 Content of the 2 Content of the 2 and 2 Content of the	arget gene (Gatt tt) 20194 arget gene (Gatt doi:10.1014/101 00.102014/0140 201000000 201000000 201000000 201000000 201000000 201000000 201000000 201000000 201000000 2010000000 2010000000 2010000000 2010000000 2010000000 20100000000 20100000000 20100000000 20100000000 2010000000000		Eleverence 2 P Second	Sectoration 2012 Sectoration	Layou: Force Reserved Re	Directed Circle	Radia
C) The I TIP evidence of TF4 Content of a Characterized and the Content of the Characterized and the Content of the Characterized and the Characterized an	arget gene (Gatt Itt) 20194 arget gene (Gatt Gatt) 20194 Gatta gene (Gatt Gatta) 20194 Gatta gene (Gatta) Gatta gene (G		Elements 2 P Parabase P Parabase </td <td>Sectoreation 2012 Sectoreation 2012 Sectoreation</td> <td>Layou: Force Calence Control of C</td> <td>Directed Circle</td> <td>Radia</td>	Sectoreation 2012 Sectoreation	Layou: Force Calence Control of C	Directed Circle	Radia

Fig 2. Search the cooperative TF pairs which regulate a gene of interest. (a) The input page. (b) The result page. (c) The publications of the experimental evidence of the regulatory associations. (d) Three types of validation (Algorithm Evidence, Physical Evidence and Genetic Evidence) for each cooperative TF pair.

doi:10.1371/journal.pone.0159213.g002

(a)

