Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Knowledge Graph-Based Hypothesis Generation and Experimental Validation in Yeast

Filip Kronström
Chalmers University of Technology and Gothenburg University
Erik Y. Bjurström
erik.bjurstrom51@gmail.com
Chalmers University of Technology
Ievgeniia A. Tiukova
Chalmers University of Technology and Gothenburg University
Ross D. King
Chalmers University of Technology and Gothenburg University

Article

Keywords:

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11469975/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Knowledge Graph-Based Hypothesis Generation and Experimental Validation in Yeast

Filip Kronström[1†], Erik Y. Bjurström[2*†], Ievgeniia A. Tiukova[1,2,3], Ross D. King[1,4]

[1]Department of Computer Science and Engineering, Chalmers University of Technology and Gothenburg University, Sweden.
[2]Department of Life Sciences, Chalmers University of Technology, Sweden.
[3]Department of Industrial Biotechnology, KTH Royal Institute of Technology, Sweden.
[4]Department of Chemical Engineering and Biotechnology, University of Cambridge, United Kingdom.

*Corresponding author(s). E-mail(s): erik.bjurstrom51@gmail.com;
Contributing authors: filip.kronstrom@gmail.com; tiukova@chalmers.se; rossk@chalmers.se;
†These authors contributed equally to this work.

Abstract

Integrating biological knowledge with automated experimentation enables the generation and testing of hypotheses for scientific discovery. However, identifying promising hypotheses from heterogeneous biological data remains challenging. Here, we present a knowledge graph-based approach to prioritising gene–chemical interactions associated with decreased chemical stress resistance in the yeast *Saccharomyces cerevisiae*. We combine information from biological databases and ontologies with a hierarchy-aware graph neural network to rank candidate interactions for experimental testing. We selected 35 gene–chemical hypotheses involving six chemical stressors and evaluated them using automated growth assays. Ten hypotheses were supported after correction for multiple testing, including nine of the 30 highest-ranked candidates for the six target chemicals and one involving a poorly characterised gene. Experimental support varied across the six chemical stressors, highlighting differences in the outcomes of the tested gene–chemical interactions. These findings demonstrate how knowledge graph predictions can be evaluated experimentally to identify gene–chemical

interactions for further investigation. Finally, we extended the Revisions for Improvements of Models in Biology Ontology (RIMBO) to represent knowledge graph updates alongside experimental evidence and provenance, retaining records of both supported and unsupported hypotheses. This approach connects knowledge graph-based prediction, experimental evaluation, and traceable updates to biological knowledge.

Keywords: Knowledge graphs, Graph neural networks, Link prediction, *Saccharomyces cerevisiae*, Chemical stress resistance, Scientific discovery

1 Introduction

Scientific discovery is increasingly driven by the combination of automation, large-scale data integration, and artificial intelligence (AI) [1, 2]. AI methods can support different stages of the scientific process, including hypothesis generation, experimental design, and data interpretation [1]. Advances in laboratory automation enable experiments to be performed at scale, while the integration of automated experimentation with computational decision-making creates opportunities for iterative scientific workflows [3, 4]. In such closed-loop systems, computational models guide experiments, automated laboratory systems perform them, and the resulting data inform subsequent predictions and hypotheses [2, 5–7].

Realising this vision, however, requires computational methods capable of identifying promising hypotheses from the large and heterogeneous body of available biological knowledge. Public biological databases contain vast amounts of information describing genes, proteins, biological processes, metabolic pathways, phenotypes, and chemical compounds, distributed across numerous specialised resources. Integrating this information manually when selecting experiments is increasingly difficult [8], motivating computational methods that can integrate heterogeneous biological knowledge and prioritise candidate hypotheses.

Knowledge graphs (KGs) provide a natural framework for addressing this challenge by integrating heterogeneous biological knowledge originating from diverse databases and ontologies [8]. By representing biological entities and their relationships in a unified graph, KGs enable machine learning models to exploit information distributed across multiple biological domains. Furthermore, ontological class hierarchies provide additional semantic information describing relationships between biological concepts, complementing the relational structure of knowledge graphs. Recent work has demonstrated that incorporating ontological hierarchies into knowledge graph learning can improve predictive performance by enabling models to exploit both relational and semantic information [9]. The ability of biological knowledge graphs to integrate diverse sources of evidence has enabled their use for computational hypothesis generation and scientific discovery [10–13]. Because biological knowledge graphs are inherently incomplete, the absence of a relationship does not necessarily imply that it is false, but rather that it has yet to be observed. Link prediction can therefore be viewed as a way of generating experimentally testable hypotheses by predicting previously

unknown relationships. A wide range of link prediction methods have been developed, including knowledge graph embedding models such as TransE [14] and ComplEx [15], and, more recently, graph neural network-based approaches that exploit both graph structure and semantic information to improve predictive performance [16].

Computational predictions, however, do not by themselves establish new biological knowledge. Experimental testing is therefore an essential step in determining whether predicted relationships are supported by biological observations. Early work demonstrated automated cycles in which computational systems generated hypotheses, designed and performed experiments, and used the results to evaluate those hypotheses [17]. Subsequent work demonstrated iterative model refinement, in which automated experiments were used to generate results that informed updates to systems biology models and subsequent experiments [18]. More recent work has integrated computational hypothesis generation, automated laboratory experimentation, and scientific knowledge representation to validate and refine biological hypotheses [19]. Such iterative processes require representations that preserve how models and knowledge change as new evidence becomes available. The Revisions for Improvements of Models in Biology Ontology (RIMBO) provides a semantic framework for representing revisions to computational models, including the changes and reasons underlying them [20]. Here, we adapt this approach to knowledge graph revisions, linking experimentally evaluated hypotheses to their evidence and resulting changes to the graph.

The yeast *Saccharomyces cerevisiae* is one of the most well-studied eukaryotic model organisms and has long served as a cornerstone of molecular and cellular biology [21]. Decades of experimental research have resulted in well-curated databases, such as the Saccharomyces Genome Database (SGD) [22] and BioCyc [23], which describe genes, proteins, pathways, phenotypes, and other biological entities using standardised ontologies and controlled vocabularies. Together, these resources provide a well-suited foundation for knowledge graph construction and machine learning. Despite the wealth of information, many yeast genes remain incompletely characterised, and the functions and environmental responses of numerous genes are still poorly understood [24]. Identifying these missing relationships therefore remains an important challenge and motivates computational hypothesis generation using biological knowledge graphs.

In this work, we develop a knowledge graph-based approach for generating and experimentally evaluating hypotheses about chemical stress resistance in *S. cerevisiae*. We combine knowledge graph-based prediction, automated experimental validation, and provenance-aware knowledge graph revision.

2 Material and Method

An overview of the closed-loop workflow is shown in Figure 1. The process begins with a knowledge graph integrating information from multiple biological databases and ontologies. A hierarchy-aware graph neural network is used to predict previously unreported gene–chemical interactions associated with decreased chemical stress resistance. We focus on a restricted set of target chemicals, C_{target}, selected based on their well-characterised biological effects, availability, ease of handling, and diversity

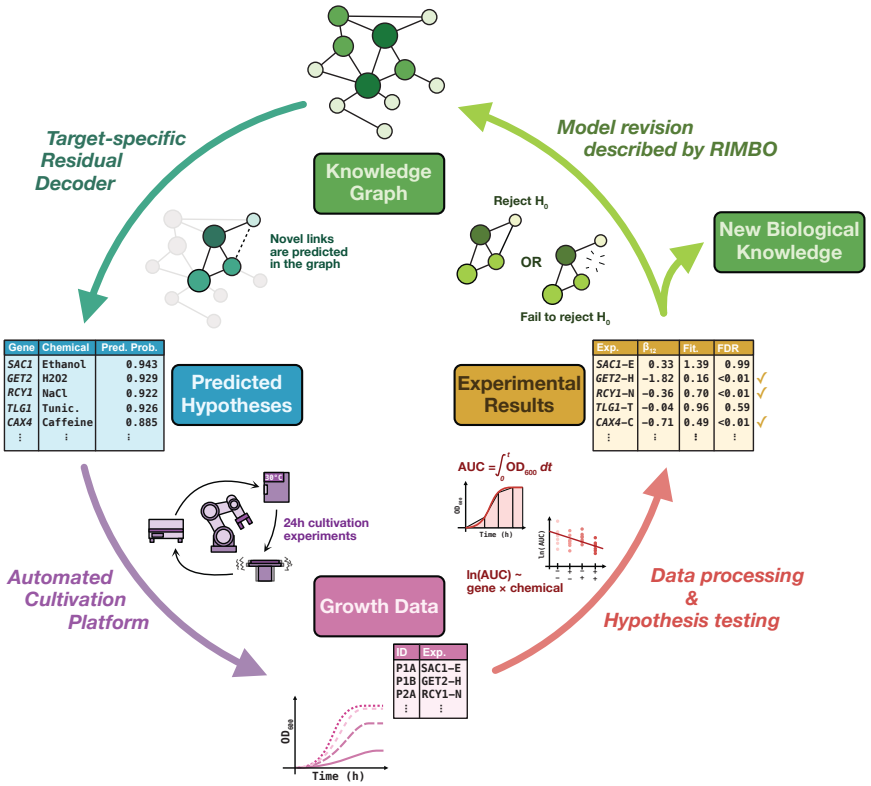


Fig. 1: The workflow presented in this study. Hypotheses are generated from the knowledge graph using a hierarchy-aware graph neural network with a target-specific residual decoder. The predicted hypotheses are experimentally evaluated using an automated cultivation platform, generating growth data. The growth data is then processed, and statistical tests are performed to evaluate the hypotheses. Interactions whose hypotheses were statistically significant are incorporated into the knowledge graph using an extension of RIMBO, thus closing the loop.

of the stresses they induce. The six selected chemical stressors were hydrogen peroxide, tunicamycin, caffeine, sodium chloride (NaCl), ethanol, and acetic acid. The highest-ranked candidate gene–chemical interactions were evaluated using automated growth assays of yeast deletion strains. Experimental measurements were statistically analysed to determine whether the predicted interactions were supported, and supported interactions were subsequently incorporated into the knowledge graph. Both supported and unsupported hypotheses were retained as provenance-aware revision records using an extension of RIMBO.

