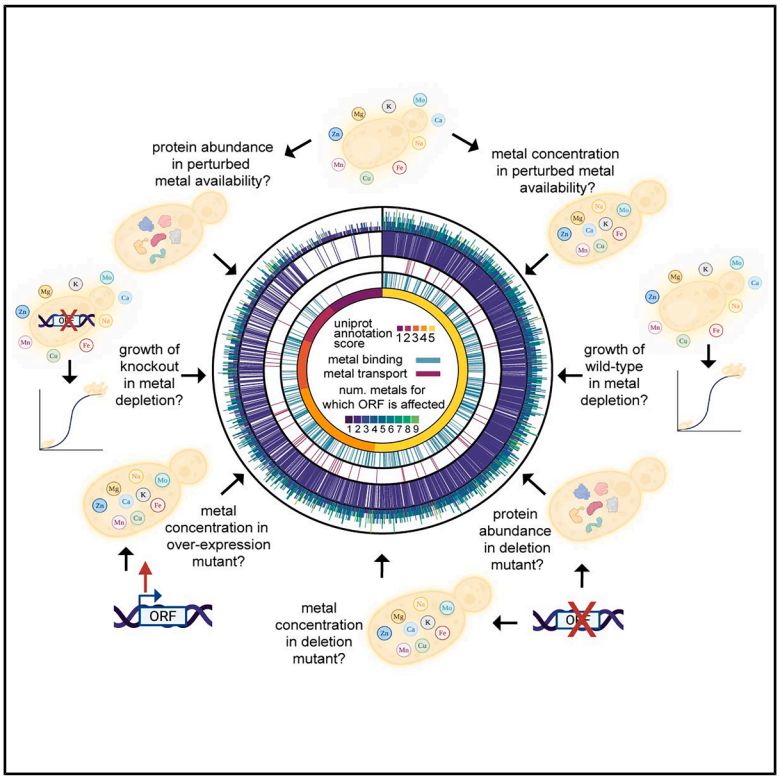
The molecular landscape of cellular metal ion biology

Article

Cell Systems

The molecular landscape of cellular metal ion biology

Graphical abstract



Highlights

- Cellular metal levels are buffered against environmental fluctuations
- Metal ion homeostasis affects >60% of yeast proteome and key signaling networks
- Many thus far uncharacterized genes function in metal ion homeostasis
- The centrality of metalloenzymes renders metabolism highly metal ion responsive

Authors

Simran Kaur Aulakh, Oliver Lemke, Lukasz Szyrwiel, ..., Michael Mülleder, Jens Nielsen, Markus Ralser

Correspondence

markus.ralser@charite.de

In brief

Aulakh et al. systematically investigated how metal ion availability impacts yeast biochemistry. Their metallomic, proteomic, and genetic screens reveal the widespread metal concentration dependence of over half the eukaryotic proteome (including key metabolic and signaling pathways), identify novel metal-related proteins, and have implications for biotechnology and medicine.

Aulakh et al., 2025, Cell Systems *16*, 101319
July 16, 2025
https://doi.org/10.1016/j.cels.2025.101319

CellPress

Cell Systems

CellPress
OPEN ACCESS

Article

The molecular landscape of cellular metal ion biology

Simran Kaur Aulakh,[1,2] Oliver Lemke,[3,4] Lukasz Szyrwiel,[3] Stephan Kamrad,[1] Yu Chen,[5,6] Johannes Hartl,[3,4] Michael Mülleder,[3] Jens Nielsen,[5,7] and Markus Ralser[1,2,3,4,8,9,*]

[1]Molecular Biology of Metabolism Laboratory, The Francis Crick Institute, 1 Midland Road, NW1 1AT London, UK
[2]Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, Oxford, UK
[3]Department of Biochemistry, Charité – Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany
[4]Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Berlin, Germany
[5]Department of Life Sciences, Chalmers University of Technology, SE41128 Gothenburg, Sweden
[6]State Key Laboratory of Quantitative Synthetic Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China
[7]BioInnovation Institute, Ole Maaløe Vej 3, DK2200 Copenhagen N, Denmark
[8]Max Planck Institute for Molecular Genetics, Ihnestrasse 73, 14195 Berlin, Germany
[9]Lead contact
*Correspondence: markus.ralser@charite.de
https://doi.org/10.1016/j.cels.2025.101319

SUMMARY

Metal ions have crucial roles in cells, but the impact of their availability on biological networks is underexplored. We systematically quantified yeast cell growth and the corresponding metallomic, proteomic, and genetic responses to perturbations in metal availability along concentration gradients of all growth-essential metal ions. We report a remarkable metal concentration dependency of cellular networks, with around half of the proteome, and most signaling pathways such as target of rapamycin (TOR), being metal responsive. Although the biological response to each metal is distinct, our data reveal common properties of metal responsiveness, such as concentration interdependencies and metal homeostasis. Furthermore, our resource indicates that many understudied proteins have functions related to metal biology and reveals that metalloenzymes occupy central nodes in metabolic networks. This work provides a framework for understanding the critical role of metal ions in cellular function, with broader implications for manipulating metal homeostasis in biotechnology and medicine.

INTRODUCTION

Metal ions played an important role during early metabolic evolution and remain crucial for various biological processes, such as cell growth, protein folding, DNA repair, neurotransmission, and immune function.[1] As catalysts in enzyme active sites, reactant cofactors in redox reactions, and mediators of protein-protein and protein-small-molecule interactions, they are critical for metabolic reactions,[2,3] protein evolution,[4] energy transformation, stress response, and cellular signaling across prokaryotic and eukaryotic metabolic biochemical networks.[5] The essential roles of metal ions are reflected in the diversity of metal ion transporters and binders encoded in the proteomes of numerous species.[6]

Metal ions fluctuate in the extracellular environment, consequently, cells must sense, control, and buffer their concentrations intracellularly.[7,8] Yet, at present, experiments addressing the role and effects of fluctuating metal ion concentrations are heavily underrepresented in the scientific literature. For example, metal ion concentration was modified in only 0.8% of ~14,500 yeast deletion (knockout) mutant screens,[2,9] and, of these, most only explored the effects of metal ion toxicity and did not consider the effects of concentration changes within a physiologically relevant range (Figure S1A). Moreover, we noted that none of these screens had assessed the impact of metal ion depletion in minimal media (devoid of amino acid supplements), which is a requirement for assessing the role of metal ions in biosynthetic metabolic pathways where metal-containing enzymes serve as catalysts (Figure S1B). However, focused studies on the depletion of individual metals have revealed widespread cellular responsiveness to altered metal ion availability on the transcriptome and proteome,[4,10–14] calling into question the practice of not varying or measuring metal concentrations in culture media in laboratory experiments. Mapping the interaction of metal ion levels and the cellular transcriptome or proteome poses numerous challenges. For instance, because of the extensive cellular buffering capacity for metal ions,[5,7] changes in metal concentrations in media do not directly translate into similar changes in cellular metal concentrations.[15–17] Moreover, metal ion transporters are promiscuous, such that altering the concentration of one metal ion can inadvertently influence the cellular

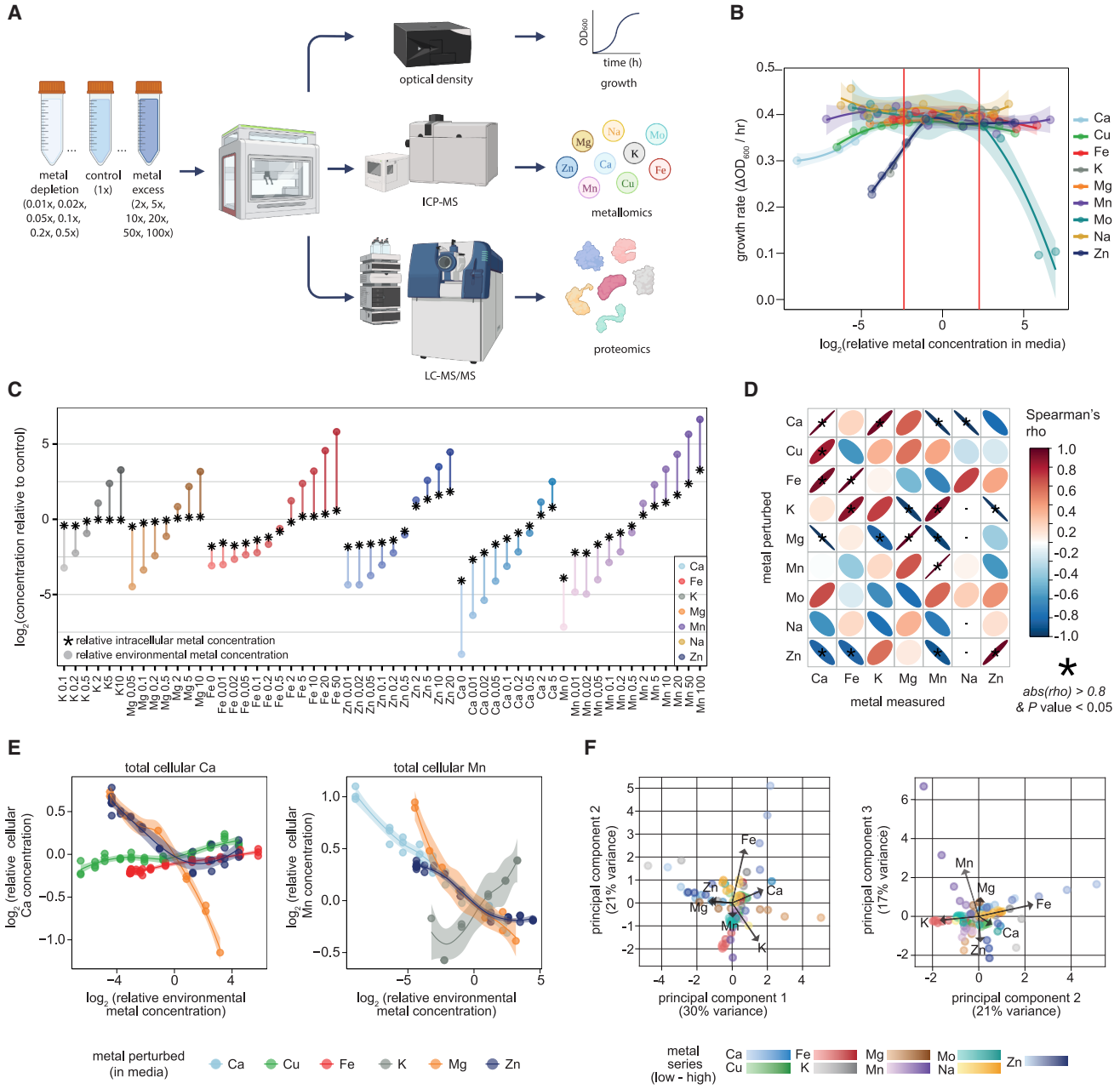


Figure 1. Metal ion homeostasis and concentration interdependence

(A) Experimental design for media generation and the growth, metallomic, and proteomic characterization of *S. cerevisiae* cells cultivated in metal perturbation conditions. The 91 media consist of a series of concentrations of Ca, Cu, Fe, K, Mg, Mn, Mo, Na, and Zn. Created in BioRender. Grüning, N. (2025) https://BioRender.com/0pwuf6j.

(B) Growth rate of a prototrophic BY4741 derivative along each metal perturbation series. Colors denote the metal concentration that was changed in the environment and the two red lines correspond to an environmental metal concentration change of 5-fold.