Browse by	Cooperative T	F Pairs								
Browse by	y Genes in									
Chrl Cl	hrll Chrill	ChrlV ChrV	ChrVI ChrVII	ChrVIII C	hrlX ChrX C	ChrXI ChrXII	ChrXIII ChrXI	V ChrXV	ChrXVI	
Regulator	ry associatio	ons are validated	by <u>TFR evidence</u>	1						
Browse by	Cooperative T	F Pairs								
Browse by	y Genes in									
Chrl Cl	hril Chrili	ChrlV ChrV	ChrVI ChrVII	ChrVIII C	hrlX ChrX C	ChrXI ChrXII	ChrXIII ChrXI	V ChrXV	ChrXVI	
Regulator	ry associatio	ons are validated	by TFB&TFR evi	dence						
Browse by	Cooperative T	F Pairs								
Browse by	y Genes in									
Chrl Cl	hril Chrili	ChrlV ChrV	ChrVI ChrVII	ChrVIII C	hrlX ChrX C	ChrXI	ChrXIII ChrXI	V ChrXV	ChrXVI	
(h)		L								
(2)										
# of co	operativ	e TF pairs	which can b	pe queried	I: 1013					
Regulate	ory associat	ions are validate	d by <u>TFB</u> & <u>TFR</u> ev	vidence						
A(2/12)+P	21+64 : [<u>2/</u>	12 Algorithm Evi	dence]+[1 Phys	ical Evidence]	+[4 Genetic Evi	idence]				
A(3/11)+P	3 : [<u>3/11</u> /	Algorithm Evider	nce]+[3 Physica	I Evidence]						
A(1/12)+G	a1 : [<u>1/12</u>	Algorithm Evide	nce]+[1 Genetic	Evidence]						
A(1/16)	: [<u>1/16 Al</u>	porithm Evidence	<u>e]</u>							
The										
The Coc	operative IF F			# of target g	genes of the <u>CTFP</u>	(TF1-TF2) valid	ated by <u>TFB</u> & <u>TFR</u>	evidence		
TF1	¢ TF2	Evidence of Co	operativity	\$						
Abf1	Cat8	A(1/4)		8						
Abf1	Cha4	A(1/15)		1						
Abf1	Ino4	A(1/15)		()		1				
Abf1	Met31	A(1/15)		4						
Abf1 Abf1	Met31 Adr1	A(1/15) A(1/15)		11						
Abf1 Abf1 The 4 tar	Met31 Adr1	A(1/15) A(1/15) f the cooperative	TF pair (Abf1-Inc	11 11 04) validated by	/ TFB&TFR evide	nce				
Abf1 Abf1 The 4 tar	Met31 Adr1 rget genes o	A(1/15) A(1/15) f the cooperative	TF pair (Abf1-Inc	4 11 04) validated by	7 TFB&TFR evide	ence	# of evidence	of Ino4-target g	jene regulatory as	sociation
Abf1 Abf1 The 4 tar	Met31 Adr1 Gene + Sy:	A (1/15) A (1/15) f the cooperative stematic Name	TF pair (Abf1-Inc # of evidence c IEB	4 11 of Abf1-target ge	7 TFB&TFR evide ene regulatory as: † TER	sociation	# of evidence	of Ino4-target g	tene regulatory as	sociation
Abf1 Abf1 The 4 tar Target G	Adr1 addr1 addr1 addr1 addr1 block addr2 addr2	Auto	TF pair (Abf1-Inc # of evidence c TEB 2	4 11 04) validated by of Abf1-target ge	r TFB&TFR evide ene regulatory as TFR 2	sociation	# of evidence # of evidence 4 4	of Ino4-target g	ene regulatory as TER 1	sociation
Abf1 Abf1 The 4 tar Target G CDS1 DPP1 FAS1	Met31 Adr1 rget genes o Gene Que Syn YBR YDR YDR	Auto Auto f the cooperative stematic Name	TF pair (Abf1-Inc # of evidence c IFB 2 1 1	4 11 of Abf1-target ge	r TFB&TFR evide ene regulatory as: IR 2 2 1	sociation	# of evidence • 1758 4 1 4	of ino4-target g	ene regulatory as e TER 1 3	sociation
Abf1 Abf1 The 4 tar Target G CDS1 DPP1 FAS1 GJT1	Met31 Adr1 Adr1 Gete \$ System VBR VDR VDR VCR VCR	Auro Auro Auro fthe cooperative stematic Name 029C 284C 882W 0098C	TF pair (Abf1-Inc # of evidence of IFB 2 1 1 1	4 11 of Abf1-target ge	y TFB&TFR evide ene regulatory ase TR 2 2 1 1	sociation	# of evidence φ 7729 4 4 1 4 2 2	of Ino4-target g	ene regulatory as present to the second sec	sociation
Abf1 Abf1 The 4 tar Target G CDS1 DPP1 FAS1 GIT1	Met31 Adr1 Gene So System YDR YDR YCR	Auns Auns I the cooperative stematic Name 2020C 284C 284C 284C 284C	TF pair (Abf1-Inco # of evidence of TFD 2 1 1 1 1	o4) validated by of Abf1-target ge	y TFB&TFR evide ene regulatory ase 178 2 2 1 1	sociation	# of evidence • 729 • 4 • 1 • 4 • 2	of Ino4-target g	e regulatory as e regulatory as a 1 1 3 3 2	sociation
Abf1 Abf1 The 4 tarr Target G CDS1 DPP1 FAS1 GJT1	Met31 Adr1 Adr1 Gene Sys YBR YDR YDR YDR YCR	Auns Auns If the cooperative stematic Name 2284C 2284C 2842 2842 2842	TF pair (Abf1-Ince # of evidence of TFB 2 1 1 1	4 11)-4) validated by of Abf1-larget ge	r TFB&TFR evide	sociation	e JFB 4 4 1 2	of Ino4-target g	e regulatory as f TER 1 1 3 3 2	sociation
Abf1 Abf1 The 4 tar CDS1 DPP1 FAS1 GIT1	Met31 Adr1 Adr1 Genes o YBR YDR YDR YCR	ACHS ACHS If the cooperative stematic Name 2020: 284C 284C 284C 284C 284C 284C	TF pair (Abf1-Inc # of evidence of TFB 2 1 1 1 1	4 11)-4) validated by of Abf1-larget ge	r TFB&TFR evide	nce	# of evidence g 110 4 1 1 1 2 2	of Inol-larget g	ene regulatory as e TER 1 1 3 2 2	sociation