2.1 The Knowledge Graph

To represent existing biological knowledge about *S. cerevisiae* we use the heterogeneous knowledge graph introduced in our previous work [9], where full details of its construction, schema, and graph statistics are provided. The graph integrates curated information from multiple public resources into a unified representation, including genes, proteins, biological processes, molecular functions, cellular components, phenotypes, chemicals, and other biologically relevant entities. These entities are connected through diverse typed relationships describing, among others, gene functions, biological processes, protein interactions, phenotype associations, and chemical associations. For the six target chemicals in C_{target}, the number of genes with an existing decreased chemical stress resistance association in the knowledge graph is summarised in Table A1, Appendix A.

The knowledge graph was constructed by integrating data from the Saccharomyces Genome Database (SGD) [22] with established biological ontologies and databases, including the Gene Ontology (GO) [25], the Ascomycete Phenotype Ontology (APO) [26], Chemical Entities of Biological Interest (ChEBI) [27], and BioCyc [23], which provides information on metabolic reactions and pathways. Entities are represented using the corresponding ontologies, allowing class hierarchies to be represented explicitly. These hierarchies provide structured biological information that is incorporated into the prediction model through hierarchy-aware semantic losses.

2.2 The prediction model

Our objective is to predict gene–chemical associations corresponding to decreased resistance to chemical stress. Given a heterogeneous knowledge graph describing *S. cerevisiae*, the model estimates a score for missing gene–chemical associations using the broader biological context encoded by the graph. The model combines a heterogeneous graph neural network encoder with a shared decoder and target-specific residual decoders. It builds upon our previously proposed hierarchy-aware GNN framework [16], which incorporates ontology-derived semantic losses into the encoder. An overview of the architecture is shown in Figure 2.

2.2.1 The heterogeneous graph encoder

Node representations are learned using a heterogeneous GraphSAGE [28] encoder operating on the yeast knowledge graph. Separate learnable embeddings are maintained for each node type, while relation-specific GraphSAGE convolutions are learned independently for every edge type. Following each message-passing layer, information from neighbouring nodes is aggregated to produce updated node representations.

The encoder operates on the full heterogeneous knowledge graph, with message passing across the different node and edge types represented in the graph. With two message-passing layers, the representation of a candidate gene or chemical incorporates information from its immediate neighbours and nodes up to two hops away. After the final message-passing layer, only the embeddings of the candidate gene and chemical are extracted and provided to the decoder. Other node types therefore contribute indirectly when they fall within this two-hop neighbourhood.

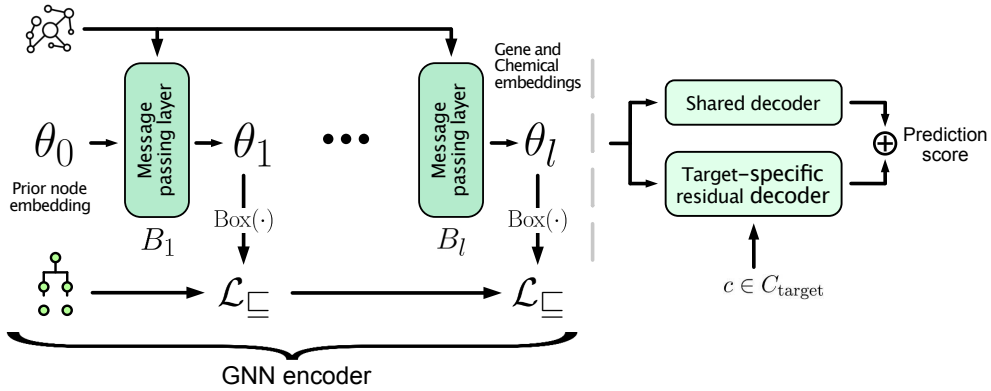


Fig. 2: Overview of the hierarchy-aware graph neural network used for gene–chemical link prediction. Initial node embeddings are updated through heterogeneous message-passing layers over the knowledge graph. Ontological class hierarchies are used to compute semantic losses at each layer, encouraging hierarchy-aware node representations. The resulting gene and chemical embeddings are passed to a shared decoder. For target chemicals $c \in C_{\text{target}}$, a target-specific residual decoder provides an additional prediction component, which is added to the shared decoder output to obtain the final prediction score.

To incorporate information from the ontological hierarchies, the encoder is jointly optimised with the hierarchy-aware semantic losses introduced in our previous work [9, 16]. These ontology-derived losses have previously been shown to improve both phenotype prediction on this knowledge graph and knowledge graph link prediction, particularly when informative ontological hierarchies are available. In this work, the same framework is employed with domain-specific weighting of the semantic losses, placing greater emphasis on the ChEBI hierarchy while omitting the gene hierarchy.

2.2.2 The gene–chemical decoder

After message passing, the learned embeddings of the candidate gene and chemical are extracted and concatenated before being provided to a multilayer perceptron that predicts a score for a decreased chemical stress resistance association. This shared decoder learns general patterns of chemical stress resistance across all genes and chemicals represented in the training data.

The primary objective is to predict and rank gene–chemical edges representing decreased chemical stress resistance for a small set of experimentally selected chemicals, denoted by C_{target}.To account for both shared and chemical-specific patterns of stress resistance, we augment the shared decoder with a residual decoder for each target chemical. These chemical-specific decoders learn corrections to the shared prediction rather than independent predictions from scratch. The final prediction score

is therefore computed as

$$s(g,c) = \begin{cases} d(g,c) + d^{(c)}(g,c), & c \in C_{\text{target}}, \\ d(g,c), & \text{otherwise}, \end{cases} \tag{1}$$

where $d(g,c)$ denotes the shared decoder and $d^{(c)}(g,c)$ denotes the decoder specific to target chemical c. This formulation enables the model to learn representations that generalise across all chemical stress resistance interactions while simultaneously adapting its predictions to the individual characteristics of the experimentally relevant chemicals.

2.3 Training

Known gene–chemical associations exhibiting decreased chemical stress resistance constitute the positive training examples, while negative examples are generated by sampling previously unobserved gene–chemical pairs. The model is trained using binary cross-entropy with logits, with a positive-class weight to account for class imbalance. Negative sampling is biased towards gene corruption, meaning that negative examples are more frequently generated by replacing the gene than the chemical. This reflects the downstream task of ranking candidate genes for a given chemical. The negative sampling strategy is described in Appendix B.

To prioritise the experimental objective while retaining information from the broader knowledge graph, training batches are balanced to include interactions involving the target chemicals in C_{target} alongside other known chemical stress resistance interactions. The GNN is additionally trained using ontology-derived semantic losses to incorporate hierarchical information into the learned representations.

Models were trained using 10-fold cross-validation, with early stopping based on Precision@5 for the target chemicals. Following this stage, models were fine-tuned on interactions involving the target chemicals. Further details of the training procedure, model architecture, and hyperparameters are provided in Appendix B, while the evaluation procedure and metrics are described in Appendix C.

2.4 Prediction of experimental hypotheses

To obtain predictions for experimental validation, candidate interaction scores were computed using an ensemble of the ten models obtained from the cross-validation, described in Section 2.3. For each candidate gene–chemical interaction, the prediction scores produced by the individual models were averaged to obtain a consensus score, which was subsequently used to rank candidate interactions. The corresponding sigmoid-transformed scores were obtained by applying the sigmoid function to the averaged scores.

Two complementary strategies were used to generate hypotheses for experimental validation. First, for each target chemical in C_{target}, all genes present both in the knowledge graph and in the available laboratory deletion collection were considered as candidate genes, yielding a total of 4,914 genes. Genes already known to exhibit decreased resistance to the corresponding chemical were excluded, and the

remaining candidate genes were ranked according to their consensus prediction scores. The five highest-ranked genes for each target chemical were selected for experimental validation.

To investigate the model's ability to identify novel functions for poorly characterised open reading frames (ORFs) whose functions are not well understood, a second analysis was performed using a predefined set of 1,033 genes with limited functional annotation. The predefined set of poorly characterised genes was obtained by selecting all entries in SGD annotated with unknown "function" in the description field. Although these genes were classified as having unknown function, they still have annotations in SGD and links to other entities in the knowledge graph. The SGD annotations used for this selection were downloaded in March 2023. Candidate interactions between these genes and all target chemicals were scored and ranked after excluding previously known associations. To ensure that the selected hypotheses involved five distinct genes, only the highest-ranked chemical interaction for each gene was considered when selecting the top predictions. The five highest-ranked gene–chemical associations involving distinct genes were selected as additional hypotheses for experimental validation.

2.5 Strain selection and cultivation conditions

The *S. cerevisiae* reference strain BY4741 (Background: MATa *his3*Δ*1 leu2*Δ*0 met15*Δ*0 ura3*Δ*0*) and the corresponding single-gene deletion strains were obtained from the EUROSCARF collection.

The strains were revived from -80 °C glycerol stocks by cultivating them in YPD media (2% w/v dextrose) at 30 °C with shaking at 200 rpm. The strains were then streaked on YPD plates and incubated at 30 °C for 72 hours to obtain single colonies.

The preculture and main culture media were prepared using modified Delft media, which contains 5 g L^{-1} $(NH_4)_2SO_4$, 3 g L^{-1} KH_2PO_4, 0.5 g L^{-1} $MgSO_4{\cdot}7H_2O$, 0.79 g L^{-1} Complete Supplement Mixture (Formedium, DCS0019), 100 µg L^{-1} Ampicillin, and 1 mL L^{-1} trace element solution and vitamin solution as described by Veruyn et al. [29]. The dextrose concentration in the main culture media was 0.25% (w/v) while the preculture media was supplemented with 2% (w/v) dextrose. The chemical treatments were prepared by adding the selected chemicals to the main culture media at concentrations determined to be sub-lethal based on literature and preliminary dose-response assays [30–35]. The chemicals were: hydrogen peroxide (Fisher Scientific, 10274050), tunicamycin (Sigma-Aldrich, T7765-1MG), caffeine (Sigma-Aldrich, C0750-100G), acetic acid (Sigma-Aldrich, A6283-2.5L), ethanol (VWR, 20816.298), and sodium chloride (Sigma-Aldrich, S3014-1KG). The concentrations for the chemical treatments in this study are listed in Table 1.