(C) The buffering capacity of cellular metal concentrations against changes in environmental metal concentrations. Different metals are color coded and the intensity of color reflects metal concentration in metal perturbation cultivation media relative to those in control media, with lighter circles indicating lower concentration and darker circles indicating higher concentrations. Black stars indicate cellular metal concentrations in each environmental condition relative to control conditons.

(D) Correlation between the metal concentration in cultivation media (environmental concentration) and the cellular concentration of other metals.

(E) Correlation between the environmental concentration of one metal (*x* axis) and the measured cellular concentration of other metals (*y* axis). Different metals are color coded and the intensity of color reflects metal concentrations, with lighter circles indicating lower concentration and darker circles indicating higher concentrations.

(legend continued on next page)

concentration of other metal ions, therefore, the transcriptome and proteome responses measured in cells under metal perturbation conditions are only indirectly related to any extracellular change in metal ion concentration. As a consequence, the role of metal ions in cellular networks remains understudied, even in well-studied model organisms like *Saccharomyces cerevisiae*.

In this study, we aimed to address the gaps in knowledge about the functions of metal ions in cellular biochemical networks and create a genome-spanning resource that systematically records cellular responses to environmental variation in essential metal ions across multiple molecular layers.

RESULTS

Homeostasis is metal specific and involves concentration interactions

We started by establishing cultivation conditions in which the nine typical metallic constituents of synthetic yeast medium—calcium (Ca), copper (Cu), iron (Fe), potassium (K), magnesium (Mg), manganese (Mn), molybdenum (Mo), sodium (Na), and zinc (Zn)—are varied in 12 steps, over five orders of magnitude (Figure 1A; Table S1; Methods Table S10). We created these series in synthetic minimal media, without amino acid supplements,[18] to ensure that the cultured cells are dependent on the amino acid biosynthetic pathways, many of which require metal-binding proteins.[19,20] We did not use chelators to completely deplete metals below trace concentrations, to avoid any confounding off-target effects[21–23] (Methods Table S11). Instead, to account for the fact that metal ions can be present as trace contaminants, we accurately determined the metal concentration for each condition generated using inductively coupled plasma mass spectrometry (Figure S1C; Table S2). Then, we selected a prototrophic, haploid *S. cerevisiae* strain derivative of BY4741[19,24,25] as our model organism. We cultured these yeast cells in each of the resulting 91 media (and control) combinations in triplicate and monitored their growth rates.

Consistent with an extensive homeostatic machinery that allows cells to buffer cellular metal ion concentrations in response to environmental changes,[7,16] the cells showed consistent growth across large parts of the concentration series (growth was maintained in 79/91 total conditions across the 9 series) (Figures 1B and S1D; Table S3). In the "depleted" media (i.e., in media with metal ion concentrations below the concentration in standard media), yeast growth rates slowed upon an ~2-fold depletion of the abundant metals K, Mg, and Zn and upon an ~8-fold depletion of Ca and Cu (Figure 1B). By contrast, no growth defects were observed when depleting Mo, Mn, Fe, and Na to the lowest levels. Because the lowest Fe concentration obtained through metal omission was 1/10th of the concentration in standard media, the absence of growth defects was expected. However, for other metals, where omission efficiently depleted the metal ion (Na to 1/100th of the typical media levels, Mn to 7/1,000th, and Mo to 4/100th), this absence of growth defects was surprising and suggests that these elements are required either in extremely low amounts, or not required at all, for yeast to grow. In the "excess" media, only changes >5-fold (1 mg/L Na_2MoO_4) of Mo and >20-fold of Cu and Fe (0.8 mg/L $CuSO_4$ and 5 mg/L $FeCl_3$) resulted in a slowing of cell growth, indicating toxicity.

We next aimed to annotate the relationship between intracellular and extracellular metal ion concentrations. First, we quantified the total Ca, Fe, K, Na, Zn, and Mg cellular concentrations (Table S4) using a modified version of an ICP-MS based metallomics protocol[26] (STAR Methods). The relative concentrations of these metals were consistent with previous reports, thus validating our approach (Figure S1F). We then determined "buffering range," the range of extracellular metal ion concentrations during which cells can maintain a consistent cellular concentration. The broadest buffering ranges were observed for K, Mg, and Fe, followed by Zn, Ca, and Mn (Figure 1C; Table S5). This pattern was consistent when cellular concentrations of each metal were measured in response to changes in environmental concentrations of all other metals that we tested, with the maximum decrease in intracellular K and Mg observed being 28% (K) and 29% (Mg), across all perturbation conditions, and the maximum increase being 8% (K) and 12% (Mg) (Table S4).

Because many proteins that interact with biologically relevant metal ions exhibit promiscuity of metal interactions due to the similar physical and chemical properties of metal ions,[7,8] we next determined the relationship between the environmental concentration of one metal and the cellular concentration of every other metal. We found that 7/9 environmental metal concentration series altered the cellular concentrations of at least one other metal (Figure 1D; Table S6). Cellular Mn and Ca levels were most sensitive to changes in the extracellular concentrations of other metals (Figure 1E), i.e., the cellular concentrations of Mn and Ca showed high correlation to changes in environmental concentrations of other metals.

Principal-component analysis (PCA) of the cellular metallomic profiles revealed that although samples could be arranged along the environmental metal concentration gradients in which the cells were cultivated, samples from each metal series could not be completely separated from other series, likely due to the interdependence of cellular metal concentrations (Figure 1F).

The yeast proteome responds globally to changes in metal availability

We next conducted a quantitative proteomic experiment to capture the molecular responses of yeast cells grown in the 88 conditions that allowed cells to grow. We employed a high-throughput proteomic pipeline that combined cell cultivation in multi-well plates, semi-automated sample preparation, microflow liquid chromatography, data-independent mass spectrometry data acquisition, and data analysis using the DIA-NN software.[27,28] We obtained precise quantitative expression values for 2,330 unique proteins (representing 22% of the annotated *S. cerevisiae* proteome), of which 1,433 were detected in at least 85% of all samples (STAR Methods; Table S7). Then, using quantitative copy-number data for 3,841 proteins (63% of the *S. cerevisiae* proteome[29]), we estimated that our dataset

(F) PCA based on all measured total cellular metal concentrations separates a subset of samples according to the metal perturbed in the environment. Colors indicate each unique cultivation condition, with darker shades indicating high amounts of each metal and lighter shades representing low amounts. Arrows indicate increasing value of PCA loadings.

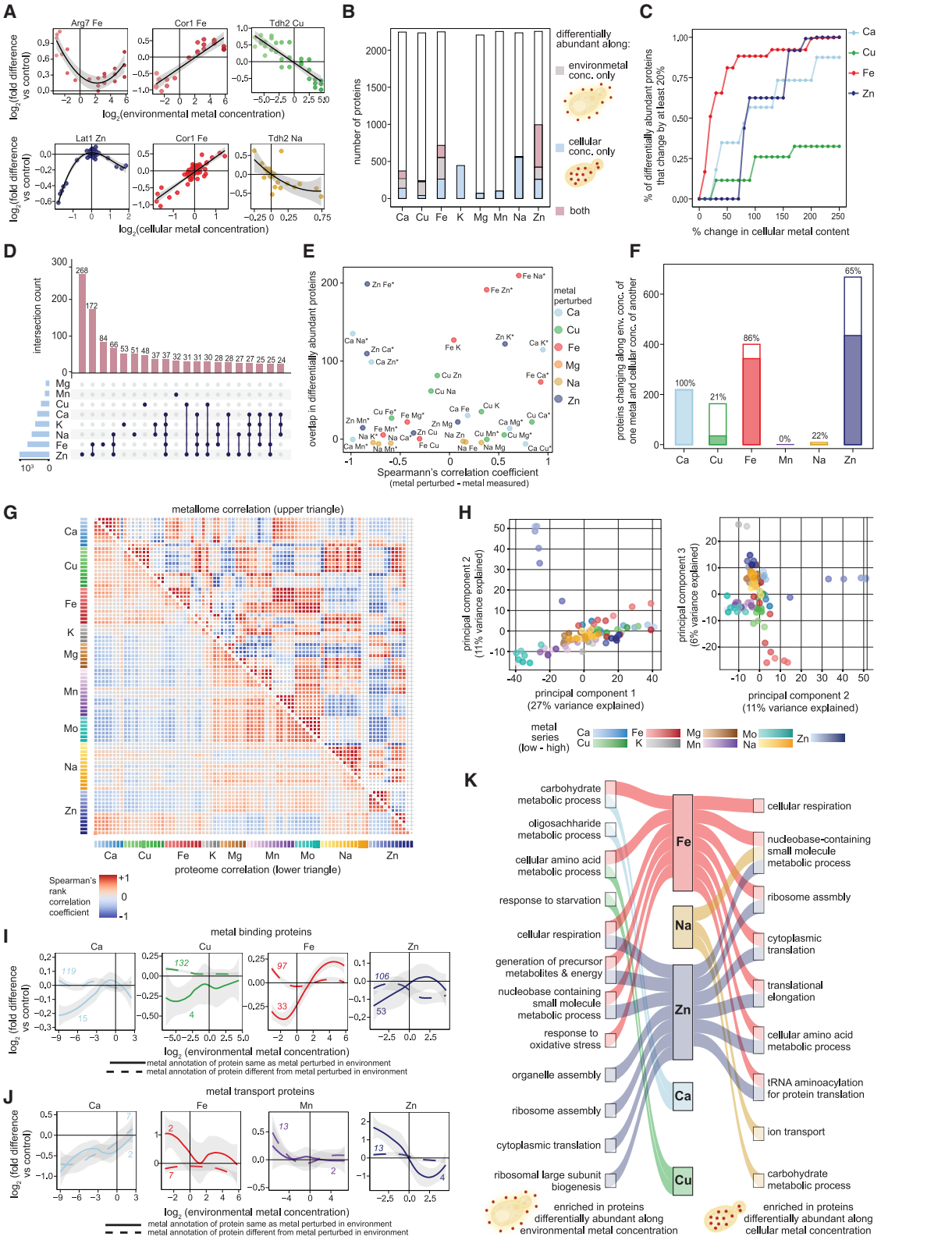


Figure 2. Identifying metal-responsive proteins

(A) Examples of six proteins that represent different metal ion-protein abundance relationships along an environmental (top row) and cellular (bottom row) metal concentration series. The plots display protein abundance changes in response to the varied metal ion concentrations.

(legend continued on next page)

quantifies around 87.9% of the proteomic mass (STAR Methods), with the technical noise being much smaller (coefficient of variation [CV] of 15.7%) than the biological variation between conditions (CV ~27.9%).

To identify differentially expressed proteins, we first determined which statistical model best represents the relationship between environmental metal concentrations and protein abundance in our cells: no relationship (null), linear (1st degree), quadratic (2nd degree), or a cubic (3rd degree) (Figures 2A and S2A; Tables S8 and S9; STAR Methods). We defined a protein as differentially abundant if the *p* value of the model that best represents the modeled relationship was <0.05 and if the difference between the minimum and maximum protein quantity along the environmental metal concentration series was at least 1.5-fold. We used the same methodology to identify proteins altered along measured cellular metal concentrations. We used binned metal concentrations and protein abundance data across the entire dataset and identified protein-metal interactions resulting from the interdependencies between different metals.

In total, 1,545 unique proteins were found to be metal responsive, representing 66% of the measured proteins and 81% of the quantified protein mass. On average, 205 proteins were altered per metal perturbation series when considering environmental concentration and 342 when considering cellular metal concentrations (Table S7). The most pronounced response was to Zn (995 responsive proteins when considering both environmental and cellular concentrations), contrasting with Mo, for which we detected no responsive proteins upon its depletion (Figure 2B). Using measured cellular concentration changes, we identified 445 additional metal-responsive proteins for K and 71 for Mg (Table S7).