Abf1 Abf1 The 4 tar Target G CDS1 DPP1 FAS1 GIT1 (C)	Met31 Adr1 Adr1 Get genes o Sys YBR YDR YDR YCR Ness in C	the cooperative stematic Name st	TF pair (Abfiling	4 11 >4) validated by of AbH-larget ge c	y TFB& TFR evide ene regulatory as: TFR 2 2 1 1	nce	B of evidence TEE TEE A	of Ino4-target g	ene regulatory as TITR 1 3 2	sociation
Abri Abri The 4 tar Target G CDS1 DPP1 FAS1 GIT1 (C) # of ge	Met31 Adr1 Adr1 Get genes o Yor Yor Yor Yor Yor Yor Ness in C	the cooperative	TE pair (Ab11-inc Ø of evidence of TE 2 1 1 1 an be queri	4 11 	rEBATER evide	sociation	# of evidence 9 2725 4 1 4 2	of Ino4-target g	ene regulatory as 1775	sociation
Abri Abri The 4 tar Target G CDS1 DPP1 FAS1 GIT1 (C) # of ge Regulato	Met31 Adr1 Adr1 Adr1 YBR YDR YDR YDR YRL YCR	the cooperative stematic Name 224C 224C 224C 224C 224C 224C 224C 224	TF pair (Ab14-inc # of evidence of TE 2 1 1 1 an be queri t by <u>IFB&TFR ev</u>	4 11 24) validated by of AbH-larget ge of abH-larget ge abH-larget ge abH-la	TFBS TFR evide ene regulatory as 2 2 1 1	sociation	# of evidence e 170 4 1 1 2 2	of Inol-target g	eneropulatory are 1777 - 2000 1777 - 2000 1777 - 2000 1777 - 2000 1777 - 2000 1777 - 2000	sociation
Abri Abri The 4 tar Target G CDS1 DPP1 FAS1 GIT1 (C) # of ge Regulate	Add1 Add1 Add1 Add1 Add1 Add1 Add1 Add1	the cooperative stematic Name 2840 2840 2840 2840 2840 2840 2840 2840	TF pair (Abff-inc # of evidence of TF 2 1 1 1 an be queri an be queri to by TESTER ev Every first first of the f	a 11 11 11 11 11 11 11 11 11 11 11 11 11	TEBS TER evide ene regulatory as TEBS 2 2 1 1	sociation	B of evidence Training A	e of Inol-target g	ere regulatory and 1 3 2 2	sociation
Add Add The 4 tar Target G COS1 DPP1 PAS1 OUT1 (C) (C) Regulato	Met31 Adri sene Sy YOR YOR YOR YOR YOR YOR	the cooperative stematic Name stematic Name coop coop coop coop coop coop coop coo	TF pair (Abff-inc # of evidence of TFB 2 1 1 1 1 1 1 1 1 1 1 1 1 1	ied: 23	TEBS TER evide ene regulatory as 2 2 1 1 1 1 1 1 1 1 1 1 1 2	sociation	B of evidence Train A	c of Ino4-target g	ene regulatory and 1 3 3 2	sociation
Auti Auti The 4 tar Target C CDS1 DPP1 FAS1 OTT (C) ¢ of ge Regulate	Add1 Get genes o Gene 9 Syd YBR YDR YRC YRC YRC Processor YRC YRC YRC YRC YRC YRC YRC YRC	the cooperative stematic Name stematic Name cooc cooc chrl which c coos are validated	TE pair (Abff-inc a of evidence of TEB 2 1 1 1 1 1 1 1 1 1 1 1 1 1	elicit 23 elicitt 23 elicitt 23 elicitt 23 elicitt 23 elicitt	TFB&TFR evide enc regulatory as a 2 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	sociation	B of evidence Tra Constraints	of Ino4-target g	ene regulatory at 1 3 2	sociation
Add Add The 4 tar Target C CDS1 DPP1 PAS1 OTT (C) ¢ of ge Regulato	Add1 Get genes o iene Syd YBR YBR YBR YBR YBR YBR YBR Orres in C ory associat	Actis Actis	TF pair (Abff-inc 5 of evidence of 128 2 1 1 1 1 1 1 1 1 1 1 1 1 1	4 4 11 11 14) validated by 14 15 14 16 14 17 14 18 14 19 14 11 14 11 14 11 14 11 14 11 14 11 14 11 14 11 14 11 14 11 14 11 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 <td>TFBS.TFR evide ene regulatory as 22 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1</td> <td>sociation</td> <td>e die Loget gene</td> <td>of Ino4-target g</td> <td>ene regulatory at 1 1 2 2</td> <td>sociation</td>	TFBS.TFR evide ene regulatory as 22 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	sociation	e die Loget gene	of Ino4-target g	ene regulatory at 1 1 2 2	sociation
Add Add The 4 tar Target G COST CAST	Met31 Adri rget genes o jene 9 Syn YBR YDR YRC YCR nes in C	Artis Artis	TF pair (Abff-inc	A A	r TFBS TFR evide ene regulatory as 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	sociation	for evidence for evidence trans interview interview the target gene	of Ino4-target g	ene regulatory at 1 1 2 2	sociation