Each gene–chemical pair was evaluated under four conditions: the reference strain with and without chemical treatment, and the corresponding deletion strain with and without chemical treatment. The reference strain without chemical treatment provided the reference condition, while the deletion strain without chemical treatment allowed fitness effects associated with the gene deletion itself to be distinguished from effects of the chemical treatment. Conversely, cultivating the reference strain under

each chemical stressor accounted for reductions in growth caused by the chemical independently of the gene deletion.

Single colonies were picked from the YPD plates and used to inoculate precultures, which were incubated at 30 °C with shaking at 200 rpm overnight. The main cultures were performed in 96-well microtitre plates (Sarstedt, 82.1581 and 82.1584). Prior to cultivation, the microtitre plates and lids were sterilised by UV irradiation for 30 minutes. Each well contained 200 µL of main culture medium. The main cultures were inoculated with the corresponding preculture to achieve an initial optical density at 600 nm (OD_{600}) of 0.025. The plate layouts were designed in a Latin square pattern to avoid systematic bias due to edge effects. Each condition was grown independently in 8 biological replicates. The plates were cultivated using the automated laboratory cell Eve [36].

The cultivation loop was performed as follows: the plate was transferred from an automated incubator (30 °C) to a Teleshaker Magnetic Shaking System, where it was agitated for 30s at 800 rpm (divided evenly between clockwise and counter-clockwise double-orbital shaking), after which the plate was transferred to a BMG Polarstar plate-reader where OD_{600} was measured in each well (30 °C), and finally the plate was transferred back into the automated incubator. The cultivation loop was repeated every 20 minutes for 24 hours, resulting in 72 loops per plate.

2.6 Growth data processing and growth parameter estimation

The growth curves were constructed from the OD_{600} measurements obtained from the plate reader. Blank wells with significant variation (coefficient of variation > 0.2) were removed from the data set. The growth curves were normalised by subtracting the median OD_{600} of the blank wells on the same plate. The area under the growth curve (AUC) was calculated for each well using the trapezoidal rule, providing a summary measure of overall fitness. For descriptive analysis of the growth phenotypes, the maximum biomass-specific growth rate ($\mu_{\max}$) and maximum OD_{600} were also calculated. The maximum biomass-specific growth rate was estimated as the maximum slope of the natural logarithm of the OD_{600} measurements over time, while the maximum OD_{600} was recorded after 24 hours as a measure of carrying capacity for strains that reached stationary phase within the cultivation period.

2.7 Statistical analysis

For each tested gene–chemical pair, growth was quantified as the area under the growth curve (AUC), which was used as a summary measure of overall growth. The AUC values were log-transformed, and outliers were identified and excluded if their log(AUC) values deviated by more than 10 median absolute deviations (MADs) from the group median. We then fitted the following linear model to the filtered log-transformed AUC measurements:

$$\ln(\text{AUC}) \sim \text{genotype} * \text{treatment}, \tag{2}$$

where genotype comprised the wild-type and corresponding deletion strain, and treatment comprised the control and chemical conditions. The model was fitted to all biological replicate measurements for the corresponding experimental conditions. To

account for potential heteroscedasticity, heteroscedasticity-consistent (HC3) standard errors were used for hypothesis testing. The genotype × treatment interaction quantifies whether the relative effect of chemical treatment differs between the deletion strain and wild type. One-sided hypothesis tests were performed for the interaction coefficient, testing whether the interaction was negative, corresponding to a greater reduction in growth under chemical treatment in the deletion strain than in the wild type. p-values were adjusted across all tested gene–chemical pairs using the Benjamini–Hochberg procedure to control the false discovery rate (FDR). Exponentiating the interaction coefficient gives the ratio of the relative treatment effects between the deletion strain and wild type. A ratio below one indicates that the deletion strain retains a smaller fraction of its baseline growth under chemical treatment than the wild type, consistent with the tested hypothesis.

2.8 Extending RIMBO for knowledge graph updates

To represent changes to the knowledge graph resulting from experimental validation, we extended the Revisions for Improvements of Models in Biology Ontology (RIMBO) [20]. RIMBO provides a semantic framework for describing revisions to computational models, including the nature of the changes and their provenance. In our extension, this framework is adapted to knowledge graph updates, allowing changes to individual triples to be represented in a structured and machine-interpretable manner.

The extension introduces representations for triple insertions, deletions, and updates, together with their associated experimental evidence. Supported hypotheses can therefore be represented as revisions that specify both the change made to the knowledge graph and its provenance. In contrast, experimentally unsupported hypotheses are represented as `UnsupportedRevision` instances, which preserve the hypothesis, experimental evidence, and provenance without specifying a knowledge graph change. This distinction allows the framework to preserve the provenance of experimentally evaluated hypotheses regardless of whether they result in an update to the knowledge graph.

3 Results

In this section we first describe the gene–chemical interactions selected for experimental validation, then report the experimental results and characterise the supported interactions. Finally, we describe how the experimental outcomes were represented as knowledge graph revisions.

3.1 Predicted candidate gene–chemical interactions

The model was used to generate candidate gene–chemical interactions for experimental validation. For each target chemical, the five highest-ranked interactions not represented in the knowledge graph were selected, yielding 30 hypotheses. A further five hypotheses involving distinct poorly characterised genes were selected across the target chemicals. In total, 35 gene–chemical interactions were selected for experimental

validation. Table 1 summarises the selected interactions, their sigmoid-transformed prediction scores, and the corresponding gene descriptions from SGD.

Table 1: Top five predicted candidate genes for each target chemical and the five selected hypotheses involving poorly characterised genes. Predictions are ordered by decreasing prediction score. Functional annotation summarises the corresponding gene description in SGD.

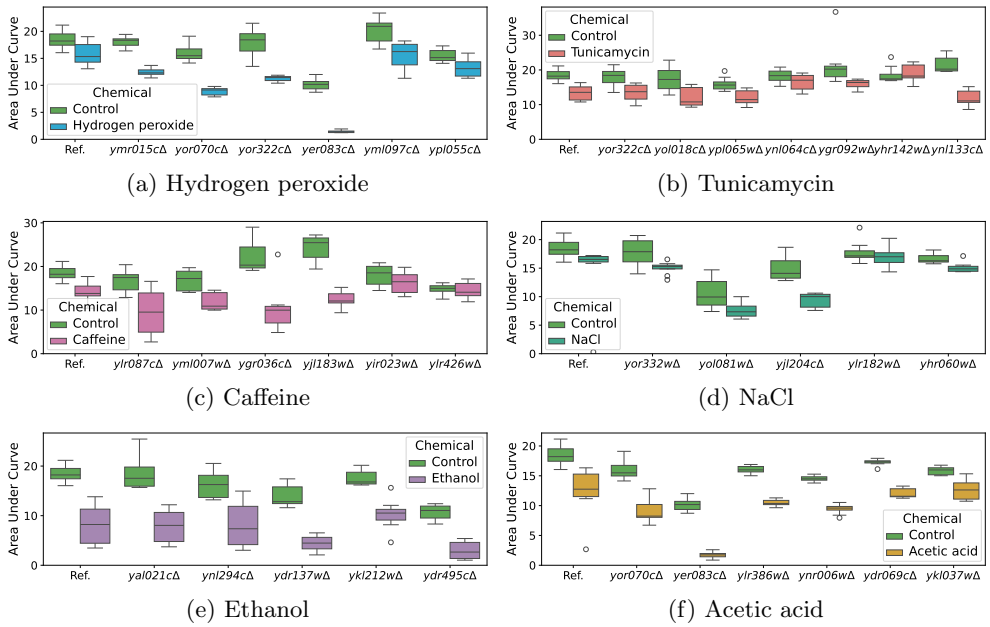
Chemical	Names Systematic / Standard	Score	Functional annotation
Hydrogen peroxide (1.5 mM)	YML097C / VPS9	0.931	Involved in ergosterol biosynthesis
	YER083C / GET2	0.929	Involved in insertion of proteins in to ER membrane
	YOR070C / GYP1	0.924	Cis-golgi GAP for yeast Rabs
	YOR322C / LBD19	0.923	Involved in endocytosis
	YMR015C / ERG5	0.915	Involved in vacuolar transport
Tunicamycin (1.5 µg L^{-1})	YNL064C / YDJ1	0.926	Regulates protein folding
	YPL065W / VPS28	0.923	Involved in endosomal protein sorting
	YOL018C / TLG2	0.920	Subunit of the SNARE complex
	YGR092W / DBF2	0.919	Kinase involved in transcription and stress response
	YOR322C / LBD19	0.918	Involved in endocytosis
Caffeine (2.5 mM)	YML007W / YAP1	0.932	Transcription factor involved in oxidative stress response
	YJL183W / MNN11	0.911	Subunit of Golgi mannosyltransferase complex
	YGR036C / CAX4	0.885	Involved in protein N-glycosylation in ER
	YIR023W / DAL81	0.885	Positive regulator of nitrogen degradation pathways
	YLR087C / CSF1	0.875	Required for fermentation at low temperatures
Acetic acid (30 mM)	YER083C / GET2	0.927	Involved in insertion of proteins in to ER membrane
	YOR070C / GYP1	0.920	Cis-golgi GAP for yeast Rabs
	YLR386W / VAC14	0.915	Involved in vacuolar transport
	YNR006W / VPS27	0.914	Required for recycling Golgi proteins
	YDR069C / DOA4	0.912	Ubiquitin hydrolase
Ethanol (8% v/v)	YKL212W / SAC1	0.943	Regulates sphingolipid biosynthesis
	YDR495C / VPS3	0.939	Required for vacuolar protein sorting and processing
	YAL021C / CCR4	0.938	Involved in regulation of gene expression
	YDR137W / RGP1	0.924	Subunit of the Golgi membrane exchange factor
	YNL294C / MSB3	0.924	Regulates endocytosis
NaCl (0.9 M)	YJL204C / RCY1	0.922	Involved in endocytic recycling
	YOL081W / IRA2	0.919	Regulate cAMP levels under nutrient limitation
	YOR332W / VMA4	0.915	Required for vacuolar H+-ATPase function
	YLR182W / SWI6	0.915	Transcription cofactor involved in meiotic gene expression
	YHR060W / VMA22	0.913	Required for vacuolar H+-ATPase function
Uncharacterised gene group			
Hydrogen peroxide	YPL055C / LGE1	0.887	Null mutants have abnormally large cells
Caffeine	YLR426W / TDA5	0.873	Detected in highly purified mitochondria
Acetic acid	YKL037W / AIM26	0.844	Elevated mitochondrial genome loss for null mutants
Tunicamycin	YHR142W / SPO12	0.813	Positive regulator of mitotic shift
Tunicamycin	YNL133C / FYV6	0.798	Proposed to be involved in pre-mRNA splicing

3.2 Experimental validation

We experimentally evaluated 35 predicted gene–chemical interactions, of which 10 were supported after correction for multiple testing (FDR < 0.05). Of the 30 hypotheses involving the six target chemicals, nine were supported. One of the five additional

hypotheses involving poorly characterised genes was also supported. All 10 supported hypotheses had negative genotype × treatment interaction coefficients, indicating a greater reduction in growth under chemical treatment in the deletion strain than in the wild-type reference.