Across all metals, 1,055 proteins (45% of the proteins we quantified) were responsive to at least one environmental metal concentration change, and 1,386 proteins were responsive to at least one cellular concentration change. By contrast, 159 and 490 proteins changed in response to environmental (Figure 2A, top) or cellular (Figure 2A, bottom) variations in only one metal, respectively. Finally, 712 proteins (~60%) responded along both the environmental and the cellular concentration series (either for the same or different metals) (Figure 2A).

We finally aimed to map the "tipping points," which we define as the minimum change in metal concentration that initiates a significant shift in the proteome relative to the control steady state condition. To achieve this, we defined the threshold for protein-level responsiveness as a change in protein abundance greater than 50% relative to the control, and then evaluated the cumulative fraction of responsive proteins at 5% increments in the magnitude of the measured cellular metal concentration (Figure 2C). We revealed notable differences in the critical concentration of metal ions that cells respond to at the protein level. For instance, at least a 75% change in the cellular Zn concentration but only a 10% change in Fe cellular concentration was required to induce a proteomic response (Figure 2C).

The cellular response to metal ion perturbation reflects concentration interdependencies between metal ions

Although altering levels of Fe or Zn individually impacts the transcriptome[10–12] and proteome,[13] our metallomic data suggest that the cellular proteomic response to a perturbation of metal availability may also depend on how metal levels interact with each other (Figure 1D). Indeed, 761 proteins, or 49% of all metal-responsive proteins, showed changes in protein abundance, possibly caused by covarying metal concentrations. To assess how interlinked the proteomic responses to perturbations of different metals were, we first visualized the intersections between differentially abundant proteins along each metal (Figure 2D). Of the 10 largest intersections, four (Fe-Zn, Fe-Zn-Na, and Na-K) corresponded to differentially abundant proteins in a combination of two or more metals (Figure 2D). Proteins varying along the Zn series, followed by those changing along both Zn and Fe, and Fe alone, triggered the largest responses (Figure 2D).

We next compared the number of differentially abundant proteins for each environmental and cellular metal concentration series with the correlation between the measured environmental (cultivation media) and cellular metal concentrations for each pair (Figure 2E). Some intersections resembled the metal-metal correlations unveiled by the metallomic data. For example, 165 of the 717 proteins that were differentially abundant along the

(B) The number of differentially abundant proteins along each metal concentration series.
(C) The cellular metal concentration thresholds at which the proteome responds to a metal perturbation.
(D) Overlap between differentially abundant proteins along each metal perturbation series (either environmental or cellular). Pink bars indicate the size of the overlap of differentially expressed proteins between the metal series listed on the left. Light blue bars indicate the number of proteins differentially abundant along each metal perturbation series.
(E) Relationship between the number of proteins that are differentially abundant both along environmental concentration of one metal as well as the cellular concentration of another metal and the correlation between the environmental concentration of the first metal and the cellular concentration of the second. The first metal in the text next to each point indicates the environmental (media) concentration series and the second indicates cellular (measured) concentration series.
(F) The fraction of differentially abundant proteins shared between the proteomics response to changes in environmental concentration of one metal and the proteomic response to changes in cellular concentration of another that can be explained by interlinked metal cellular concentration changes. *y* axis: border indicates total number of differentially abundant metal-binding proteins along environmental concentrations of metals indicated on *x* axis that are also differentially abundant along cellular concentrations of any other metal. *y* axis fill indicates the number of such proteins that can be explained by a correlation between environmental concentration of metal on *x* axis and cellular concentration of another metal.
(G) Comparison of correlation coefficients between metallomic and proteomic profiles for each pair of unique cultivation conditions.
(H) PCA of proteomic data, showing separation of cultivation conditions along the first three principal components.
(I and J) Average abundance of metal-binding proteins (I) or metal transporters (J) in *S. cerevisiae* cells cultivated in media with a series of metal concentration variations. Numbers in upright font indicate the total number of proteins bearing annotations for the same metal being perturbed, those in italics indicate the total number of proteins that have an annotation that is different from the metal being perturbed.
(K) Gene Ontology Slim (biological process) terms enriched in differentially abundant proteins, depending on the extracellular (left) and cellular (right) metal ion concentration. The perturbed metal is indicated in color in the center.

Fe series (environmental and cellular concentrations considered together) were also significantly altered in abundance along the Zn or Na series (Figure 2E). This overlap coincided with a high correlation coefficient between environmental Fe concentration changes and cellular Na and Zn concentrations, as well as the link between environmental Zn concentration and cellular Fe concentration (Figure 2E). Ca also showed a strong relationship with other metals at the proteome level: all differentially abundant proteins along the environmental Ca series and the cellular concentration of any of the other metals could be explained by a metal-metal connection (correlation between environmental concentration of the metal perturbed and the cellular concentration of another metal) discovered in the metallomic data (Figure 2F).

We then evaluated how closely the proteome of a sample is related to its metal content by measuring the similarity between the proteomic and metallomic profiles across all samples. We computed correlation coefficients for all possible condition pairs (3,828 unique pairs from combinations of 88 conditions) based on the metallome and proteome separately and then compared the correlation coefficients for each paired sample to assess whether the proteome and metallome capture different aspects of metal ion biology (Figure 2G). This revealed an overall agreement between the proteomic and metallomic data about the relative similarity of pairwise samples, with a few minor differences and more granularity in groups of similar samples being captured by the proteomic data. PCA of the proteomic dataset revealed that Zn depletion, Fe depletion, and K depletion samples were clearly distinct from each other, whereas the remaining samples formed biologically relevant clusters of conditions, with a small overlap (Figure 2H). While PCA of the metallomic dataset showed a clear separation between Mg depletion and Mn excess samples were distinct groups, PCA of the proteomic dataset indicated that Mn excess samples clustered near the Mo and control samples and Mg depletion clustered centrally among all other samples.

Altered metal availability induces specific cellular responses

Metal-binding proteins were responsive to variation in the concentration of the metal they bind (Figure 2I). Our data show that the extent of this relationship differs across metal ions. For example, Zn depletion had a lower impact upon the abundance of Zn-binding proteins relative to the effect Ca, Cu, or Fe depletion on proteins that bind each corresponding metal (Figures 2I and S2H). Also, the responses within metal ion transporters were distinct. While the abundance of Fe, Zn, and Mn transporters increased in response to the depletion of the metal that they transport, the abundance of the two quantified Ca transporters decreased at high Ca levels, indicating the presence of a negative feedback response (Figures 2J and S2I). Intriguingly, the abundance of metal transporters annotated to transport metals other than Ca also decreased in the context of Ca depletion, further reflecting the interlinked nature of metal ion homeostasis (Figure 2J).

Because the metal requirement for the mitochondrial respiratory chain is well established, we next queried our dataset to explore the relationship between metal availability and the generation and maintenance of the proton motive force (PMF) across various cellular membranes. Our findings reveal that alterations in metal availability impact proteins involved in PMF generation, not only across mitochondrial membranes but also within the vacuolar, plasma, Golgi, and the interconnected endosomal system membranes (Figures S2J and S2K).

To explore the proteomic responses in a systematic manner, we conducted a gene set enrichment analysis using the Gene Ontology (GO), GO slim, and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases. Cellular respiration, translation and transcriptional processes, stress response pathways, metabolic pathways, and ion homeostasis processes were over-represented among the differentially abundant proteins (Figure 2K). While our analysis recapitulated known metal-specific molecular functions: for example, Fe binding, Fe-S cluster binding, heme binding, and enzymes with oxidoreductase activity were enriched along the environmental Fe perturbation series; lyase and oxidoreductase activity were enriched along the Cu series; and many ribosomal and oxidoreductase processes were enriched along the Zn series (Figure S2G; Table S10), it also revealed less-well-documented responses to changes in metal ion availability. Notably, we observed broad crosstalk between metal ion concentrations and cellular signaling pathways. In 28/34 signaling pathways (GO biological process annotations) for which proteins were quantified in our dataset, at least one protein was responsive to a perturbation of metal availability (Table S11). These included known metal response pathways, such as the calcineurin signaling pathway,[30] both quantified proteins of the osmosensory pathway,[31] but also signaling pathways with other canonical functions, including all four quantified proteins (Gpa1, Gpa2, Asc1, and Cyr1) of the G-protein-coupled receptor pathway (involved in the pheromone response and glucose sensing[32]) and four of the five proteins quantified that map to protein kinase A signaling.[33] Furthermore, our dataset revealed a strong metal response within the target of rapamycin (TOR) pathway, which, to the best of our knowledge, has not been associated with metal ion responsiveness thus far. Specifically, seven out of eight quantified TOR-related proteins (Sit4, Ksp1, Kog1, Slm1, Stm1, Tap42, and Tip41) were differentially expressed in metal perturbation media. For example, Kog1, a subunit of TORC1, responded to Cu, Fe, and Zn availability, whereas Sit4 responded to Fe and Na.

Due to their low abundance, we could quantify only 22 transcription factors. Twelve of these, such as Yap1, which has a known role in Fe homeostasis,[34,35] and Zn-finger or Zn cluster transcriptional activators such as Cat8, Gat1, and Gts1, were differentially expressed in at least one metal perturbation series.

Our dataset quantified at least one component of 340 protein complexes (GO cellular compartment annotation). We detect a metal ion response in 289 of these (Table S12). For 145 complexes, we quantified at least 75% of the components (Table S12). In 128 of these (~88%), at least one component was metal responsive, with 112 (~77%) showing a change in at least 50% of their components. All the 38 large (five or more proteins involved) complexes (including the proteasome, vacuolar proton-transporting V-type ATPase, retromer complex, mannan polymerase, and GPI-anchor transamidase complex) that were quantified with over 75% coverage contained at least one protein that was metal responsive.

Finally, our dataset revealed a network-wide responsiveness of metabolism-to-metal availability. In 38/39 KEGG metabolic

pathways, for which we quantified more than 75% of the enzymes, at least one enzyme was differentially expressed along the metal concentration series. In 35 of these pathways, at least 50% of the quantified enzymes were metal responsive (Table S13). Highly metal-responsive KEGG pathway terms include steroid biosynthesis (85% proteins responsive), glycolysis (75%), tri-carboxylic acid (TCA) cycle (72%), and biosynthesis of secondary metabolites (81%). All proteins mapping to fatty acid biosynthesis and elongation (5), histidine metabolism (12), thiamine metabolism (4), and propanoate metabolism (9); 14/15 enzymes of tryptophan metabolism; and 10/11 enzymes of the lysine biosynthesis pathway were differentially expressed alongside a metal concentration series. Notably, the only four KEGG terms that we quantified at a high coverage, but for which we did not observe a high responsiveness, are indirect participants in the metabolic network (i.e., ABC transporters, protein export, and RNA polymerase) and, thus, essentially, all measured primary metabolic processes were responsive to at least one metal ion perturbation.

Quantitative proteomic responses to perturbations of metal availability cluster proteins according to function

Having cataloged the broad set of cellular responses to altered metal availability, we next explored our datasets to obtain novel insights into protein function. Previous studies, including our own, on the yeast metabolome and proteome[27,36] have revealed that the clustering of "omic" profiles can be an effective strategy for protein functional annotation. Therefore, we employed an ensemble clustering approach[37] and analyzed the proteomes in two parallel analysis pipelines: in the first, we clustered the proteomes of cells cultivated in each individual metal series separately, whereas in the second, we clustered all proteomic data obtained in the metal series together.