Add Add The 4 tar Target G COST DPP1 PP1 PP1 PP1 CC) CC) CC) CC) CC) CC) CC) CC	Met31 Adri rget genes o Jene Syv YBR YDR YDR YCR Ness in C Dry associat	the cooperative see cooperativ	TF pair (Abff-inc a of evidence of THE 2 1 1 1 1 1 1 1 1 1 1 1 1 1	a a 11 11 b4) validated by b4 b5) validated by b5 c b4 c c c c <t< td=""><td>r TFB&TFR evide ene regulatory as 2 2 2 1 1 2 1 2 2 1 1 0 0 0 0 0 0 0 0 0</td><td>sociation</td><td>For evidence of the strengt gene of the str</td><td>of Ino4-target g</td><td>ene regulatory as</td><td>sociation</td></t<>	r TFB&TFR evide ene regulatory as 2 2 2 1 1 2 1 2 2 1 1 0 0 0 0 0 0 0 0 0	sociation	For evidence of the strengt gene of the str	of Ino4-target g	ene regulatory as	sociation
Add Add The 4 tar Target 0 COS1 DPP1 FAS1 GUT (C) ¢ of ge Regulato	Add1 Add1 Seene 9 Sy YOR YOR YOR YOR YOR	the cooperative stematic Name of the cooperative stemator name of the cooperative stemator name of	TE pair (Abrilance d d evidence of TE 2 1 1 1 1 1 1 1 1 1 1 1 1 1	4 1 11 1 14) validated by 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 11 1 11 1 11 1 11 1 11 1 11 1 11 1 11 1 11 1 11 1 11 1 11 1 11 1 12 1 13 1 14 1 15 1 16 1 17 1 18 1 19 1 10 1 10 1 10 1 10 1 <	Image: state of the s	sociation 1725 which regulat	e the target gene	of Inol-target g	ene regulatory as 1375 1376 1 3 2 2	sociation
Add Add The 4 tar Target G COS1 DPP1 FAS1 G(T) For get Regulato	Add1 Add1 Gene So YBR YDR YDR YDR YDR YDR YDR YDR YDR YDR YD	the cooperative stematic Name 1 0220 0220 0220 0220 0220 0220 0220 02	TF pair (Abr1-inc # d evidence of TB 2 1 1 1 1 1 1 1 1 1 1 1 1 1	* * 11 * ************************************	Items Items 1 1 1 1 1 1 2 2 1 1 2 2 1 1 1 1 2 2 1 1 2 1 1 2 1 1 2 2 1 1 2 2 1 1 2 1 2 1 3 1 4 1 5 2 0 1 1 5 2 0 1 1	sociation	e the target gene	of Inol-target g	ene regulatory as 1 1 2 2 2	sociation
Add Add The 4 tar Target G COST DPP1 PAST arr (c) ¢ of ge Regulato	Add1 Add1 Seene 0 Syd YBR YDR YDR YCR YCR	the cooperative seematic Name 1 2220 22400 2240 2240 2240 2240 2240 2	TF pair (Abrf-inc # d evidence of TB 2 1 1 1 1 1 2 2 1 1 1 1 2 2 1 1 1 1 2 2 1 1 1 2 2 1 1 1 1 2 2 1 1 1 2 2 1 1 1 2 2 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2	* * 11 11 ************************************	TEBS TER evide ene regulatory as 2 2 1 1 1 2 2 1 1 1 2 2 1 1 1 2 2 1 1 1 5 6 0 2 2 1 1 1 6 0 2 2 1 1 1 2 2 1 1 1 2 2 1 1 1 1 2 2 1 2 1 1 1 2 2 2 2 1 2 2 1 2 2 1 2 2 2 2 2 1 2	ITEPs which regulat	e the target gene	e e	ene regulatory as 1000	sociation
Add Add The 4 tar Target G COS1 CPP1 PAS1 CIT (C) ¢ of ge Regulate	Add1 Add1 Gene of Syd YBR YDR YDR YRC YRC Nos in CC Ory associat	the cooperative stematic Name stematic Name coop coop coop coop coop coop coop coo	TF pair (Abff-inc d of evidence of TF 2 1 1 1 1 1 1 1 1 1 1 1 1 1	• • • •	TEBS TER evide ene regulatory as TEBS Z Z Z T TEBS TER T	Ince Sociation	B of evidence Train A	of Inol-target g	ene regulatory and 1 1 3 2 2	sociation
Aufi Aufi The 4 tar Target Q COST DPPT ARST OFT (C) (C) Regulate	Met31 Adri Siget genes o Syn YBR YDR YRL YCR Ness in CC Dry associat	the cooperative stematic Name stematic Name cooc cooc cooc cooc cooc cooc cooc co	TE pair (Abff-inc a of evidence of TEB 2 1 1 1 1 1 1 1 1 1 1 1 1 1	4 4 11 11 14) validated by 14 15) validated by 14 16) validated by 14 17) validated by 14 18) validated by 14 18) validated by 14 19) validated by 14 10) validated by 14 11) validated by 14 11) validated by 14 11) validated by 14 11) validated by 14 12) validated by 14 12) validated by 14 12) validated by 14	TEB&TER evidence		 B of evidence TB 4 1 4 2 	of Inol-target g	ene regulatory at 1 3 2 2	sociation
Aufi Aufi The 4 tar COST DPP1 FAST OUT (C) COST COST COST COST COST COST COST COST	Met31 Adr1 get genes o iene yBR yBR yBR yRR yRR yRR yRR yRR yRR yRR	a cose a cooperative a tematic Name a tematic Name a cose b frI which c	TE pair (Abfflanc a of evidence of TER 2 1 1 1 1 1 1 1 1 1 1 1 1 1	4 1 11 1 14) validated by 1 15 1 16 1 17 1 18 1 19 1 10 1 11 1 12 1 12 1 13 1	TFB&TFR evidence		B of evidence Constraints Constant Constraints Constraints Constraints Constraint	of Ino4-target g	ene regulatory at 1 2 2 2	sociation
Auti Auti The 4 tar Target G COS1 DPP1 FAS1 GUT (C) (C) (C) (C) (C) (C) (C) (C) (C) (C)	Met31 Adr1 get genes o person of the second secon	Auto	TE pair (Ab11-inc g & of evidence of TE 2 1 1 1 1 1 1 1 1 1 1 1 1 1	a a 11 11 a 11 b a a a b a a b a a a a a a a a a a a a a a a a b a	TEB&TER evide TEB TER TE TE T T	Ince sociation	Point Control Con	of Ino4-target g	ere regulatory as TTR TTR TTR TTR TTR TTR TTR TTR TTR TT	sociation