Figure 3 shows the distributions of AUC measurements for the wild-type reference and deletion strains under control and chemical treatment conditions. Table 2 summarises the supported hypotheses, their estimated interaction effects, and the corresponding statistical significance. Complete results for all 35 hypotheses, including those not supported after multiple-testing correction, are provided in Table D5 in Appendix D.



(a) Hydrogen peroxide (b) Tunicamycin

(c) Caffeine (d) NaCl

(e) Ethanol (f) Acetic acid

Fig. 3: Area under the growth curves (AUC) for the experimentally evaluated gene–chemical interactions. Boxplots show AUC measurements for the wild-type reference and corresponding deletion strains under control and chemical treatment conditions.

3.3 Characterisation of experimentally supported gene–chemical interactions

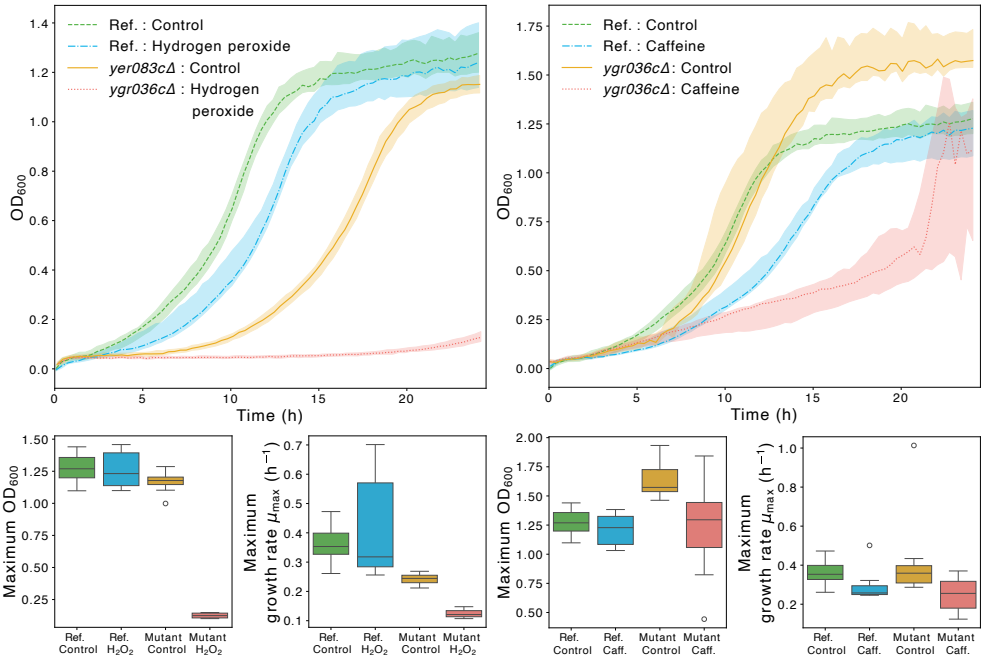
As shown in Table 2, hydrogen peroxide had the highest number of experimentally supported gene–chemical interactions, with four of six interactions supported. Two interactions were supported under each of acetic acid and caffeine stress (2/6), while one was supported under each of NaCl (1/6) and tunicamycin stress (1/7). No interactions were supported under ethanol stress (0/5).

Table 2: Experimentally supported gene–chemical interactions. Rank denotes the prediction rank within the corresponding candidate set. Interaction coefficient is the interaction coefficient β_{12} from the linear model using log-transformed AUC as response. Fitness reduction represents the reduction beyond that expected from the independent effects of gene deletion and chemical stress and is calculated as 1-$e^{\beta_{12}}$. Unadjusted and adjusted p-values are reported as p-values and FDR, respectively. The shaded row corresponds to the supported hypothesis involving a poorly characterised gene.

Chemical	Systematic name	Rank	Interaction coefficient	Fitness reduction	p-value	FDR
Hydrogen peroxide	YER083C	2	-1.797	83.4%	<0.001	<0.001
Acetic acid	YER083C	1	-1.517	78.1%	<0.001	<0.001
Hydrogen peroxide	YOR070C	3	-0.430	34.9%	<0.001	<0.001
NaCl	YJL204C	1	-0.357	30.0%	<0.001	<0.001
Caffeine	YJL183W	2	-0.398	32.9%	<0.001	0.002
Hydrogen peroxide	YOR322C	4	-0.296	25.6%	0.001	0.006
Hydrogen peroxide	YMR015C	5	-0.217	19.5%	0.002	0.010
Caffeine	YGR036C	3	-0.580	44.0%	0.003	0.014
Tunicamycin	YNL133C	5	-0.275	24.0%	0.008	0.031
Acetic acid	YOR070C	2	-0.270	23.7%	0.014	0.049

Hydrogen peroxide is a reactive oxygen species (ROS) and induces oxidative [37]. Excess ROS disturbs the cell's redox homeostasis which causes widespread damage, such as strand breaks and base modifications in DNA, protein misfolding, lipid peroxidation, and mitochondrial dysfunction. The largest fitness reduction observed under any of the tested chemical stresses was for the *YER083C/GET2* deletion strain under hydrogen peroxide treatment, with an 83.4% reduction in fitness. As shown in Figure 4*a*, this strain exhibited almost no growth under hydrogen peroxide stress. *YER083C/GET2* encodes a subunit of the GET complex, which is involved in inserting tail-anchored proteins into the endoplasmic reticulum membrane [38]. Three other hydrogen peroxide interactions were experimentally supported: deletion of *YOR070C/GYP1*, which encodes a GTPase-activating protein for yeast Rab GTPases involved in vesicle docking and fusion, was associated with a 34.9% fitness reduction [39]; deletion of *YOR322C/LBD19*, which encodes an alpha-arrestin involved in regulating the endocytosis of selected plasma membrane-localised amino acid transporters, was associated with a 25.6% reduction [40]; and deletion of *YMR015C/ERG5*, which encodes a C-22 sterol desaturase involved in ergosterol biosynthesis, was associated with a 19.5% reduction [41].

Acetic acid is an organic acid that diffuses into the yeast cells during pH below its pK_a ($\leq$4.75 at 25°C) [33]. While acetic acid is a minor natural co-product during ethanol fermentation in *S. cerevisiae*, excess acetic acid leads to cytosolic acidification and induces oxidative stress. Both gene–chemical interactions supported under acetic acid stress were also supported under hydrogen peroxide stress. The *YER083C/GET2* deletion strain showed a 78.1% reduction in fitness under acetic acid treatment, while the *YOR070C/GYP1* deletion strain showed a 23.7% reduction. The overlap in these



(a) *yer083c*Δ : Hydrogen peroxide (b) *ygr036c*Δ : Caffeine

Fig. 4: Growth data from experiments. The line plot on at the top represent the growth curves. The lines represent median OD_{600} and the fill represent the interquartile range. The box plots in the bottom show empirically derived growth parameters.

responses suggests that oxidative stress may contribute to the effects of acetic acid, although the experiments do not establish the underlying mechanism.

Caffeine functions as a highly specific small-molecule inhibitor of Target of Rapamycin Complex 1 (TORC1) [42]. TORC1 is a crucial nutrient sensing complex that coordinates cell growth, protein synthesis, and metabolism with the availability of nutrient. TORC1 also contributes to aligning cell wall remodelling with nutrient availability. A 44.0% reduction in fitness was observed in the *YGR036C/CAX4* deletant when grown under caffeine stress. As shown in Figure 4*b*, the maximum OD_{600} was greater for the *YGR036C/CAX4* deletion strains compared to the BY4741 reference strain under control conditions, but exhibited an extended lag-phase (15-20 hours), followed by a sudden and unstable increase in OD. *YGR036C/CAX4* encodes a protein indirectly involved in N-linked glycosylation of proteins in the ER, where Cax4p has a luminally oriented active site [43]. The other supported interaction under caffeine stress involved *YJL183W/MNN1* (32.9% reduction) which is a subunit of the Golgi mannosyltransferase complex [44]. Both genes are involved in the biosynthesis of the mannoproteins, contributing to the structural integrity of the cell wall.

Tunicamycin is an antibiotic that inhibits N-linked glycosylation of proteins in the endoplasmic reticulum (ER), leading to ER stress and misfolded protein accumulation [45]. This in turn induces the unfolded protein response (UPR). The

YNL133C/FYV6 deletion strain exhibited a modest, yet statistically significant, 24.0% reduction in fitness under tunicamycin stress. Notably, *YNL133C/FYV6* was relatively poorly characterised, with its function annotated as unknown in SGD at the time of data collection.

Sodium chloride (NaCl) exerts osmotic pressure (hyperosmosis) at elevated concentrations by pulling water across semi-permeable membranes—e.g., cell membranes—to balance the solute concentrations [46]. This causes the cell volume to decrease. As the plasma membrane tension drops, endocytosis is upregulated and endocytic recycling is inhibited to reduce the cell surface area. The *YJL204C/RCY1* deletion strain exhibited a 30.0% reduction in fitness under NaCl stress. *YJL204C/RCY1* is involved in endocytic recycling, where membrane proteins that were internalised by endocytosis are recycled back to the plasma membrane [47].