For the former (metal-wise clustering), only proteins that were identified as differentially abundant along each metal series were retained for that specific metal. For the latter (all-metal clustering), all proteins detected in at least 85% of the entire dataset were included. In both instances, we used three clustering algorithms—density-based CommonNN,[38,39] spatial k-Means(++),[40,41] and a community-detection algorithm (Leiden[42]). Then, we integrated the co-clustering matrices into a singular matrix with equal weighting, followed by a final hierarchical clustering (Ward's method[43]) (Figure S3A). We obtained a total of 96 "fuzzy" (partially overlapping) clusters (with a range from 4 clusters for Mg and Mn to 27 clusters along the Zn concentration series) from the metal-wise clustering pipeline and 35 clusters from all-metal clustering. The coarse structure of the clustering was mainly driven by the Leiden-clustering algorithm and the fine structure by the CommonNN- and the k-Means-clustering algorithms.

Next, we performed a functional enrichment analysis for each cluster using the GO, GO slim, KEGG, and Enzyme Commission (EC) databases (Figures 3A–3C). Twenty of the 35 all-metal clusters and 26/96 of the metal-wise clusters, cumulatively representing ~60% (1,061/1,764) of the proteins, were enriched in at least one functional term (Table S14). For example, a cluster of 29 proteins displayed an increase in abundance at low Cu concentrations (Figure 3A). These proteins included enzymes for amino acid biosynthesis, mitochondrial function, and metabolism, thus reflecting the key role of Cu for mitochondrial respiratory chain proteins and amino acid precursor production. Another cluster obtained via the all-metal clustering pipeline identified a group of 90 proteins with a complex profile, involving protein abundance changes along the Ca, Zn, Mn, Cu, and Fe series. The cluster was enriched for terms related to cation transport activity (Figure 3B).

Overall, the largest functional categories represented in the clusters were metal binding, mitochondrial proteins, and ribosomal functions. The remaining clusters were enriched for a range of GO biological process and KEGG pathway terms not usually linked to metal ions, such as the organization of the cytoskeleton and the assembly of organelles (Figure 3C; Table S14). Notably, 44 of the 72 previously poorly characterized proteins that were included in the clustering pipeline were placed within clusters enriched in specific functions (Figures S3B and S3C), suggesting novel functional links between the poorly characterized proteins and those with well-defined molecular functions. Taken together, these observations reflect the critical involvement of metal ions in major organelles of eukaryotic cells, as well as the potential of our approach to propose novel protein function annotations.

Incorporating functional genomics datasets to elucidate protein function

Having discovered that many of the poorly characterized proteins were assigned to clusters enriched for a functional term, we sought to validate these functional propositions with complementary molecular data. To evaluate whether the clustering-based associations provide relevant functional information, we incorporated additional and complementary genome-scale datasets relevant to metal biology. We cultivated a genome-wide deletion mutant collection consisting of 4,850 single-gene deletion mutants of *S. cerevisiae* in a prototrophic background[19,25] in 16 different metal omission media (see STAR Methods) for 48 h and obtained 357,972 colony size measurements, from which we calculated the effect sizes for the growth of each mutant under each cultivation condition.

We identified 734 significant genetic interactions with metal ion availability, involving 642 unique gene deletions (Figure 4A; Table S15; STAR Methods). At the individual metal level, most genetic interactions were discovered for K (516) and Mg (175), followed by Zn (26) and Ca (16) (Figures 4B and S4A). The identified genetic interactions were enriched for metal protein binding, endosomes, protein complexes, ribosomes, translation, mitophagy, amino acid, amide, and peptide biosynthetic pathways (Table S16). Translation, gene expression, and the nitrogenous compound and peptide metabolic processes were over-represented in the deletions that led to growth aberrations in K and Mg. For Ca depletion, we detected a specific signature for endosomal transport, vesicle-mediated transport, and Golgi-vesicle transport (Figure S4B; Table S16). To test whether vacuolar buffering could contribute to the maintenance of overall growth despite reductions in total cellular metal content that we observed earlier, we analyzed genes encoding vacuolar proteins. We identified 33 such genes with reduced fitness (effect size $<-\log_2(1.2)$, $p < 0.1$ after Benjamini-Hochberg correction). These genes encompass various functional roles, including vacuolar transporters, regulators of vacuolar morphology, and components of vacuolar

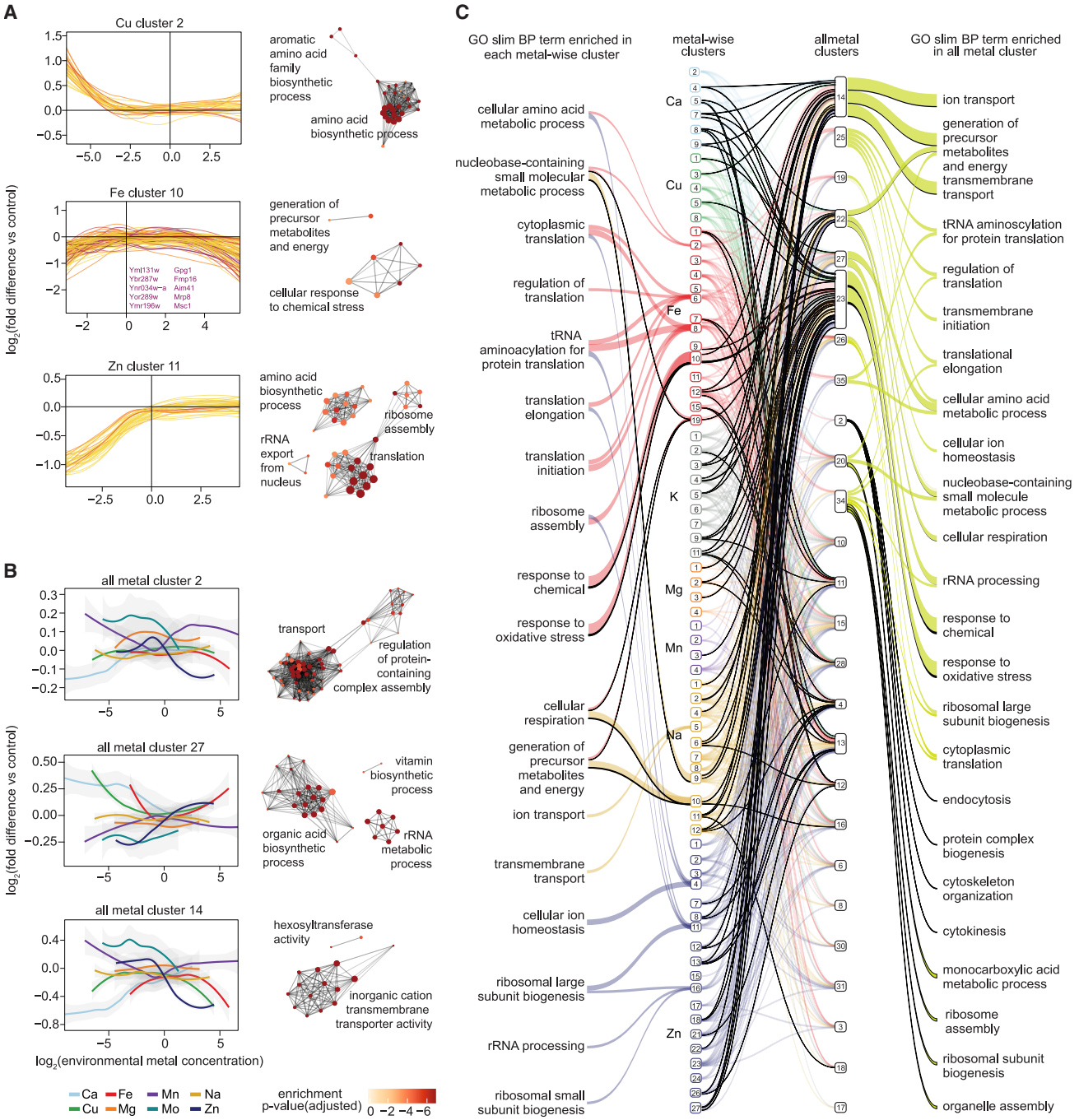
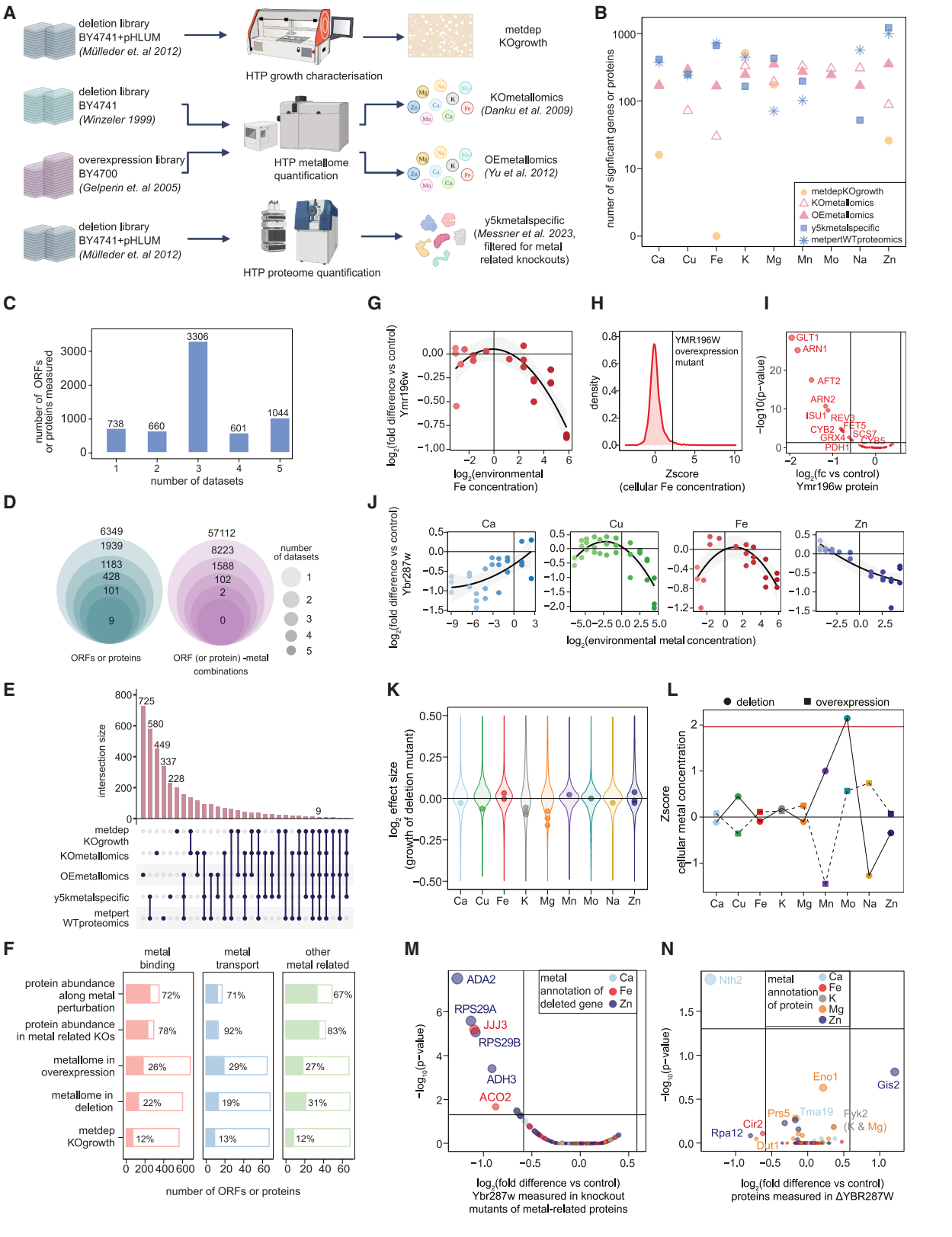


Figure 3. Metal responsiveness clusters proteins according to function

(A and B) Examples of clusters of proteins resulting from ensemble clustering analysis of proteomes from each metal series considered individually (A) and all metal series combined. In (A) the color indicates the UniProt annotation score and the names of any protein bearing UniProt annotation score < 3 (poorly characterized proteins). In (B) the color indicates the metal perturbation series, and the gray zone indicates the 95% confidence interval around the mean abundance of all proteins in each metal perturbation series. Network plots indicate the GO biological processes over-represented in the cluster, which are shown on the right: the circles represent individual gene set terms, the circle size corresponds to the number of proteins mapping to the term, gray lines connect gene sets with shared proteins, clusters are annotated with names that summarize groups of gene set nodes.