Add Add The 4 tar Target 0 COS1 DPP1 PAS1 GUT (C) (C) (C) (C) (C) (C) (C) (C)	Met31 Add1 Get genes o Gene Syv Yor Yor Yor Yor Yor Yor Sene Syv Yor Syv Yor Syv Yor Syv Yor Syv Syv Syv Syv Syv Syv Syv Sy	the cooperative tematic Name tematic Name	TE pair (Abrilance d d avidance of TE 2 1 1 1 1 1 1 1 1 1 1 1 1 1	a a 11 1 a 1 b a b a c b c c c c c c c c c c c c c c c c c c c c	TEB& TER evidence	Ince sociation	f of evidence f	of Inol-target g	Image: state	sociation
Add Add The 4 tar Target G COS1 DPP1 FAS1 G(C) ¢ of ge Regulato COS1 ¢ df ge Regulato COS1 ¢ df ge Regulato	Met31 Add1 Get genes o Gene Syue Yor	a vois a	TF pair (Abr14inc # d evidence of TB 2 1 1 1 1 1 1 1 1 1 1 1 1 1	4 4 11 11 13 11 14 11 15 11 16 11 17 11 18 11 19 11 19 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 12 11 13 14 14 14 14 14 15 14 14 14 15 14 14 14 15 <td< td=""><td>TTERS TER evidence</td><td>ITEPs which regulated and a second seco</td><td>A of evidence Tro Tro</td><td>of Inol-target g</td><td>ene regulatory asolitatory aso</td><td>sociation</td></td<>	TTERS TER evidence	ITEPs which regulated and a second seco	A of evidence Tro	of Inol-target g	ene regulatory asolitatory aso	sociation
Add Add The 4 tar Target G COST DPP1 FAST OCC COST	Met31 Adr1 get genes of Sene 0 Syn YDR YDR <t< td=""><td>Rottis the cooperative seematic Name 0000</td><td>TF pair (AbrH-inc # d evidence of 1 1 1 1 1 1 1 1 1 1 1 1 1</td><td>4 4 11 11 14) validated by 14 15) Validated by 14 16) Validated by 14 17) Validated by 14 18) Validated by 14 19) Validated by 14 10) Validated by 14 11) Validated by 14 11) Validated by 14 11) Validated by 14 11) Validated by 14 12) Validated by 14 12) Validated by 14 13) Validated by 14 14 14 14 14 14 14 15 14 14 14 15 14 14 14 14 14 15 14</td><td>TTER stree evidence TTER evide</td><td>TEPs which regulations in the second se</td><td>B of evidence Constant of the second of the secon</td><td>of Inol-target g</td><td>reputatory association of the second of the</td><td>sociation</td></t<>	Rottis the cooperative seematic Name 0000	TF pair (AbrH-inc # d evidence of 1 1 1 1 1 1 1 1 1 1 1 1 1	4 4 11 11 14) validated by 14 15) Validated by 14 16) Validated by 14 17) Validated by 14 18) Validated by 14 19) Validated by 14 10) Validated by 14 11) Validated by 14 11) Validated by 14 11) Validated by 14 11) Validated by 14 12) Validated by 14 12) Validated by 14 13) Validated by 14 14 14 14 14 14 14 15 14 14 14 15 14 14 14 14 14 15 14	TTER stree evidence TTER evide	TEPs which regulations in the second se	B of evidence Constant of the second of the secon	of Inol-target g	reputatory association of the second of the	sociation
Addi Addi Addi The 4 tar Target G COST DPPT PAST OTT COST COST Regulato Regulato	Met31 Adr1 Spene c Spene c <td>Rottis the cooperative use tops://www.cooperative use use</td> <td>Image: The pair (Abit-lance of the section of the sectiono</td> <td>4 4 11 11 13 11 14 11 15 11 16 11 17 11 18 11 19 11 19 11 10 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 12 11 13 11 14 11</td> <td>TTER evidence TTR evidence</td> <td>TETE which regulate</td> <td></td> <td>e of inol-target g</td> <td>Image: Transmission of the second o</td> <td>sociation</td>	Rottis the cooperative use tops://www.cooperative use	Image: The pair (Abit-lance of the section of the sectiono	4 4 11 11 13 11 14 11 15 11 16 11 17 11 18 11 19 11 19 11 10 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 12 11 13 11 14 11	TTER evidence TTR evidence	TETE which regulate		e of inol-target g	Image: Transmission of the second o	sociation
Add Add The 4 tar Target G COST COST COST COST COST COST COST COST COST COST Cost	Met31 Add1 Gene G Gene G VBR VBR VBR VBR VBR VBR VBR VB	Acting It the cooperative acting	TF pair (Ablf-lanc 2 1 1 1 2 1 1 2 1 1 2 1 1 2 1 1 2 1 1 2 1 2 1 1 2 2 1 2<	4 11 11 11 12 11 13 11 14 11 15 11 16 11 17 11 18 11 19 14 10 14 11 14 12 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14	TTERS TER evidence TTERS TER evidence TTER	ITER which regulation	Image: section of the sectio	e e e e e e e e e e e e e e e e e e e	Image: Transmitter of the second se	sociation