At elevated concentrations, ethanol induces cellular stress and can cause widespread damage, including disruption of cell membranes, protein denaturation, and oxidative stress [48]. However, none of the five tested gene–ethanol interactions were experimentally supported after correction for multiple testing.

3.4 The knowledge graph revisions

The experimentally evaluated hypotheses were represented as semantic revision records using the RIMBO extension. Supported hypotheses resulted in knowledge graph patches specifying the corresponding gene–chemical relationships, while unsupported hypotheses were retained as revision records without an associated patch. Examples of both cases are provided in Figure F1 in Appendix F.

This representation allows experimentally supported relationships to be incorporated into the knowledge graph while preserving a record of all evaluated hypotheses, including those that were not supported, together with their experimental evidence and provenance.

4 Discussion

In this study, we used a knowledge graph-based model to prioritise gene–chemical interactions associated with decreased chemical stress resistance in *S. cerevisiae* and evaluated 35 hypotheses using automated growth assays. Ten interactions were supported after correction for multiple testing. We also extended RIMBO to represent the experimental evidence and resulting knowledge graph revisions, retaining both supported and unsupported hypotheses with their provenance.

The experimental validation of the KG-derived hypotheses has elucidated novel gene–chemical interactions in *S. cerevisiae*. Here, some of the more notable insights are discussed in further detail. The gene–chemical interaction with the largest decrease in fitness was *YER083C/GET2* under hydrogen peroxide stress. The GET complex enables the insertion of tail-anchored proteins into the ER membrane [49]. Get3p is a cytosolic protein which binds and transport the tail-anchored proteins to the ER membrane, where Get1p (human homologue *WRB*) and Get2p (human hologue *CAMLG*) form a receptor complex that facilitates the insertion of tail-anchored proteins into the ER membrane. The role of Get3p in oxidative stress has been studied previously

and mechanisms of Get3p-mediated stress response have been proposed [50]. Under oxidative stress, the function of Get3p switches from guiding tail-anchored proteins to the ER, to acting as a chaperone that prevents irreversible protein aggregation. While *get2*Δ has been associated with decreased resistance to oxidative stress [38], the extent of this contribution has not been fully investigated. The observations in this study suggest that Get2p could play a more important role in the response to oxidative stress than previously reported, although the mechanism underlying this interaction remains unclear. Dysfunction in the human homologue *CAMLG* has been reported to be linked to congenital disorder of glycosylation where the carrier displays neurological defects such as psychomotor disability, hypotonia, epilepsy and structural brain abnormalities [51].

The *YGR036C/CAX4* (human homologue *DOLPP1*) deletant showed increased biomass yield under control condition while having significantly reduced fitness during caffeine stress, marked by an extended lag-phase. Dolichyl pyrophosphate (Dol-P-P) is a byproduct of the transfer of lipid-linked oligosaccharides to proteins, while dolichyl phosphate (Dol-P) is a substrate for many early reactions in N-linked glycosylation. The availability of Dol-P is often a bottleneck in this process [52]. *YGR036C/CAX4* encodes a phosphatase that converts Dol-P-P into Dol-P, hereby supporting protein glycosylation by recycling a byproduct into a rate-limiting substrate [43]. Protein glycosylation—especially mannoprotein synthesis—is one of the most resource-intensive processes in yeast [53]. The increased maximum OD_{600} observed under the control condition may therefore be attributed to a reduced allocation of chemical and metabolic resources towards mannoprotein production, which may be less necessary for cellular survival under minimal stress. However, under environmental stress the mannoprotein production increases to ensure cell wall integrity [54]. The extended lag-phase in the deletant may therefore be attributed to the the need to synthesise the expensive Dol-P *de novo* in order to sustain the mannoprotein production required to withstand environmental stress. The observed interaction between *YGR036C/CAX4* and caffeine stress therefore suggests that Cax4p may play a role in maintaining cell wall integrity under TORC1 inhibition, induced by caffeine stress. Pathogenic mutations in *DOLPP1* have not been reported to be linked to a specific disorder. However, it is still part of the dolichol phosphate biosynthesis pathway, where significant defects can cause severe disorders with death in early infancy [55].

The *YNL133C/FYV6* deletion strain showed a modest reduction in fitness under tunicamycin stress. *YNL133C/FYV6* is a poorly characterised gene in yeast—whose function has only been inferred from its homology to the human splicing factor *FAM192A* [56]. While mRNA splicing is structurally rare in yeast—only 4-5% of yeast genes contain introns—it is notable that *HAC1*, a transcription factor that regulate the unfolded protein response, is activated when its intron is spliced in response to ER stress, which tunicamycin induces [57, 58]. Therefore, the findings from this study support a hypothesis that Fyv6p may play a minor yet functionally significant role in the activation of *HAC1*. The human homologue *FAM192A* is currently not associated with any confirmed established genetic human diseases. However, a recent study has flagged it as a high-probability potential driver in lung adenocarcinoma [59].

The proposed model achieved higher Precision@5 for the six target chemicals than the ComplEx baseline (0.303 vs. 0.237), despite lower overall MRR (0.221 vs. 0.306). This indicates that overall link prediction performance does not necessarily reflect performance on the target-chemical prioritisation task. Experimental validation provides a complementary assessment: cross-validation measures the recovery of held-out associations within the knowledge graph, whereas the experiments assess predicted interactions against observed phenotypic responses. Nine of the 30 target-chemical hypotheses were supported (30.0%), numerically similar to the cross-validation Precision@5 of 0.303. However, performance varied substantially between chemicals, with cross-validation Precision@5 ranging from 0.12 for NaCl to 0.44 for tunicamycin, while experimental support ranged from 0% for ethanol to 80% for hydrogen peroxide (see Appendix C). This numerical similarity at the aggregate level does not imply correspondence at the individual-chemical level. Moreover, the limited number of experiments prevents a reliable assessment of whether higher-ranked predictions were more likely to be supported.

The prediction scores also did not provide reliable estimates of experimental support. Almost all experimentally evaluated predictions received high sigmoid-transformed scores, despite only 30.0% of the target-chemical hypotheses being experimentally supported. Thus, the scores should not be interpreted as calibrated probabilities of experimental support. Future work could investigate probability calibration methods to better align predicted scores with observed experimental outcomes.

Beyond model performance, the experimental results revealed an overrepresentation of genes encoding membrane-bound proteins, as well as proteins involved in their synthesis, transport, and membrane insertion, among the experimentally supported hypotheses. As shown in Appendix E, these genes were implicated in 9 of the 10 experimentally supported hypotheses (90%) and 19 of the 35 hypotheses tested (54%). Fisher's exact test indicated a significant enrichment ($p = 0.0098$). One possible explanation is that the growth-based experimental procedure was particularly sensitive to perturbations affecting these processes. However, some unsupported hypotheses may also have affected cellular processes in ways that could not be detected through growth profiling alone. Further investigation using transcriptomics and metabolomics could help characterise these effects.

In addition to generating and experimentally evaluating hypotheses, this work represents experimental outcomes explicitly as knowledge graph revisions. The RIMBO extension records the proposed change, the supporting evidence, and whether the change was incorporated into the knowledge graph, while retaining unsupported hypotheses and their provenance rather than discarding them. This representation could support iterative scientific workflows by preserving the outcomes of previous experiments as context for subsequent hypothesis generation.

Several limitations should be considered. First, the study does not establish how the proposed hypothesis-generation strategy compares with alternative approaches. Although the model was compared with ComplEx using cross-validation, hypotheses

generated by alternative methods, model ablations, random selection, or human prioritisation were not experimentally evaluated. Such comparisons would help clarify the contribution of the proposed architecture and computational prioritisation.

Second, the experimental evaluation was relatively small, with five candidates tested per target chemical. Testing more distinct hypotheses would provide stronger evidence about the relationship between prediction rank and experimental support, while additional biological replicates could improve the statistical power of individual tests.

Third, the approach depends on the coverage and quality of the underlying knowledge graph. Although the graph is derived from established biological databases, the source data may contain inconsistencies, conflicting associations, or varying levels of experimental support. Furthermore, representing biological findings as triples necessarily simplifies the underlying evidence, potentially omitting experimental conditions and other contextual information relevant to a gene–chemical association. These limitations may introduce noise into the graph and lead to uncertain or conflicting predictions. In addition, the graph may not capture all gene–chemical relationships already reported in the literature. Predictions should therefore be interpreted in the context of the knowledge represented in the graph, and relationships absent from it should not necessarily be considered novel in the broader biological literature.

Fourth, the cultivation method introduces experimental variability. Although the automated cultivation loop enables multiple experiments to be performed simultaneously, limited control over cultivation conditions allows small variations in temperature and oxygenation to affect AUC, with greater variability under chemical stress (Figure 3). This variability reduces the statistical power of the hypothesis tests. Conducting experiments in more tightly controlled reactors, such as chemostats, could reduce variability and improve the reliability of the resulting conclusions [60].

Finally, the study covers only a single iteration of the proposed closed-loop process. Although supported hypotheses can be incorporated into the knowledge graph, subsequent rounds of hypothesis generation using the updated graph were not evaluated. Repeating the process would allow new experimental evidence to inform subsequent predictions, while retraining the model as evidence accumulates could provide a way to incorporate these updates into its learned representations.

The experimentally supported gene–chemical interactions identified in this study provide a basis for further investigation using additional data modalities, including multi-omics data. In previous work, we demonstrated how hypotheses concerning poorly characterised genes can be translated into quantifiable multi-omic predictions and evaluated using automated laboratory platforms [61]. A similar approach could be used to investigate the experimentally supported interactions identified here using transcriptomics and metabolomics.

Another direction is to extend the approach beyond individual edge additions to more complex hypotheses involving multiple related edge additions or deletions. Such hypotheses may require evaluating a proposed knowledge graph revision as a whole. Previous work explored changes in the graph's embedding space as a potential measure for evaluating revisions [9]. Although preliminary, this approach could help

prioritise revisions according to their impact on the graph representation and consistency with hierarchical constraints. More broadly, combining predicted likelihood with the expected information gained from an experiment could provide a basis for active-learning strategies that prioritise hypotheses both for their likelihood of being supported and their potential to improve the knowledge graph.