(C) Summary of the Gene Ontology slim biological process (GO slim BP) terms enriched in each cluster obtained via ensemble clustering. The two columns in the middle indicate the cluster number. The links between the outer text column and clusters (two middle columns) connect each cluster to the gene set term that was enriched in the cluster. Links between the two middle columns connect each protein to itself across the two types of clusters. Colored links represent proteins belonging to each metal-wise cluster with a UniProt annotation score >2; black links represent proteins with a UniProt annotation score < 3. Only proteins that are part of a cluster (metal-wise or all-metal) that is enriched for at least one GO slim-BP term are included.



(legend on next page)

protein sorting, among others, suggesting potential involvement of vacuolar mechanisms in growth maintenance under metal-limiting conditions.

Next, to determine how the deletion of a metal-binding or metal-transport protein impacts other proteins that are connected to the same metal, we integrated the quantitative proteomes of deletion mutants.[27] 1,391 unique proteins were differentially expressed in the 304 deletion mutants bearing a metal-related GO annotation (Table S17). In this dataset, all metals showed an average of six to eight differentially abundant proteins per deletion, except for K and Mn, which had 11 and 4 (Table S18). Further, we found that ~72% (48/66) of metal-binding protein deletions only impacted the abundance of one or two other metal-binding proteins. However, 9 deletions (like *ACO2*, *LEU3*, and *IDH1*), affected the expression levels of five or more metal-binding proteins (Figure S4C). These outliers can be partly explained by of the slow growth rate of the deletion mutants and the previously reported negative correlation between growth rate of deletion mutants and high proteome impact.[27] For Ca- and Cu-related proteins, a decrease in the abundance of metal-binding proteins upon the deletion of other proteins annotated to bind the same metal was common. By contrast, Fe- and Zn-binding proteins exhibited mixed patterns, i.e., whereas some metal-binding proteins increased in abundance upon deletion of a protein that binds the same metal as the measured protein, others decreased in abundance in response to the deletion, indicating complex regulatory interactions within the cell (Figure S3E).

Lastly, we incorporated gene-metal concentration interactions by including two datasets comprising cellular quantities of metal ions in the *S. cerevisiae* gene deletion and overexpression collections.[26,44] Using *Z* scores computed by Iacovacci et al.,[45] we found that the number of gene deletions that affected the cellular concentration of a metal was quite variable, with the highest number identified for Mn, followed by Mo, Na, K, Ca, Zn, Cu, and then Fe (Figures 4B and S4D).The original metallomic experiments to generate this dataset were conducted using heavy metal supplemented rich yeast peptone dextrose (YPD) media. Mn and Mo are toxic at the supplemented concentrations, which accounts for the high number of hits for Mn and Mo in these screens.

When we analyzed the metallomic profiles of deletion and overexpression mutants of genes encoding metal-binding proteins, we observed the following types of cellular metal abundance patterns. The first was a group of genes encoding Cu-, Fe-, or Mn-binding proteins for which a deletion leads to a decrease in the cellular quantity of the corresponding metal and vice versa (e.g., Cu concentration in CUP2 mutants and Mn concentration in YFR006W mutants) (Figure S3F). This likely reflects the loss of metal storage capacity upon deletion of the gene that encodes the metal binding protein. For another group

Figure 4. Data integration to generate a comprehensive resource for studying metal-responsive proteins

(A) Summary of genetic, metallomic, and proteomic data that were integrated with metallomic and proteomic data captured from cells cultivated along metal concentration gradients. Deletion mutant libraries have been abbreviated as KO for knockout and overexpression mutants as OE. Created in BioRender. Grüning, N. (2025) https://BioRender.com/e17pk2v.

(B) Number of open reading frames (ORFs) or proteins encoded by ORFs identified as significantly affected across the five datasets as summarized in (A). The *x* axis indicates the metal that was perturbed or the metal that was connected (based on Gene Ontology database molecular function [GO-MF] annotations) to the gene being deleted. The color indicates the type of assay (yellow, fitness inferred from end point colony size measurements; pink, cellular metal concentration; blue, proteomes); the shape indicates each individual dataset. Deletion mutant libraries have been abbreviated as KO for knockout and overexpression mutants as OE.

(C) Number of ORFs or proteins encoded by ORFs assayed across the five datasets. Deletion mutant libraries have been abbreviated as KO for knockout and overexpression mutants as OE.

(D) Intersection between ORFs (or protein encoded by ORF) and ORF (or protein encoded by ORF)-metal combinations that were identified as a significant hit across the five datasets. The size of the circle indicates how many datasets are considered, with the largest circle representing the set of ORFs or ORF-metal pairs that are significant in any one dataset and the smallest representing those that are significant in all five. Numbers outside the largest circle represent the total number of unique ORFs (or protein encoded by ORF) and ORF (or protein encoded by ORF)-metal pairs that were measured in all five datasets, cumulatively.

(E) Intersection between the list of ORFs (or protein encoded by ORF) that were a significant hit in any of the five datasets. Black circles (and lines) indicate the identity of the datasets in the overlap; pink bars indicate the number of ORFs (or protein encoded by ORF) that are shared between the datasets indicated by the black circles and lines. Deletion mutant libraries have been abbreviated as KO for knockout and overexpression mutants as OE.

(F) Metal-related proteins identified as significantly affected in each dataset. Panels represent the type of annotation a protein has in the GO molecular function (MF) database. The outer bar outline represents the total number of ORFs (or protein encoded by ORF) measured or assayed in each dataset and the inner filled-up bar represents the number that was significantly affected in at least one metal perturbation condition or metal-related mutant). Deletion mutant libraries have been abbreviated as KO for knockout and overexpression mutants as OE.

(G) Protein abundance changes of Ymr196w in *S. cerevisiae* cells cultivated under varying Fe concentrations.

(H) *Z* score of the cellular Fe concentration of the YMR196W overexpression mutant relative to *Z* scores of all other mutants. Black vertical line indicates the *Z* score of the YMR196W overexpression mutant (2.23).

(I) Impact of Fe related protein deletions on Ymr196w protein abundance in each mutant strain. The points correspond to Ymr196w protein abundance in each deletion mutant (indicated by gene names on the figure).

(J) Protein abundance of Ybr298w in wild-type (WT) *S. cerevisiae* cells cultivated in each metal concentration series (indicated by panel) along which it was deemed to exhibit a significant change.

(K) Impact of Ybr298w gene deletion on *S. cerevisiae* cell growth in metal depletion media.

(L) Impact of Ybr298w deletion and overexpression on cellular concentrations of various metals. The *x* axis shows the metal that was quantified using ICP-MS by Danku et al.[26] (deletions) and Winzeler et al.[25] (overexpression). *y* axis: *Z* score of the concentration of each metal in the YBR287WW mutants.[27] The color indicates the metal that was quantified.

(M) Impact of deletion of metal-related proteins on the abundance of Ybr298w protein in each deletion mutant. Points correspond to Ybr298w protein abundance in each deletion mutant (indicated by gene names). Colors indicate the metal annotation of the deleted gene based on the GO molecular function (MF) database.

(N) Impact of the deletion of Ybr298w gene on the abundance of metal-related proteins. Points correspond to protein abundance of each metal-related protein (indicated by gene names) in the Ybr298w deletion mutant. Colors indicate the metal annotation of the measured protein.

of deletion mutants, we saw an increase in the cellular concentration of a metal upon deletion of a gene encoding a protein that binds the same metal, while displaying a concomitant small increase in the metal or no change in the overexpression mutant (e.g., *CNE1*, *EDE1*, *CNB1* [for Ca], *SOD1* [for Cu], and *IDH1* and *IDH2* [for Mg]) (Figure S3F). These findings might reflect a disruption of metal homeostasis and the activation of compensatory mechanisms. For a minority of genes, exemplified by Cu concentrations in *COX11* mutants, we observed similar cellular metal concentration changes in both deletion and overexpression mutants (Figure S3F).

Multi-layer data integration improves the functional annotation of understudied proteins

Out of a total of 6,349 putative and confirmed protein-coding open reading frames (ORFs) represented in the five datasets (proteomics of metal perturbation samples, proteomics of deletion mutants, metallomics of deletion mutants, metallomics of overexpression mutants, and growth of overexpression mutants), we could assess about half in at least three datasets and 1,044 across all five datasets (Figure 4C). In total, around half (3,662/6,349) of the tested ORFs were associated with a metal ion in at least one dataset (Figure 4D, left), with 110 ORFs showing a phenotype in at least one metal condition in three or more datasets; only nine showed phenotypes across all 5 datasets (Figure 4D, left). When ORF-metal pairs were considered, only two exhibited phenotypes across 4 datasets, 104 across $\geq$3 datasets, and 1,692 across $\geq$2 datasets (Figure 4D, right). The intersections between lists of ORFs (or encoded proteins) observed to be linked to a metal across each dataset are represented in Figure 4E. Overall, the proteomic datasets captured the largest fraction of responses within the known metal-related proteins, followed by the metallomics study of overexpression mutants, the metallomics study of the deletion mutants, and the growth screen of the deletion mutants capturing the lowest signal for known metal-related genes (Figure 4F).

We then sought to determine to what extent the orthogonal nature of the five datasets improves the functional annotation of the yeast proteome. Understudied proteins (UniProt annotation score of 1 or 2) produced a similar number of hits in our datasets compared with well-studied genes (UniProt annotation score > 2), indicating that our resource can help mitigate annotation biases. Indeed, 470 poorly characterized proteins, including the 55 proteins that were functionally annotated in our ensemble clustering of the proteome, were a hit in at least one of the five datasets (Tables S19 and S20). To illustrate how the proteomic, metallomic, and genetic interactions provide signals for complementary sets of genes, we studied two examples of proteins that were identified in different datasets to generate hypotheses about their function: Ymr196wp and Ybr287wp.