Fig 3. The browse mode. (a) The browse page. (b) The result page of browsing YCRD by the name of a cooperative TF pair. (c) The result page of browsing YCRD by a gene name.

doi:10.1371/journal.pone.0159213.g003

(a)

Search important <u>CTFPs</u> which regulate the inp	put genes			
	Regulatory Associations Validated by	The Input Genes		
• IFB	evidence	YBR009C YNL030W YNL031C YDR24C	-	
• <u>TFR</u>	<u>R evidence</u>	YBR010W YDR225W YDR355C		
○ TFB	3&TFR evidence	YGL093W YOR084W YKL127W	*	
• Bonfe	erroni-corrected p-value threshold: 10 ⁻¹²			
EDR ((False Discovery Rate): 10-2			
	0.00.UT		Example	
	SUBMIT			

(b)

The identified important cooperative TF pairs which regulate the input genes

R	Regulatory associations are defined by TFB&TFR evidence										
X:	X: # of input genes										
Y:	Y: # of target genes of TF1-TF2										
Z:	# of g	enes in	the yeast genome								
	A(2/12)+P1+01 : [2/12 Algorithm Evidence]+[1 Physical Evidence]+[1 Genetic Evidence] A(3/15)+P2 : [5/15 Algorithm Evidence]+[2 Physical Evidence]										
	The Cooperative TF Pair (TF1-TF2) $X \land Y$ $X \land Y$ Z Observed Ratio Expected Ratio Expected Ratio $X \cap Y$ $X \land Y$ Z $X \cap Y$ $X \cap Y$ $X \cap Y$ Y $X \cap Y$ Y $X \cap Y$ Y $X \cap Y$ Y										
		1172									
M	Mbp1 Swi4 A12/17/1924/G10 22 787 23 6572 2.80% 0.35% 8.407e-17										
S	Swi4 Tec1 A(2/12)+P1+G1 25 787 31 6572 3.18% 0.47% 2.360e-15										
м	cm1	Ste12	A(5/15)+P2	28	787	40	6572	3.56%	0.61%	1.379e-14	
S	wi4	Swi6	A(14/17)+P19+G10	19	787	21	6572	2.41%	0.32%	4.254e-13	



doi:10.1371/journal.pone.0159213.g004

However, the identified CTFP Gat1-Gcn4 has no physical or genetic interaction. We suspect that the cooperativity between Gat1 and Gcn4 may act through an intermediate TF Gln3 since both Gat1 and Gcn4 can separately form a protein complex with Gln3 [25].

In the third case study, we would like use a tool in YCRD to identify important CTFPs which regulate a set of ~800 cell cycle genes retrieved from Spellman et al.'s study [31]. So we (i) input a set of ~800 cell cycle genes, (ii) choose TFB&TFR as the experimental evidence of the regulatory association, and (iii) set 10^{-12} as the threshold of the Bonferroni-corrected p-value (see Fig 4A). YCRD then returns four important CTFPs (Mbp1-Swi4, Swi4-Tec1, Mcm1-Ste12 and Swi4-Swi6) which regulate the input set of genes (see Fig 4B). We validate the biological relevance of the four identified CTFPs as follows. First, the two TFs of each CTFP have physical interaction [25], implying that they may cooperatively regulate target genes. Second, all the TFs in the four identified CTFPs are either well known (Mbp1, Mcm1, Swi4, Swi6) [32] or predicted (Ste12 and Tec1) cell cycle TFs [8,33–35], indicating that our tool can identify important cell cycle-related CTFPs which regulate a set of cell cycle genes.

Conclusions

In this article, we present the YCRD which provides 434,197 experimentally validated regulatory associations between 2535 cooperative TF pairs and 6243 target genes. YCRD has an easyto-use interface for biologists to retrieve the target genes of a cooperative TF pair of interest or the cooperative TF pairs which regulate a gene of interest. Moreover, YCRD provides a tool for identifying important cooperative TF pairs which regulate the given set of genes. This is a very useful tool because biologists often have a set of genes of interest (e.g. upregulated genes under a specific biological condition) and would like to know the important cooperative TF pairs which regulate this set of genes. By identifying important cooperative TF pairs, biologists can form hypotheses of the combinatorial regulation of gene expression and have candidates to do further experimental investigation. YCRD will be regularly updated based on the newly published CTFPs identification algorithms and the latest releases of the BioGRID and YEAS-TRACT databases. We believe that the experimentally validated regulatory associations between cooperative TF pairs and their target genes deposited in YCRD will be a very useful resource for yeast biologists to study combinatorial regulation of gene expression.