5 Conclusions

In this study, we used a knowledge graph and a hierarchy-aware graph neural network to identify promising gene–chemical interactions in the yeast *Saccharomyces cerevisiae*. Experimental evaluation of 35 hypotheses identified 10 interactions associated with decreased chemical stress resistance, including one involving a poorly characterised gene. These findings illustrate how knowledge graph predictions can guide experiments to identify biologically relevant interactions. We also extended RIMBO, an ontology for representing computational model revisions, to capture knowledge graph updates alongside supporting experimental evidence and provenance, providing a foundation for iterative scientific discovery. Future work could integrate multi-omics data to further investigate the supported interactions.

Acknowledgements. The computations were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS), partially funded by the Swedish Research Council through grant agreement no. 2022-06725. This work was supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation, the UK Engineering and Physical Sciences Research Council (EPSRC) [EP/R022925/2, EP/W004801/1 and EP/X032418/1], the Chalmers AI Research Centre, and Swedish Research Council Formas [2020-01690].

Declarations

Author contributions

Conceptualisation: F.K., E.Y.B., I.A.T, R.D.K. **Data curation:** F.K., E.Y.B. **Formal analysis:** F.K., E.Y.B. **Investigation:** F.K., E.Y.B. **Methodology:** F.K., E.Y.B. **Supervision:** I.A.T., R.D.K. **Visualization:** F.K., E.Y.B. **Writing — original draft:** F.K., E.Y.B. **Writing — review & editing:** I.A.T., R.D.K.

Data availability

The code and data can be found at https://github.com/filipkro/yeast-kg-discovery

Competing interests

The authors declare no competing interests.

Funding

Open access funding was provided by Chalmers University of Technology.

References

[1] Wang, H. *et al.* Scientific discovery in the age of artificial intelligence. *Nature* **620**, 47–60 (2023).

[2] King, R. D. *et al.* The Automation of Science. *Science* **324**, 85–89 (2009).

[3] Abolhasani, M. & Kumacheva, E. The rise of self-driving labs in chemical and materials sciences. *Nature Synthesis* **2**, 483–492 (2023).

[4] Tom, G. *et al.* Self-Driving Laboratories for chemistry and materials science. *Chemical Reviews* **124**, 9633–9732 (2024).

[5] Sparkes, A. *et al.* Towards Robot Scientists for autonomous scientific discovery. *Automated Experimentation* **2**, 1 (2010).

[6] Gower, A. H. *et al.* The use of AI-robotic systems for scientific discovery (2026). 2406.17835 [cs.LG].

[7] Zenil, H. *et al.* The future of fundamental science led by generative closed-loop artificial intelligence. *Frontiers in Artificial Intelligence* **9** (2026).

[8] Nicholson, D. N. & Greene, C. S. Constructing knowledge graphs and their biomedical applications. *Computational and Structural Biotechnology Journal* **18**, 1414–1428 (2020).

[9] Kronström, F., Gower, A. H., Brunnsåker, D., Tiukova, I. A. & King, R. D. Graph neural network-based hierarchy-aware embeddings of knowledge graphs: Applications to yeast phenotype prediction. *Neurosymbolic Artificial Intelligence* **2** (2026). URL https://doi.org/10.1177/29498732261469314.

[10] Sosa, D. N. & Altman, R. B. Contexts and contradictions: a roadmap for computational drug repurposing with knowledge inference. *Briefings in Bioinformatics* **23**, bbac268 (2022). URL https://doi.org/10.1093/bib/bbac268.

[11] Himmelstein, D. S. *et al.* Systematic integration of biomedical knowledge prioritizes drugs for repurposing. *eLife* **6**, e26726 (2017).

[12] Sudhahar, S. *et al.* An experimentally validated approach to automated biological evidence generation in drug discovery using knowledge graphs. *Nature Communications* **15**, 5703 (2024). URL https://www.nature.com/articles/s41467-024-50024-6.

[13] Kronström, F., Brunnsåker, D., Tiukova, I. A. & King, R. D. Gilpin, L. H., Giunchiglia, E., Hitzler, P. & van Krieken, E. (eds) *Ontology-based box embeddings and knowledge graphs for predicting phenotypic traits in Saccharomyces cerevisiae.* (eds Gilpin, L. H., Giunchiglia, E., Hitzler, P. & van Krieken, E.) *Proceedings of The 19th International Conference on Neurosymbolic Learning and*

Reasoning, 891–912 (PMLR, 2025). URL https://proceedings.mlr.press/v284/kronstrom25a.html.

[14] Bordes, A., Usunier, N., Garcia-Durán, A., Weston, J. & Yakhnenko, O. Burges, C. J. C., Bottou, L., Welling, M., Ghahramani, Z. & Weinberger, K. Q. (eds) *Translating embeddings for modeling multi-relational data.* (eds Burges, C. J. C., Bottou, L., Welling, M., Ghahramani, Z. & Weinberger, K. Q.) *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*, Vol. 2 of *NIPS'13*, 2787–2795 (Curran Associates Inc., Red Hook, NY, USA, 2013).

[15] Trouillon, T., Welbl, J., Riedel, S., Gaussier, E. & Bouchard, G. Balcan, M. F. & Weinberger, K. Q. (eds) *Complex Embeddings for Simple Link Prediction.* (eds Balcan, M. F. & Weinberger, K. Q.) *Proceedings of The 33rd International Conference on Machine Learning*, 2071–2080 (PMLR, 2016). URL https://proceedings.mlr.press/v48/trouillon16.html.

[16] Kronström, F. & King, R. D. Pesquita, C., Mileo, A., Shimizu, C. & Passerini, A. (eds) *Hierarchy-aware semantic losses for knowledge graph link prediction.* (eds Pesquita, C., Mileo, A., Shimizu, C. & Passerini, A.) *Proceedings of The 20th International Conference on Neurosymbolic Learning and Reasoning* (PMLR, 2026). URL https://openreview.net/pdf?id=mRfGldY4SX.

[17] King, R. D. *et al.* Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature* **427**, 247–252 (2004). URL https://www.nature.com/articles/nature02236.

[18] Coutant, A. *et al.* Closed-loop cycles of experiment design, execution, and learning accelerate systems biology model development in yeast. *Proceedings of the National Academy of Sciences* **116**, 18142–18147 (2019). URL https://www.pnas.org/doi/10.1073/pnas.1900548116.

[19] Brunnsåker, D. *et al.* Agentic AI integrated with scientific knowledge: laboratory validation in systems biology. *Journal of The Royal Society Interface* **23**, 20260043 (2026). URL https://doi.org/10.1098/rsif.2026.0043.

[20] Kronström, F., Gower, A. H., Tiukova, I. A. & King, R. D. Bifet, A., Lorena, A. C., Ribeiro, R. P., Gama, J. & Abreu, P. H. (eds) *RIMBO - an ontology for model revision databases.* (eds Bifet, A., Lorena, A. C., Ribeiro, R. P., Gama, J. & Abreu, P. H.) *Discovery Science*, Lecture Notes in Computer Science, 523–534 (Springer Nature Switzerland, 2023).

[21] Botstein, D., Chervitz, S. A. & Cherry, J. M. Yeast as a model organism. *Science* **277**, 1259–1260 (1997).

[22] Engel, S. R. *et al.* Saccharomyces genome database: Advances in genome annotation, expanded biochemical pathways, and other key enhancements. *Genetics*

iyae185 (2024).

[23] Karp, P. D. *et al.* The BioCyc collection of microbial genomes and metabolic pathways. *Briefings in Bioinformatics* **20**, 1085–1093 (2019).

[24] Wood, V. *et al.* Hidden in plain sight: what remains to be discovered in the eukaryotic proteome? *Open Biology* **9**, 180241 (2019).

[25] Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. *Nature Genetics* **25**, 25–29 (2000). Number: 1.

[26] Costanzo, M. C. *et al.* New mutant phenotype data curation system in the saccharomyces genome database. *Database: The Journal of Biological Databases and Curation* **2009**, bap001 (2009).

[27] Hastings, J. *et al.* ChEBI in 2016: Improved services and an expanding collection of metabolites. *Nucleic acids research* **44**, D1214–9 (2016).

[28] Hamilton, W. L., Ying, R. & Leskovec, J. von Luxburg, U., Guyon, I., Bengio, S., Wallach, H. & Fergus, R. (eds) *Inductive representation learning on large graphs.* (eds von Luxburg, U., Guyon, I., Bengio, S., Wallach, H. & Fergus, R.) *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, 1025–1035 (Curran Associates Inc., 2017).

[29] Verduyn, C., Postma, E., Scheffers, W. A. & van Dijken, J. P. Effect of benzoic acid on metabolic fluxes in yeasts: A continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast* **8**, 501–517 (1992).

[30] Hou, J. *et al.* Management of the endoplasmic reticulum stress by activation of the heat shock response in yeast. *FEMS Yeast Research* **14**, 481–494 (2014). URL https://academic.oup.com/femsyr/article-lookup/doi/10.1111/1567-1364.12125.

[31] Ji, X.-X. *et al.* Response mechanism of ethanol-tolerant Saccharomyces cerevisiae strain ES-42 to increased ethanol during continuous ethanol fermentation. *Microbial Cell Factories* **24**, 33 (2025). URL https://microbialcellfactories.biomedcentral.com/articles/10.1186/s12934-025-02663-7.

[32] Kuranda, K., Leberre, V., Sokol, S., Palamarczyk, G. & François, J. Investigating the caffeine effects in the yeast *Saccharomyces cerevisiae* brings new insights into the connection between TOR, PKC and Ras/cAMP signalling pathways. *Molecular Microbiology* **61**, 1147–1166 (2006). URL https://onlinelibrary.wiley.com/doi/10.1111/j.1365-2958.2006.05300.x.

[33] Ribeiro, R. A. *et al.* Yeast adaptive response to acetic acid stress involves structural alterations and increased stiffness of the cell wall. *Scientific Reports* **11**, 12652 (2021). URL https://www.nature.com/articles/s41598-021-92069-3.

[34] Martins, D. & English, A. M. Catalase activity is stimulated by H2O2 in rich culture medium and is required for H2O2 resistance and adaptation in yeast. *Redox Biology* **2**, 308–313 (2014). URL https://linkinghub.elsevier.com/retrieve/pii/S2213231714000020.