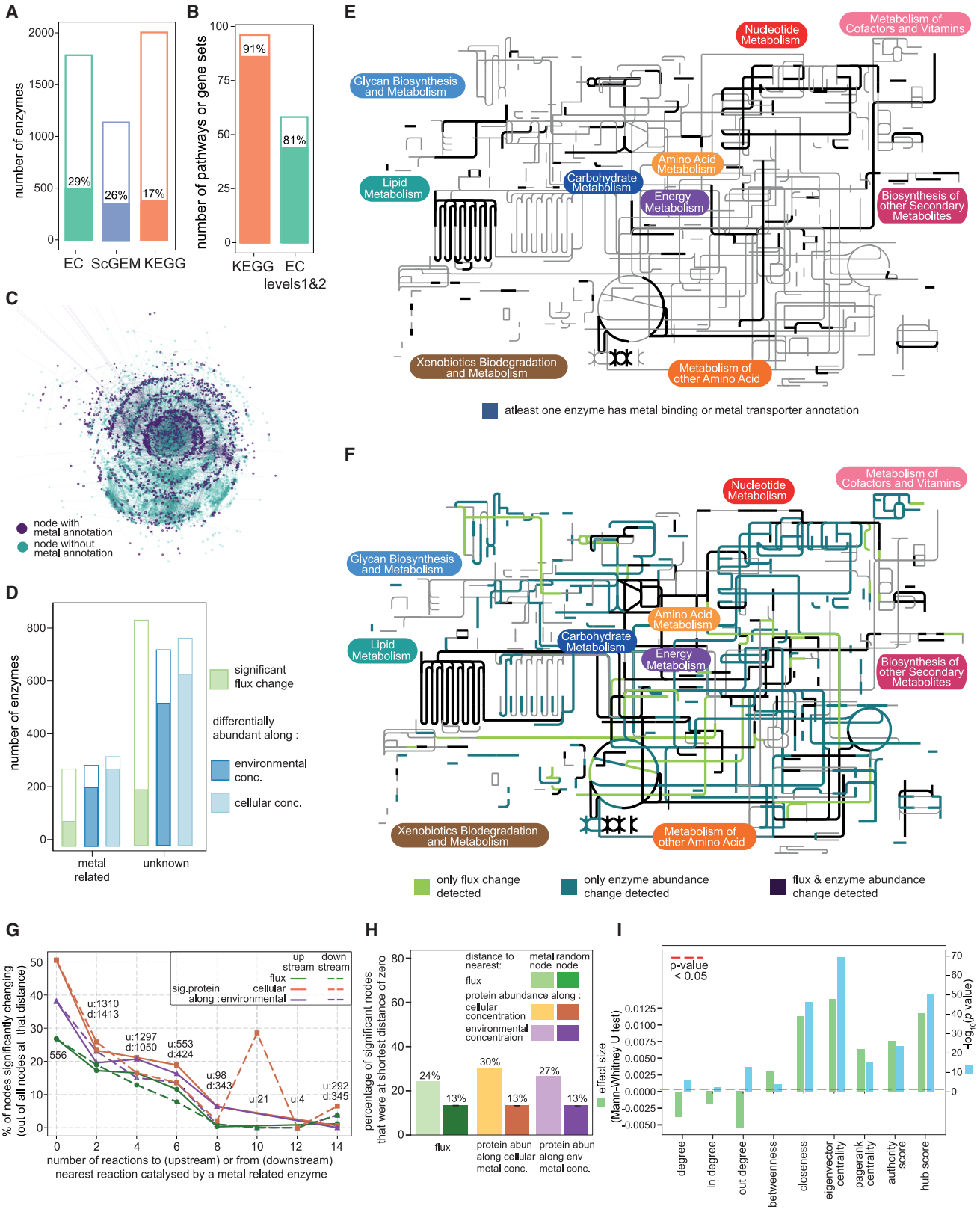
First, we report that Ymr196wp decreases in abundance when exposed to excess Fe (Figure 4G). The ensemble clustering pipeline assigned it to Fe cluster 10 based on metal-wise clustering and to cluster 23 in the all-metal clustering. Both clusters were enriched for functional terms related to oxidative stress and chemical stress, but Fe cluster 10 was also enriched for proteins that localize to the mitochondria (Figure 3A; Table S14). Overexpression of Ymr196wp led to an increase in the cellular Fe concentration (Figure 4H), whereas its abundance decreased upon the deletion of seven metal-binding proteins (Figure 4I). Proteomic profiles of the *S. cerevisiae* gene deletion collection showed that Ymr196wp was associated with the respiratory chain and the TCA cycle.[27] We thus posit from its comprehensive metal-linked molecular profile that Ymr196wp is linked to mitochondrial Fe metabolism.

Our second example, Ybr287wp, localizes to the endoplasmic reticulum[46] and contains eight transmembrane domains.[47,48] The abundance of Ybr287wp was correlated to environmental Ca, Cu, Fe, and Zn concentrations (Figure 4J) and was placed in a cluster enriched in Fe transport and cation channel terms (cluster 14; see Figure 3B) by our clustering analysis. Deletion of YBR287W resulted in decreased growth rates in the context of K and Mg depletion media, with milder changes observed in the context of Cu, Fe, Mn, and Na depletion media (Figure 4K). Ybr287wp deletion and overexpression mutants revealed mirrored cellular metal concentration profiles with altered cellular concentrations of Cu, Mn, Mo, and Na and minor alterations in Ca, Fe, and Zn (Figure 4L). Furthermore, the deletion of eight other metal-binding proteins led to the downregulation of Ybr287wp (Figure 4M), whereas the deletion of Ybr287w itself led to a significant decrease in the abundance of the Ca-binding protein Nth2 and smaller changes in Zn- and Mg-binding proteins (Figure 4N). Collectively, these data suggest that Ybr287wp is associated with metal biology, with a molecular profile suggestive of a promiscuous metal ion transporter.

Centrality of metal-requiring enzymes translates to high metal responsiveness of metabolism

Metabolism was identified by our functional enrichment analyses (Figures 2K, S2G, and S4B) as one of the cellular networks most affected by metal availability. We therefore selected the *S. cerevisiae* metabolic network to exemplify the utility of our dataset to assess the system-wide impact of metal availability. Annotations from various metabolic databases (GO, KEGG, EC, and Yeast8 metabolic model) suggest that between 13% and 29% of enzymes and ~13% of reactions in the genome-scale metabolic model (Table S21) are linked to at least one metal through direct binding, transport, or complex-metal-containing cofactor binding (Figures 5A left and 5B). Furthermore, 81% of all EC numbers and 91% of all KEGG pathways involve at least one metal-associated protein (Figure 5A, right). Oxidative phosphorylation contains the highest fraction of metalloenzymes (80%), closely followed by folate biosynthesis (75%) (Table S22).

To complement our protein-level changes, we simulated metabolic flux changes in response to metal depletion and excess using the CofactorYeast framework[20] (STAR Methods). Then, we represented the Yeast8 genome-scale metabolic model as a directed, bipartite graph and annotated the enzyme nodes with known metal related functional annotations (Figures 5C and 5D), the simulated flux changes, and the experimentally observed protein-level responses from our experiments (Figure 5E). The simulated flux-level responses were lower (32% of simulated reaction fluxes for metal-related enzymes and 25% for non-metal-annotated enzymes were changed by more than 50%) than those experimentally observed (~73% of both metal-related and non-metal-annotated enzymes) (Figure 5F). This indicates a potential flux-level buffering of the effects of



(legend on next page)

changes in metal availability enabled by compensatory changes in protein abundance.

Notably, over half of the nodes without metal annotations also showed significant protein abundance changes (Figure 5F), prompting us to examine whether this reflects still-missing metal-binding annotations or is a consequence of inherent network connectivity. We noted some trends between the shortest distance to or from a metal-requiring reaction node and the reaction nodes identified as showing significant simulated flux changes and protein abundance changes along cellular and environmental metal concentration gradients (Figure 5G). Although only 13% of reactions were annotated as being catalyzed by a metal-requiring enzyme, 46% of reaction nodes are one reaction away from a metal-dependent enzyme (Table S21) and 60% of metabolites are directly connected to, or just one reaction away from, a metal-requiring reaction node (Table S23). This close connectivity to metal-binding enzymes appears to be a consequence of the network structure, as calculating shortest distances to a random selection of 13% of reaction nodes produced similar results. Highly connected metabolites like ATP can skew biologically meaningful shortest distances by reducing average path lengths and creating dense connections. Even though we excluded H^+ and H_2O from the graphs before calculating shortest distances, over 50% of shortest distances in the network were just one reaction, likely due to other frequently occurring metabolites that would require hypothesis-specific manual curation to exclude. Therefore, we restricted further analysis to determining whether significantly altered fluxes and protein abundances directly coincided with metal nodes more than with random nodes. Our analysis confirmed that nodes with significant changes in flux and protein abundance were indeed more likely to directly coincide (distance zero) with metal-requiring nodes than with a random set of notes (Figure 5H). These results were robust to testing 10, 100, or 1,000 iterations of random node picking.

A feature of certain metal-requiring pathways is the presence of alternative enzymes or bypasses that help maintain flux through critical pathways during metal limitation. For example, upon Fe deprivation, *S. cerevisiae* cells switch from the Fe-dependent NADPH-dependent glutamate synthase (Glt1) to the NADH-dependent glutamate dehydrogenase (Gdh1 or Gdh2), enabling continued glutamate synthesis despite reduced Fe availability.[49] To explore the presence of enzyme pairs that catalyze the same reaction but show complementary metal responses, we identified enzyme pairs catalyzing the same reaction (isozymes) but showing different responses to metal perturbations. Of 246 unique metal-binding isozyme pairs, 43 showed anti-correlated abundance patterns (Figure S5B), i.e., we observed a simultaneous increase in the abundance of one isoenzyme and a decrease in the abundance of another, or vice versa (Table S24), suggesting complementary metal responses. For instance, Zn-binding *Adh3* and Fe/Zn-binding *Adh4* exhibit anti-correlated abundance profiles under Zn limitation (Table S24), with Fe^{2+}-dependent Adh4 being induced upon Zn limitation conditions and decreases in abundance in Fe limitation (Figure S5C). We identified 28 such isozyme pairs along the Zn perturbation series and 15 pairs along each of the Ca, Cu, Fe, and Na perturbation gradients, 11 of which were negatively correlated along more than one metal perturbation series (Table S24).

Although the presence of isozymes with differential metal requirements explains some of the observed protein abundance patterns along metal perturbation gradients, it does not fully account for the significant changes observed in over 60% of the metabolic network under metal perturbation, which far exceeds the fraction of metal-requiring nodes. Therefore, we turned to methods of analyzing node importance in networks that cannot be captured by shortest distances alone. We compared nine centrality metrics that quantify a node's influence across the

Figure 5. Central role of metal-dependent enzymatic reactions results in a network-wide metal dependency of metabolism

(A and B) The fraction of enzymes in the Enzyme Commission (EC) and KEGG databases and the Yeast8 genome-scale metabolic model (ScGEM) with annotated connections to metals. (A) Summarizes annotations at the protein level. The outer rectangle of the bars indicates the total number of proteins in the database; the inner bar indicates the number of proteins that have any metal-related annotations in the Gene Ontology molecular function (GO-MF) database. The percentage of proteins with metal-related functions in each database is given. (B) Summarizes annotations at the pathway and enzyme-class level. The outer rectangle of the bars indicates the number of pathways (for KEGG) and the number of unique EC numbers (up to level 2), whereas the inner filled bar represents the number of KEGG pathways or EC number categories for which at least one member protein had a metal-related annotation. The percentage of KEGG or EC categories that contained at least one protein with a metal-related annotation is given.

(C) Visualization of the directed bi-partite graph representation of the *S. cerevisiae* Yeast8 GEM. Each metabolite and reaction is represented by a unique node in the graph. Links in the graph are directed (from substrate metabolite to reaction and from reaction to product), with irreversible reactions added twice (to link each reaction - metabolite node in both directions).

(D) The fraction of the metabolic network that responds to perturbations of metal availability at the flux (simulated) and protein-abundance (experimentally quantified) levels. The outer rectangle of the bars represents the total number of enzymes measured or fluxes assessed in the simulation; the inner filled bar represents the number of enzymes for which a significant flux change through the enzyme was detected in the simulation (light green); a protein abundance change was detected in linear models fit along the environmental (cultivation media) metal concentration (dark blue) or the cellular metal concentration (light blue) for at least one metal. The percentage of fluxes or protein abundances that were significantly altered is given.