Supporting Information

S1 File. For each cooperative TF pair, the number of target genes validated by TFB evidence without/with the same experimental condition is provided. (XLSX)

Acknowledgments

The physical/genetic interaction data and a TF's experimentally validated target genes were retrieved from BioGRID and YEASTRACT in March 2016, respectively. We greatly appreciate the effort of both research teams in collecting and curating such valuable data from the literature. This work was supported by National Cheng Kung University and Ministry of Science and Technology of Taiwan [MOST-103-2221-E-006-174-MY2]. Funding for open access charge: National Cheng Kung University and 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.

Author Contributions

Conceived and designed the experiments: WSW. Performed the experiments: YCH FJL. Analyzed the data: WSW YCH FJL. Contributed reagents/materials/analysis tools: YCH FJL. Wrote the paper: WSW YCH.

References

- 1. Tanguay RM. Transcriptional activation of heat-shock genes in eukaryotes. Biochem Cell Biol 1998; 66(6):584-593.
- 2. Nemer G, Nemer M. Regulation of heart development and function through combinatorial interactions of transcription factors. Ann Med 2001; 33(9):604-610. PMID: 11817655
- 3. Lelli KM, Slattery M, Mann RS. Disentangling the many layers of eukaryotic transcriptional regulation. Annu Rev Genet 2012; 46:43–68. doi: 10.1146/annurev-genet-110711-155437 PMID: 22934649
- 4. Teixeira MC, Monteiro PT, Guerreiro JF, Gonçalves JP, Mira NP, dos Santos SC, et al. The YEAS-TRACT database: an upgraded information system for the analysis of gene and genomic transcription regulation in Saccharomyces cerevisiae. Nucleic Acids Res 2014; 42:D161–D166. doi: 10.1093/nar/ gkt1015 PMID: 24170807
- 5. Banerjee N, Zhang MQ. Identifying cooperativity among transcription factors controlling the cell cycle in yeast. Nucleic Acids Res 2003; 31:7024-7031. PMID: 14627835