[35] Olz, R., Larsson, K., Adler, L. & Gustafsson, L. Energy flux and osmoregulation of Saccharomyces cerevisiae grown in chemostats under NaCl stress. *Journal of Bacteriology* **175**, 2205–2213 (1993). URL https://journals.asm.org/doi/10.1128/jb.175.8.2205-2213.1993.

[36] Williams, K. *et al.* Cheaper faster drug development validated by the repositioning of drugs against neglected tropical diseases. *Journal of The Royal Society Interface* **12**, 20141289 (2015).

[37] De Almeida, A. J. P. O. *et al.* ROS: Basic Concepts, Sources, Cellular Signaling, and its Implications in Aging Pathways. *Oxidative Medicine and Cellular Longevity* **2022**, 1225578 (2022). URL https://onlinelibrary.wiley.com/doi/10.1155/2022/1225578.

[38] Helsen, J. *et al.* Gene Loss Predictably Drives Evolutionary Adaptation. *Molecular Biology and Evolution* **37**, 2989–3002 (2020). URL https://academic.oup.com/mbe/article/37/10/2989/5870834.

[39] Du, L. L., Collins, R. N. & Novick, P. J. Identification of a Sec4p GTPase-activating protein (GAP) as a novel member of a Rab GAP family. *The Journal of Biological Chemistry* **273**, 3253–3256 (1998).

[40] Lin, C. H., MacGurn, J. A., Chu, T., Stefan, C. J. & Emr, S. D. Arrestin-related ubiquitin-ligase adaptors regulate endocytosis and protein turnover at the cell surface. *Cell* **135**, 714–725 (2008).

[41] Skaggs, B. A. *et al.* Cloning and characterization of the Saccharomyces cerevisiae C-22 sterol desaturase gene, encoding a second cytochrome P-450 involved in ergosterol biosynthesis. *Gene* **169**, 105–109 (1996).

[42] Rallis, C., Codlin, S. & Bähler, J. TORC1 signaling inhibition by rapamycin and caffeine affect lifespan, global gene expression, and cell proliferation of fission yeast. *Aging Cell* **12**, 563–573 (2013).

[43] Fernandez, F. *et al.* The CWH8 gene encodes a dolichyl pyrophosphate phosphatase with a luminally oriented active site in the endoplasmic reticulum of Saccharomyces cerevisiae. *The Journal of Biological Chemistry* **276**, 41455–41464 (2001).

[44] Jungmann, J. & Munro, S. Multi-protein complexes in the cis Golgi of Saccharomyces cerevisiae with alpha-1,6-mannosyltransferase activity. *The EMBO journal* **17**, 423–434 (1998).

[45] Hernández-Elvira, M. *et al.* Tunicamycin Sensitivity-Suppression by High Gene Dosage Reveals New Functions of the Yeast Hog1 MAP Kinase. *Cells* **8**, 710 (2019).

[46] Szopinska, A., Degand, H., Hochstenbach, J.-F., Nader, J. & Morsomme, P. Rapid Response of the Yeast Plasma Membrane Proteome to Salt Stress. *Molecular & Cellular Proteomics* **10**, M111.009589 (2011). URL https://linkinghub.elsevier.com/retrieve/pii/S153594762030102X.

[47] Wiederkehr, A., Avaro, S., Prescianotto-Baschong, C., Haguenauer-Tsapis, R. & Riezman, H. The F-box protein Rcy1p is involved in endocytic membrane traffic and recycling out of an early endosome in Saccharomyces cerevisiae. *The Journal of Cell Biology* **149**, 397–410 (2000).

[48] Sahana, G. R. *et al.* A review on ethanol tolerance mechanisms in yeast: Current knowledge in biotechnological applications and future directions. *Process Biochemistry* **138**, 1–13 (2024). URL https://linkinghub.elsevier.com/retrieve/pii/S1359511323004245.

[49] Schuldiner, M. *et al.* The GET Complex Mediates Insertion of Tail-Anchored Proteins into the ER Membrane. *Cell* **134**, 634–645 (2008). URL https://linkinghub.elsevier.com/retrieve/pii/S0092867408007770.

[50] Ulrich, K. *et al.* From guide to guard-activation mechanism of the stress-sensing chaperone Get3. *Molecular Cell* **82**, 3226–3238.e7 (2022).

[51] Wilson, M. P. *et al.* CAMLG-CDG: a novel congenital disorder of glycosylation linked to defective membrane trafficking. *Human Molecular Genetics* **31**, 2571–2581 (2022). URL https://academic.oup.com/hmg/article/31/15/2571/6545759.

[52] Schenk, B., Fernandez, F. & Waechter, C. J. The ins(ide) and outs(ide) of dolichyl phosphate biosynthesis and recycling in the endoplasmic reticulum. *Glycobiology* **11**, 61R–70R (2001). URL https://academic.oup.com/glycob/article-lookup/doi/10.1093/glycob/11.5.61R.

[53] Tanner, W. & Lehle, L. Protein glycosylation in yeast. *Biochimica Et Biophysica Acta* **906**, 81–99 (1987).

[54] Ribeiro, R. A., Bourbon-Melo, N. & Sá-Correia, I. The cell wall and the response and tolerance to stresses of biotechnological relevance in yeasts. *Frontiers in Microbiology* **13**, 953479 (2022). URL https://www.frontiersin.org/articles/10.3389/fmicb.2022.953479/full.

[55] Kranz, C. *et al.* A Defect in Dolichol Phosphate Biosynthesis Causes a New Inherited Disorder with Death in Early Infancy. *The American Journal of Human Genetics* **80**, 433–440 (2007). URL https://linkinghub.elsevier.com/retrieve/pii/S0002929707600925.

[56] Zhan, X., Lu, Y., Zhang, X., Yan, C. & Shi, Y. Mechanism of exon ligation by human spliceosome. *Molecular Cell* **82**, 2769–2778.e4 (2022).

[57] Kalra, S. & Coolon, J. D. Decoding RAP1's Role in Yeast mRNA Splicing (2025). ISSN: 2692-8205.

[58] Sidrauski, C. & Walter, P. The transmembrane kinase Ire1p is a site-specific endonuclease that initiates mRNA splicing in the unfolded protein response. *Cell* **90**, 1031–1039 (1997).

[59] Yuan, F. *et al.* Integrative multi-omics machine learning reveals novel driver genes associations in lung adenocarcinoma. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1874**, 141113 (2026). URL https://linkinghub.elsevier.com/retrieve/pii/S1570963925000512.

[60] Tiukova, I. A. *et al.* Genesis: Towards the automation of systems biology research (2024). URL https://arxiv.org/abs/2408.10689. ArXiv: 2408.10689 [cs.AI].

[61] Bjurström, E. Y. *et al.* Investigating uncharacterised genes in Saccharomyces cerevisiae using robot scientists. *Scientific Reports* **16**, 10999 (2026). URL https://www.nature.com/articles/s41598-026-46236-z.

Appendix A Genes associated with C_{target}

Table A1: Number of genes annotated with 'Decreased stress resistance to chemical' in the knowledge graph, to the chemicals in C_{target}.

Chemical	Hydrogen peroxide	Tunicamycin	Caffeine	NaCl	Ethanol	Acetic acid
Number of linked genes	815	591	775	589	912	273

Appendix B Model Implementation and Training Details

The heterogeneous GraphSAGE encoder consists of two message-passing layers. Initial node embeddings have dimension 64 for all node types, and both layers produce 128-dimensional node representations. For each edge type, GraphSAGE convolutions use max aggregation, with learnable root transformations, input feature projection, and L2 normalisation of the output representations. Outputs from different edge-type convolutions are combined using mean aggregation.

The shared decoder is a multilayer perceptron (MLP) with layer dimensions 256-32-8-4-1, where the input comprises the concatenated 128-dimensional gene and chemical embeddings. Each chemical-specific residual decoder takes the same input and has layer dimensions 256-8-4-1. ReLU activation is applied after each hidden layer, while the output layers have no activation function.

The ComplEx baseline was trained on a flattened version of the knowledge graph, in which nodes from all domains were represented in a single graph. The embedding dimension was 64. The GNN model and ComplEx baseline contained 14,470,007 and 16,217,856 trainable parameters, respectively (Table C2).

Both models were trained using the AdamW optimiser, with an initial learning rate of 10^{-2} and L2 regularisation of 10^{-3}. The learning rate was adjusted using cosine annealing, with a minimum learning rate of 10^{-5} and $T_{\max} = 500$ epochs. Binary cross-entropy with logits was used as the training objective, with a positive-class weight of 400. Models were trained for a maximum of 500 epochs, with early stopping based on Precision@5 across the five target chemicals and a patience of 50 epochs. Following early stopping, each model was fine-tuned for 20 epochs using only interactions involving the target chemicals, with a learning rate of 5×10^{-5}.

For each training batch, 90% of the training edges were used for message passing and the remaining 10% for supervision. Batches were constructed so that every training edge was used as a supervision edge once per epoch. To prioritise the experimental objective while retaining general predictive performance, 40% of positive supervision examples were sampled from interactions involving the target chemicals. For each positive supervision edge, 600 negative examples were sampled: 10% by corrupting

both the head and tail entities, 10% by sampling random tails for the true heads, and 80% by sampling random heads for the true tails.

For the GNN, the ontology-derived semantic losses were assigned a base weight of 0.01, with the losses associated with the ChEBI domain weighted ten times more heavily than those associated with other domains. The semantic loss formulation follows our previous work [16].

Appendix C Model evaluation

Model performance was evaluated using 10-fold cross-validation, where known decreased chemical stress resistance associations were randomly partitioned into training and validation sets. Models were trained as described in Appendix B. Performance was assessed using two complementary evaluation measures.

Overall link prediction performance was measured using mean reciprocal rank (MRR). For each positive validation interaction, rankings were computed against 1000 sampled negative gene–chemical pairs, consisting of equal numbers of corrupted genes and corrupted chemicals while respecting the corresponding entity domains.

To assess the model's ability to prioritise candidate genes for the target chemicals, Precision@5 was computed separately for each target chemical. For each chemical, genes without a known decreased chemical stress resistance association were ranked according to their predicted scores, together with held-out positive interactions. Precision@5 was calculated as the fraction of the five highest-ranked genes that corresponded to held-out positive interactions.