(E) Metal-related enzymes in the *S. cerevisiae* metabolic network.

(F) Metal responsiveness of the *S. cerevisiae* metabolic network, highlighting changes in simulated fluxes and measured protein abundance changes upon perturbations in metal availability.

(G) The relationship between the shortest distance from a reaction node to a metal-related reaction node (upstream), or from a metal-related reaction node to a reaction node (downstream), and the fraction of nodes at each distance in the metabolic network will be affected by perturbed metal availability. Zero indicates that the reaction node itself maps to at least one gene or encoded protein with a metal-related annotation. The percentage of fluxes or protein abundances that were significantly altered at each distance from the nearest metal-related reaction node is visualized.

(H) Metal-requiring nodes are more likely to coincide (same node or distance zero) with the reaction nodes showing significant simulated flux changes and significant protein abundance changes along cellular and environmental conditions than a random set of nodes.

(I) Centrality metrics of metal-related reaction nodes as compared with a randomly selected set of nodes. The green bars indicate effect size of each metric (metal-related nodes versus randomly selected nodes), whereas the blue bars indicate *p* values between metal-binding nodes and randomly selected nodes.

network (Figure 5I). For example, betweenness centrality measures the frequency with which a node appears on the shortest paths between other nodes, highlighting its role as a connector or bridge within the network. Our results indicate that nodes with metal-related annotations occupied more important positions in the metabolic network. Interestingly, the metal-binding enzyme nodes tended to have fewer substrate connections (lower degree) compared with a randomly selected set of nodes from the network (Figure 5I), suggesting that metal-requiring enzymes catalyze reactions with a lower-than-average number of substrates and products. However, all metrics of centrality and node importance that consider paths traversed through nodes suggested that metal-binding nodes were more important in the network compared with random nodes (Figure 5I). Thus, although metal-binding nodes do not directly participate in reactions involving a higher number of substrates or products, the reactions they participate in have a higher importance in the metabolic network. To summarize, our network analysis results suggest that a combination of the low average shortest path between nodes in the metabolic network and the high centrality of metal-requiring reaction nodes is the likely cause of the high metal responsiveness of the metabolic network.

DISCUSSION

The availability of metal ions not only varies across evolutionary periods and geological landscapes but also across ecological and physiological niches. At the cellular level, the importance of metal ions for metabolic and protein function and the variability in metal availability has driven the evolution of metal-ion sensing, transport, and buffering systems. Indeed, many transporters, chaperones, and metal-responsive transcriptional elements that maintain metal ion balance have been described.[5,7] Nevertheless, our understanding of how biological networks respond to physiologically relevant changes in metal availability on a broader scale, and the signaling systems contributing to metal ion homeostasis and their connection to cellular phenotypes, remains limited. Experiments conducted with metal concentrations that deviate from standard media formulations are highly underrepresented in molecular biology and systems biology experiments. A further limitation emerges from (1) the capacity of cells to buffer against environmental fluctuations, and (2) the promiscuity of metal ion transport and sensing systems disentangles extracellular from cellular metal ion levels. The concentration of metal ions in a cell does not correlate directly with concentration changes induced in the growth media, which greatly complicates the interpretation of experimental results obtained in any metal concentration perturbation experiment. To address the first limitation, we varied all major metallic components of minimal media for *S. cerevisiae* over five orders of magnitude in concentration. To address the second, we applied quantitative metallomics to compare cellular and environmental metal concentrations and to systematically detect the interactions between environmental and cellular metal concentrations in the context of eliciting cellular responses.

By using *S. cerevisiae* as our model organism, we could leverage prior knowledge to assess the responses of proteins with metal binding, metal transport, metabolic, or other functions to changes in metal availability and integrate uniquely available genomic data for yeast (e.g., genome-scale metallomic profiles,[26,44] proteomes of a genome-scale deletion mutant library,[27] and a genome-scale metabolic model that includes metal ions as cofactors[20]). Moreover, a key factor that distinguishes our study from previous systematic analyses of metal biology (besides its scale) is that we utilized a prototrophic *S. cerevisiae* strain that could grow in minimal media lacking amino acids. As a result, we could assay the impact of altered metal ion availability under conditions where metals fulfill one of their most important biochemical functions, i.e., as cofactors in biosynthetic enzymes.[5] These combined efforts yielded a comprehensive picture of the effect of metal ion concentrations on cellular networks and revealed, at the proteomic scale, the remarkable interdependence of cellular processes and their metallomic environment. Interestingly, although we observed high correlations between the cellular concentrations of different metals, the proteomic responses to these metal perturbations were highly specific to each metal. This suggests that despite interlinked fluctuations in cellular metal concentrations, each metal perturbation triggers distinct biological responses, likely due to the unique catalytic roles, specific sensing mechanisms, and biomolecular interactions associated with each metal.

Although we intend for our dataset to serve as a resource for the research community to study metal ion biology at various molecular layers, we have derived several general principles that govern cellular responsiveness to metal ion perturbations. For example, we have revealed that metal ion homeostasis strongly varies between metals and is evident only for those that are physiologically important. For instance, Mo was not buffered and elicited cellular responses only when present at high, toxic concentrations, whereas its depletion caused no growth defects or protein responses. We thus conclude that *S. cerevisiae* cells do not require Mo and so its addition to common synthetic yeast media formulation[50] is unnecessary. Indeed, our results also suggest that the concentrations of most essential metals in synthetic minimal media exceed what is required for normal cell growth, consistent with earlier conclusions.[51] Although this practice is not inherently problematic (as the routinely supplied concentrations of essential metals are within the physiological, non-toxic range), relevant phenotypes, such as ion transport defects, might be masked in experiments conducted in these media. We anticipate that decreasing the concentration of metal ions in standard media recipes might lead to new discoveries.

It is important to note that our measurements of intracellular metal concentrations represent an average of both cytoplasmic and organellar contents, including significant stores in the vacuole. The vacuole is a major reservoir for metals such as Zn, Fe, Mn, and Ca, and fluctuations in vacuolar storage and release can impact cytoplasmic metal levels and cellular homeostasis. Indeed, we find several vacuolar genes to influence metal ion responses, underlining the important role of metal ion storage in homeostasis and the buffering process. This implies that the buffering capacities we describe at the whole-cell level are a likely underestimation of the degree of buffering that occurs in specific organelles and the cytoplasm. In parallel, our comprehensive quantitative metallomic data have revealed the interdependence between the cellular concentrations of several metals. These results help not only the interpretation of our own data,

such as the proteomes, but could also provide key context for the interpretation of results of other metal ion perturbation experiments.

Overall, our study provides a comprehensive blueprint for understanding how cells adapt to variations in metal ion concentrations at the molecular level. The systematic nature by which we collected our data allowed us to unveil the comprehensive cellular response to changes in metal availability and how these responses are integrated across different molecular layers of the cell. Even though our experiments addressed the response to metal concentration changes in a single environmental condition and within a single genetic background of one eukaryotic species, we discovered that the abundance of approximately 60% of proteins is influenced by metal availability, with Zn, Fe, Ca, and Cu inducing the most widespread responses. We speculate that most biological responses are dependent on cellular metal availability. In this context, it is interesting that major components of cellular transcription and signaling machinery, especially most kinase pathways, are among the metal-responsive pathways. For example, proteins in 28/34 of the signaling pathways captured by our proteomes seem to change in abundance to several metal ions. Notably, these pathways included TOR,[52] a signaling pathway that functions at the crossroads of cellular transport, the lysosome, and energy metabolism, which are all processes related to metal biology. Additionally, our findings reveal an interplay between metal availability and membrane potentials. Alterations in metal ion concentrations led to substantial changes in the abundance of key components involved in the generation and maintenance of the proton motive force across various cellular membranes. This suggests that metal availability can influence membrane potentials, either through direct modulation or by necessitating compensatory changes in protein abundance to maintain homeostasis. We can exclude changes in growth rate as a main driver for most of these responses: proteomic responses explained by a change in growth rate were only detected in the case of Zn and K depletion and under conditions with extremely high levels of Fe and Cu and represent only a subset of the proteomic signature of altered metal availability.

Our resource also provides a fresh perspective on a common problem in molecular biology—the high number of understudied proteins. Even in the most well-studied organisms, approximately 20% of the proteome lacks functional annotation, and, for many more proteins, our functional annotation remains incomplete. We previously speculated that the limited number of experimental conditions tested in most laboratory experiments might be a leading cause of missing protein functional annotation.[53] The results of our present study support this conjecture. We find that understudied proteins, except for those with very low evidence for their expression (UniProt annotation score of 1), are as likely to be identified as being a hit (having a genetic or proteomic interaction with metal availability) across our datasets as proteins with high annotation scores. Many of the thus far unknown proteins thus have functions related to metal biology, and systematically varying metal ion levels can thus help to mitigate the annotation bias and provide testable hypotheses about new protein functions. We highlight two examples that are linked to metal biology across multiple datasets: Ymr196w, for which we obtained a molecular profile consistent with regulator of in Fe homeostasis, and Ybr287w, for which we obtained a molecular profile consistent with it being a promiscuous metal ion transporter.

The essential roles of metal ions in cellular biochemistry are thought to be derived from their functions as critical catalysts early in evolution of metabolic networks. For example, the central role of Fe(II)-based catalysis can be correlated to the fact that Fe(II) was readily available both as an electron donor and as a catalyst during planet Earth's period of low atmospheric oxygen levels.[54,55] The widespread availability of this catalyst likely shaped the evolution of the first metabolic networks, as suggested by the discovery that Fe(II) can catalyze non-enzymatic reactions in a manner that closely resembles the biochemistry underlying central metabolic pathways.[56–61] Similarly, the emergence of new protein domains throughout life has also likely been influenced by changes in the availability of metals on Earth's surface.[2,62] The fundamental importance of metal ions for metabolic reactions was reflected in the high responsiveness of the metabolic network across our analyses. Although metal-dependent reaction nodes comprise only ~13% of the metabolic network, they are present at key locations in the network (as reflected by a higher betweenness, closeness, and centrality), which, combined with the low average shortest distances between nodes in the metabolic network, is the likely cause of a high metal responsiveness at the proteome level. We speculate that this centrality of metal-related nodes stems from the use of metals as catalysts in early metabolic evolution, which is conserved as the core of modern-day metabolic networks. It is also likely that the evolution of metal-requiring enzymes is more constrained relative to other proteins due to the essential catalysis they enable. Finally, we report here a subset of reactions catalyzed by pairs of isoenzymes that catalyze similar reactions but use different metal cofactors and show negatively correlated protein abundance patterns, which suggests either divergent evolution or a mechanism to compensate for variability in metal availability for essential reactions. We posit that the central role of metalloenzymes, combined with dramatic changes in metal availability across time and ecological niche, is a driver of enzyme evolution.

Despite the systems-scale and quantitative nature of all our experiments, our study has limitations that need to be considered while interpreting and querying our dataset. The first is that we restricted the scope of our study to a single species, a single reference growth condition, and excluded metal toxicity. The biological response to metal ion perturbation is therefore more extensive than that reported herein. Second, we measured total cellular metal concentrations and did not address differences in metal ion concentrations between subcellular compartments, which can vary significantly and impact protein function within compartments. The buffering capacities we report here are likely modulated by the organelle-level reserves. For example, our data as well as previous literature assign metal buffering functions to the vacuole. Future studies could build on our work and assess the impact of subcellular metal distribution and resolve cellular responses to perturbed metal availability at the subcellular scale. Third, we faced analytical constraints while varying and quantifying certain metals. We varied K and Mg in this growth experiment for not more than one order of magnitude in order to maintain general properties of the media,

especially ionic strength. We made this choice because K and Mg salts are present at much higher levels than the other metals. Therefore, it is likely that we have missed the metal responsiveness of some genes, pathways, and processes relevant for K and Mg homeostasis. Fourth, and last, although our findings lend strong support to many proposed functions of poorly characterized proteins, the mutant libraries used to validate our hypotheses suffer from limitations common to such libraries, such as secondary mutations. Hence, we advocate for the use of our resource to derive system-level insights into the role of metals in biology and as a foundation for hypothesis generation to be validated by future studies.