- Harbison CT, Gordon DB, Lee TI, Rinaldi NJ, Macisaac KD, Danford TW, et al. Transcriptional regulatory code of a eukaryotic genome. Nature 2004; 431:99–104. PMID: <u>15343339</u>
- Nagamine N, Kawada Y, Sakakibara Y. Identifying cooperative transcriptional regulations using protein-protein interactions. Nucleic Acids Res 2005; 33:4828–4837. PMID: <u>16126847</u>
- 8. Tsai HK, Lu HHS, Li WH. Statistical methods for identifying yeast cell cycle transcription factors. Proc Natl Acad Sci USA 2005; 102:13532–13537. PMID: <u>16157877</u>
- Balaji S, Babu MM, Iyer LM, Luscombe NM, Aravind L. Comprehensive analysis of combinatorial regulation using the transcriptional regulatory network of yeast. J Mol Biol 2006; 360(1):213–227. PMID: 16762362
- Chang YH, Wang YC, Chen BS. Identification of transcription factor cooperativity via stochastic system model. Bioinformatics 2006; 22:2276–2282. PMID: <u>16844711</u>
- 11. He D, Zhou D, Zhou Y. Identifying synergistic transcriptional factors involved in the yeast cell cycle using Microarray and ChIP-chip data. In Proceedings of the Fifth International Conference on Grid and Cooperative Computing Workshops (GCCW) 2006; 357–360.
- Wang J. A new framework for identifying combinatorial regulation of transcription factors: a case study of the yeast cell cycle. J Biomedical Informatics 2006; 40:707–725.
- Yu X, Lin J, Masuda T, Esumi N, Zack DJ, Qian J. Genome-wide prediction and characterization of interactions between transcription factors in Saccharomyces cerevisiae. Nucleic Acids Res 2006; 34:917–927. PMID: 16464824
- Elati M, Neuvial P, Bolotin-Fukuhara M, Barillot E, Radvanyi F, Rouveirol C. LICORN: learning cooperative regulation networks from gene expression data. Bioinformatics 2007; 23:2407–2414. PMID: 17720703
- Datta D, Zhao H. Statistical methods to infer cooperative binding among transcription factors in Saccharomyces cerevisiae. Bioinformatics 2008; 24:545–552. PMID: <u>17989095</u>
- Chuang CL, Hung K, Chen CM, Shieh GS. Uncovering transcriptional interactions via an adaptive fuzzy logic approach. BMC Bioinformatics 2009; 10:400. doi: <u>10.1186/1471-2105-10-400</u> PMID: <u>19961622</u>
- Wang Y, Zhang XS, Xia Y. Predicting eukaryotic transcriptional cooperativity by Bayesian network integration of genome-wide data. Nucleic Acids Res 2009; 37:5943–5958. doi: <u>10.1093/nar/gkp625</u> PMID: <u>19661283</u>
- Yang Y, Zhang Z, Li Y, Zhu XG, Liu Q. Identifying cooperative transcription factors by combining ChIPchip data and knockout data. Cell Res 2010; 20:1276–1278. doi: <u>10.1038/cr.2010.146</u> PMID: <u>20975739</u>
- Chen MJ, Chou LC, Hsieh TT, Lee DD, Liu KW, Yu CY, et al. De novo motif discovery facilitates identification of interactions between transcription factors in Saccharomyces cerevisiae. Bioinformatics 2012; 28:701–708. doi: 10.1093/bioinformatics/bts002 PMID: 22238267
- Lai FJ, Jhu MH, Chiu CC, Huang YM, Wu WS. Identifying cooperative transcription factors in yeast using multiple data sources. BMC Systems Biology 2014; 8(Suppl 5):S2. doi: <u>10.1186/1752-0509-8-</u> S5-S2 PMID: 25559499
- Wu WS, Lai FJ. Properly defining the targets of a transcription factor significantly improves the computational identification of cooperative transcription factor pairs in yeast. BMC Genomics 2015; 16(Suppl 12):S10. doi: <u>10.1186/1471-2164-16-S12-S10</u> PMID: <u>26679776</u>
- Lai FJ, Chang HT, Huang YM, Wu WS. A comprehensive performance evaluation on the prediction results of existing cooperative transcription factors identification algorithms. BMC Systems Biology 2014; 8(Suppl 4):S9. doi: <u>10.1186/1752-0509-8-S4-S9</u> PMID: <u>25521604</u>
- Lai FJ, Chang HT, Wu WS. PCTFPeval: a web tool for benchmarking newly developed algorithms for predicting cooperative transcription factor pairs in yeast. BMC Bioinformatics 2015; 16(Suppl 18):S2. doi: 10.1186/1471-2105-16-S18-S2 PMID: 26677932
- Costanzo MC, Engel SR, Wong ED, Lloyd P, Karra K, Chan ET, et al. Saccharomyces Genome Database provides new regulation data. Nucleic Acids Res 2014, 42(1):D717–D725.
- Chatr-Aryamontri A, Breitkreutz BJ, Oughtred R, Boucher L, Heinicke S, Chen D, et al. The BioGRID interaction database: 2015 update. Nucleic Acids Res 2015; 43:D470–D478. doi: <u>10.1093/nar/</u> <u>gku1204</u> PMID: <u>25428363</u>
- Wu WS, Li WH. Identifying gene regulatory modules of heat shock response in yeast. BMC Genomics 2008; 9:439. doi: <u>10.1186/1471-2164-9-439</u> PMID: <u>18811975</u>
- Hofman-Bang J. Nitrogen catabolite repression in Saccharomyces cerevisiae. Mol Biotechnol 1999; 12(1):35–73. PMID: <u>10554772</u>

- Crespo JL, Daicho K, Ushimaru T, Hall MN. The GATA transcription factors GLN3 and GAT1 link TOR to salt stress in Saccharomyces cerevisiae. J Biol Chem 2001; 276(37):34441–34444. PMID: <u>11457832</u>
- 29. Dabas N, Morschhäuser J. Control of ammonium permease expression and filamentous growth by the GATA transcription factors GLN3 and GAT1 in Candida albicans. Eukaryot Cell 2007; 6(5):875–888. PMID: <u>17369441</u>
- Hernández H, Aranda C, Riego L, González A. Gln3-Gcn4 hybrid transcriptional activator determines catabolic and biosynthetic gene expression in the yeast Saccharomyces cerevisiae. Biochem Biophys Res Commun 2011; 404(3):859–864. doi: <u>10.1016/j.bbrc.2010.12.075</u> PMID: <u>21184740</u>
- Spellman PT, Sherlock G, Zhang MQ, Iyer VR, Anders K, Eisen MB, et al. Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. Mol Biol Cell 1998; 9(12):3273–3297. PMID: <u>9843569</u>
- Simon I, Barnett J, Hannett N, Harbison CT, Rinaldi NJ, Volkert TL, et al. Serial regulation of transcriptional regulators in the yeast cell cycle. Cell 2001; 106(6):697–708. PMID: <u>11572776</u>
- Wu WS, Li WH, Chen BS. Computational reconstruction of transcriptional regulatory modules of the yeast cell cycle. BMC Bioinformatics 2006; 7:421. PMID: <u>17010188</u>
- Cheng C, Li LM. Systematic identification of cell cycle regulated transcription factors from microarray time series data. BMC Genomics 2008; 9:116. doi: 10.1186/1471-2164-9-116 PMID: 18315882
- Wu WS, Li WH. Systematic identification of yeast cell cycle transcription factors using multiple data sources. BMC Bioinformatics 2008; 9:522. doi: 10.1186/1471-2105-9-522 PMID: 19061501