The performance of the proposed architecture was compared with a baseline knowledge graph embedding model based on ComplEx [15]. The baseline was trained and evaluated using the same training and validation folds, allowing direct comparison with the proposed model.

Table C2 summarises the cross-validation performance of the proposed model and the ComplEx baseline. The proposed model achieved a higher average Precision@5 across the six target chemicals (0.3033 vs. 0.2367), while ComplEx achieved a higher overall MRR (0.3058 vs. 0.2212). Thus, although the proposed model did not achieve the highest overall link prediction performance, it performed better when evaluated according to the target-chemical prioritisation task. Table C3 shows the average Precision@5 for each of the six target chemicals across the cross-validation folds.

Table C2: Cross-validation performance of the proposed model and the ComplEx baseline. Results are reported as mean ± standard deviation over the ten cross-validation folds.

Model	Trainable parameters	Target chemicals Precision@5	Overall MRR
ComplEx	16,217,856	0.2367 ± 0.0233	**0.3058 ± 0.0273**
Proposed GNN	14,470,007	**0.3033 ± 0.0433**	0.2212 ± 0.0159

Table C3: Average Precision@5 from the 10 cross validation folds within each target chemical group.

Hydrogen peroxide	Tunicamycin	Caffeine	NaCl	Ethanol	Acetic acid
0.22	0.44	0.36	0.12	0.42	0.24

Table C4: Hit-rate among the experimentally tested hypotheses.

Hydrogen peroxide	Tunicamycin	Caffeine	NaCl	Ethanol	Acetic acid	Poorly characterised
0.8	0.0	0.4	0.2	0.0	0.4	0.2

Appendix D Full experimental results

Table D5: All experimentally evaluated gene–chemical interactions. Rank denotes the prediction rank within the corresponding candidate set and among all genes considered by the unrestricted prediction (candidate / all). Interaction coefficient is the interaction coefficient β_{12} from the linear model using log-transformed AUC as response. Interaction$_{95\%}$ bound is the one-sided 95% upper confidence bound for the interaction coefficient. Fitness reduction represents the reduction beyond that expected from the independent effects of gene deletion and chemical stress and is calculated as 1-$e^{\beta_{12}}$. Unadjusted and adjusted p-values are reported as p-values and FDR, respectively. A ✓ indicates an experimentally supported hypothesis (FDR < 0.05). Shaded rows correspond to hypotheses involving poorly characterised genes.

Chemical	Gene	Rank candidate/all	Interaction coefficient	Interaction$_{95\%}$ bound	Fitness reduction	p-value	FDR
Hydrogen peroxide	YER083C ✓	2 / 2	-1.797	-1.636	83.4%	<0.001	<0.001
Acetic acid	YER083C ✓	1 / 3	-1.517	-1.254	78.1%	<0.001	<0.001
Hydrogen peroxide	YOR070C ✓	3 / 3	-0.430	-0.291	34.9%	<0.001	<0.001
NaCl	YJL204C ✓	1 / 2	-0.357	-0.215	30.0%	<0.001	<0.001
Caffeine	YJL183W ✓	2 / 3	-0.398	-0.226	32.9%	<0.001	0.002
Hydrogen peroxide	YOR322C ✓	4 / 4	-0.296	-0.147	25.6%	0.001	0.006
Hydrogen peroxide	YMR015C ✓	5 / 5	-0.217	-0.100	19.5%	0.002	0.010
Caffeine	YGR036C ✓	3 / 5	-0.580	-0.244	44.0%	0.003	0.014
Tunicamycin	YNL133C ✓	5 / 121	-0.275	-0.093	24.0%	0.008	0.031
Acetic acid	YOR070C ✓	2 / 5	-0.270	-0.072	23.7%	0.014	0.049
NaCl	YOL081W	2 / 4	-0.208	-0.000	18.8%	0.050	0.148
Caffeine	YLR087C	5 / 8	-0.457	0.009	36.7%	0.053	0.148
Acetic acid	YNR006W	4 / 9	-0.133	0.005	12.4%	0.057	0.148
Ethanol	YDR495C	2 / 2	-0.537	0.030	41.5%	0.059	0.148
Acetic acid	YLR386W	3 / 8	-0.119	0.011	11.2%	0.065	0.152
Ethanol	YDR137W	4 / 5	-0.296	0.148	25.6%	0.133	0.288
Hydrogen peroxide	YML097C	1 / 1	-0.111	0.061	10.5%	0.140	0.288
NaCl	YOR332W	3 / 7	-0.060	0.062	5.8%	0.205	0.399
Caffeine	YML007W	1 / 2	-0.083	0.100	8.0%	0.223	0.412
Acetic acid	YDR069C	5 / 12	-0.055	0.078	5.3%	0.244	0.427
Tunicamycin	YOL018C	3 / 4	-0.044	0.182	4.3%	0.371	0.619
Ethanol	YAL021C	3 / 3	-0.010	0.455	1.0%	0.486	0.754
Hydrogen peroxide	YPL055C	1 / 23	-0.001	0.144	0.1%	0.496	0.754
NaCl	YHR060W	5 / 9	0.005	0.084	-0.5%	0.540	0.788
Ethanol	YNL294C	5 / 6	0.071	0.618	-7.4%	0.587	0.800
Tunicamycin	YOR322C	5 / 6	0.029	0.232	-2.9%	0.594	0.800
Tunicamycin	YPL065W	2 / 2	0.035	0.217	-3.5%	0.626	0.812
Tunicamycin	YGR092W	4 / 5	0.059	0.266	-6.1%	0.685	0.856
Acetic acid	YKL037W	3 / 92	0.071	0.225	-7.4%	0.782	0.937
NaCl	YLR182W	4 / 8	0.057	0.169	-5.9%	0.803	0.937
Ethanol	YKL212W	1 / 1	0.333	0.750	-39.5%	0.907	1.000
Caffeine	YIR023W	4 / 6	0.174	0.355	-19.0%	0.944	1.000
Tunicamycin	YNL064C	1 / 1	0.217	0.382	-24.2%	0.983	1.000
Caffeine	YLR426W	2 / 9	0.244	0.398	-27.7%	0.994	1.000
Tunicamycin	YHR142W	4 / 101	0.385	0.540	-47.0%	1.000	1.000

Appendix E Over-representation of membrane related proteins

Table E6: Evaluated gene–chemical interactions grouped by their relation to membrane-bound proteins. Checkmarks indicate experimentally supported interactions (FDR < 0.05). Genes associated with membrane-bound proteins were implicated in 9 of the 10 experimentally supported hypotheses and 19 of the 35 suggested hypotheses. Fisher's exact test indicated a statistically significant association between membrane-relatedness and experimental support ($p = 0.0098$).

Membrane-related genes	Unrelated genes
Hydrogen peroxide	
YER083C / GET2 ✓	YML097C / VPS9
YOR070C / GYP1 ✓	YPL055C / LGE1
YOR322C / LBD19 ✓	
YMR015C / ERG5 ✓	
Tunicamycin	
YNL064C / YDJ1	YGR092W / DBF2
YPL065W / VPS28	YHR142W / SPO12
YOL018C / TLG2	YNL133C / FYV6 ✓
YOR322C / LBD19	
Caffeine	
YJL183W / MNN11 ✓	YML007W / YAP1
YGR036C / CAX4 ✓	YIR023W / DAL81
	YLR087C / CSF1
	YLR426W / TDA5
Acetic acid	
YER083C / GET2 ✓	YLR386W / VAC14
YOR070C / GYP1 ✓	YKL037W / AIM26
YNR006W / VPS27	
YDR069C / DOA4	
Ethanol	
YKL212W / SAC1	YDR495C / VPS3
YDR137W / RGP1	YAL021C / CCR4
YNL294C / MSB3	
NaCl	
YJL204C / RCY1 ✓	YOL081W / IRA2
YHR060W / VMA22	YOR332W / VMA4
	YLR182W / SWI6

Appendix F RIMBO extension

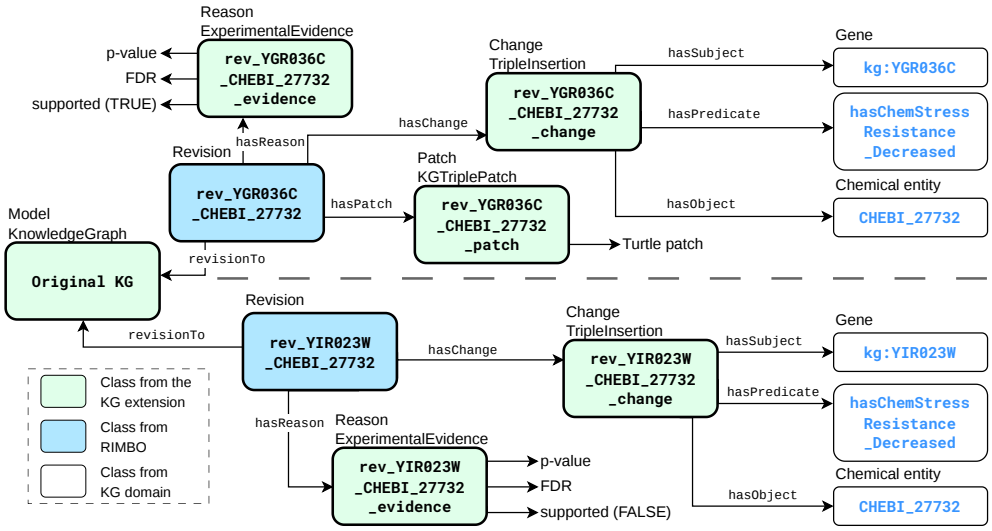


Fig. F1: Examples of knowledge graph revisions represented using the RIMBO extension. The upper example shows an experimentally supported hypothesis proposing the triple (`YGR036C`, `hasChemStressResistance_Decreased`, `CHEBI_27732`), for which a knowledge graph patch is generated and can be applied to construct the revised knowledge graph. The lower example shows an experimentally unsupported hypothesis proposing the triple (`YIR023W`, `hasChemStressResistance_Decreased`, `CHEBI_27732`). In this case, the proposed triple insertion and experimental evidence are retained as provenance, but no knowledge graph patch is generated and the proposed change is therefore not applied. Colours distinguish classes originating from the KG extension, RIMBO, and the KG domain.