Our comprehensive resource highlights the central role of metal ion homeostasis within the regulatory and functional landscape of the cell. We envision that this dataset will serve as a foundational reference for unraveling the connections between metal ions and a spectrum of biological processes, facilitating the integration of metal ions into a system-wide understanding of cellular function. Our dataset opens avenues for exploring the roles of previously understudied genes and proteins, enriching our comprehension of signaling pathways and gene regulatory networks.

Our findings also advocate for a paradigm shift in current laboratory practices that rarely vary metal ion concentrations within the physiologically relevant ranges. This practice starkly contrasts the dynamics of metal ion availability in natural settings, and our data show that the cellular biochemical network responds to these fluctuations. Thus, current laboratory practices both mask biological discoveries and potentially hamper the reproducibility of laboratory research. By varying and diligently reporting metal ion concentrations, researchers can unlock new biological insights and enhance experiment reproducibility.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to, and will be fulfilled by, the lead contact, Prof. Markus Ralser (markus.ralser@charite.de).

Materials availability

This study did not generate new unique reagents. All materials used for this work are already openly available.

Data and code availability

- All supplementary information and result data files are available at Zenodo (https://zenodo.org/records/10708992). All raw proteomic data files are available at ProteomeXchange: PXD057956.
- All code used to analyze data and associated files required as input for scripts is available at https://github.com/Ralser-lab/metallica/releases/tag/v1.0 (archived Zenodo:15047222). A table mapping each script to specialized analyses and each figure and supplementary figure is also available within this repository and as Methods Table S14.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

The authors would like to thank Benjamin Heineike, Christoph Messner, Lucia Herrera-Dominguez, Clara Correia-Melo, and Enrica Calvani for their support throughout the execution of this project; James Macrae, Luiz Carvalho, Jürg Bähler, and Gavin Kelly for key inputs. This work was supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (no. FC001134), the UK Medical Research Council (no. FC001134), and the Wellcome Trust (no. FC001134). We further acknowledge project funding from the Wellcome Trust (IA 200829/Z/16/Z) and the European Research Council ERC-SyG-2020 (951475).

AUTHOR CONTRIBUTIONS

S.K.A. and M.R. conceptualized and designed the study. S.K.A. and J.H. conducted growth rate measurements. S.K.A. and L.S. collected and analyzed ICP-MS measurements. S.K. and S.K.A. designed, conducted, and analyzed the deletion mutant growth screen. O.L. performed ensemble clustering analysis. Y.C. simulated metal perturbations using flux balance analysis. S.K.A. performed all other data analysis and visualization. M.M. provided key inputs for designing the study. S.K.A. and M.R. wrote the manuscript. All authors reviewed and revised the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL DETAILS
  - Strains and mutant libraries
- METHOD DETAILS
  - Cultivation of wild type *S. cerevisiae* cells
  - ICP-MS measurements of cultivation media
  - Growth rate measurement of WT *S. cerevisiae* cells
  - ICP-MS measurements of WT *S. cerevisiae* cells
  - Proteomics sample preparation
  - Liquid chromatography–mass spectrometry
  - Growth screen of deletion mutants
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Analysis of growth curves of WT cells
  - Analysis of ICP-MS data
  - Processing of LC-MS data using DIA-NN
  - Normalisation, batch correction, filtering, and protein quantification
  - Identification of proteins differentially abundant along environmental metal concentration
  - Identification of proteins differentially abundant along cellular metal concentration
  - Correlation analysis using proteomics and metallomics profiles
  - Focused analysis of metal-related proteins
  - Focused analysis of proteins involved in proton motive force generation
  - Ensemble clustering analysis
  - Gene set enrichment analysis
  - Analysis of growth screen of deletion mutants
  - Defining metal-related genes
  - Comparison and integration with published datasets
  - Simulations of metabolic flux using CofactorYeast
  - Metabolic network analysis using *igraph*
  - Publicly available databases used for analysis
  - Data visualisation

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cels.2025.101319.

Received: February 28, 2024
Revised: November 25, 2024

Accepted: May 19, 2025
Published: June 13, 2025

REFERENCES

1. Monteith, A.J., and Skaar, E.P. (2021). The impact of metal availability on immune function during infection. Trends Endocrinol. Metab. *32*, 916–928. https://doi.org/10.1016/j.tem.2021.08.004.
2. Aulakh, S.K., Varma, S.J., and Ralser, M. (2022). Metal ion availability and homeostasis as drivers of metabolic evolution and enzyme function. Curr. Opin. Genet. Dev. *77*, 101987. https://doi.org/10.1016/j.gde.2022.101987.
3. Andreini, C., Bertini, I., Cavallaro, G., Holliday, G.L., and Thornton, J.M. (2009). Metal-MACiE: a database of metals involved in biological catalysis. Bioinformatics *25*, 2088–2089. https://doi.org/10.1093/bioinformatics/btp256.
4. Bromberg, Y., Aptekmann, A.A., Mahlich, Y., Cook, L., Senn, S., Miller, M., Nanda, V., Ferreiro, D.U., and Falkowski, P.G. (2022). Quantifying structural relationships of metal-binding sites suggests origins of biological electron transfer. Sci. Adv. *8*, eabj3984. https://doi.org/10.1126/sciadv.abj3984.
5. Maret, W. (2016). The Metals in the Biological Periodic System of the Elements: Concepts and Conjectures. Int. J. Mol. Sci. *17*, 66. https://doi.org/10.3390/ijms17010066.
6. Gene Ontology Resource. http://geneontology.org/.
7. Cyert, M.S., and Philpott, C.C. (2013). Regulation of cation balance in Saccharomyces cerevisiae. Genetics *193*, 677–713. https://doi.org/10.1534/genetics.112.147207.
8. Ma, Z., Jacobsen, F.E., and Giedroc, D.P. (2009). Coordination chemistry of bacterial metal transport and sensing. Chem. Rev. *109*, 4644–4681. https://doi.org/10.1021/cr900077w.
9. Turco, G., Chang, C., Wang, R.Y., Kim, G., Stoops, E.H., Richardson, B., Sochat, V., Rust, J., Oughtred, R., Thayer, N., et al. (2023). Global analysis of the yeast knockout phenome. Sci. Adv. *9*, eadg5702. https://doi.org/10.1126/sciadv.adg5702.
10. Shakoury-Elizeh, M., Tiedeman, J., Rashford, J., Ferea, T., Demeter, J., Garcia, E., Rolfes, R., Brown, P.O., Botstein, D., and Philpott, C.C. (2004). Transcriptional remodeling in response to iron deprivation in Saccharomyces cerevisiae. Mol. Biol. Cell *15*, 1233–1243. https://doi.org/10.1091/mbc.e03-09-0642.
11. Shakoury-Elizeh, M., Protchenko, O., Berger, A., Cox, J., Gable, K., Dunn, T.M., Prinz, W.A., Bard, M., and Philpott, C.C. (2010). Metabolic response to iron deficiency in Saccharomyces cerevisiae. J. Biol. Chem. *285*, 14823–14833. https://doi.org/10.1074/jbc.M109.091710.
12. Puig, S., Askeland, E., and Thiele, D.J. (2005). Coordinated remodeling of cellular metabolism during iron deficiency through targeted mRNA degradation. Cell *120*, 99–110. https://doi.org/10.1016/j.cell.2004.11.032.
13. Navarrete-Perea, J., Guerra-Moreno, A., Van Vranken, J., Isasa, M., Paulo, J.A., and Gygi, S.P. (2021). Iron Deficiency and Recovery in Yeast: A Quantitative Proteomics Approach. J. Proteome Res. *20*, 2751–2761. https://doi.org/10.1021/acs.jproteome.1c00035.
14. Wang, Y., Weisenhorn, E., MacDiarmid, C.W., Andreini, C., Bucci, M., Taggart, J., Banci, L., Russell, J., Coon, J.J., and Eide, D.J. (2018). The cellular economy of the Saccharomyces cerevisiae zinc proteome. Metallomics *10*, 1755–1776. https://doi.org/10.1039/c8mt00269j.
15. Volkov, V. (2015). Quantitative description of ion transport via plasma membrane of yeast and small cells. Front. Plant Sci. *6*, 425. https://doi.org/10.3389/fpls.2015.00425.
16. Ke, R., Ingram, P.J., and Haynes, K. (2013). An integrative model of ion regulation in yeast. PLoS Comput. Biol. *9*, e1002879. https://doi.org/10.1371/journal.pcbi.1002879.
17. Aleksander, P., Piotr, A., Tadeusz, T., and Makarewicz, M. (2009). Accumulation and release of metal ions by brewer's yeast during successive fermentations. J. Inst. Brew. *115*, 78–83. https://doi.org/10.1002/j.2050-0416.2009.tb00347.x.
18. Wickerham, L.J. (1951). Taxonomy of Yeasts (United States Department of Agriculture).
19. Mülleder, M., Capuano, F., Pir, P., Christen, S., Sauer, U., Oliver, S.G., and Ralser, M. (2012). A prototrophic deletion mutant collection for yeast metabolomics and systems biology. Nat. Biotechnol. *30*, 1176–1178. https://doi.org/10.1038/nbt.2442.
20. Chen, Y., Li, F., Mao, J., Chen, Y., and Nielsen, J. (2021). Yeast optimizes metal utilization based on metabolic network and enzyme kinetics. Proc. Natl. Acad. Sci. USA *118*, e2020154118. https://doi.org/10.1073/pnas.2020154118.
21. Britigan, B.E., Rasmussen, G.T., and Cox, C.D. (1998). Binding of iron and inhibition of iron-dependent oxidative cell injury by the "calcium Chelator" 1,2-bis(2-aminophenoxy)ethane N,N,N′,N′-tetraacetic acid (BAPTA). Biochem. Pharmacol. *55*, 287–295. https://doi.org/10.1016/s0006-2952(97)00463-2.
22. Tang, Q., Jin, M.-W., Xiang, J.-Z., Dong, M.-Q., Sun, H.-Y., Lau, C.-P., and Li, G.-R. (2007). The membrane permeable calcium chelator BAPTA-AM directly blocks human ether a-go-go-related gene potassium channels stably expressed in HEK 293 cells. Biochem. Pharmacol. *74*, 1596–1607. https://doi.org/10.1016/j.bcp.2007.07.042.
23. Li, S., Crooks, P.A., Wei, X., and de Leon, J. (2004). Toxicity of dipyridyl compounds and related compounds. Crit. Rev. Toxicol. *34*, 447–460. https://doi.org/10.1080/10408440490503143.
24. Brachmann, C.B., Davies, A., Cost, G.J., Caputo, E., Li, J., Hieter, P., and Boeke, J.D. (1998). Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast *14*, 115–132. https://doi.org/10.1002/(SICI)1097-0061(19980130)14:2<115::AID-YEA204>3.0.CO;2-2.
25. Winzeler, E.A., Shoemaker, D.D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J.D., Bussey, H., et al. (1999). Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science *285*, 901–906. https://doi.org/10.1126/science.285.5429.901.
26. Danku, J.M.C., Gumaelius, L., Baxter, I., and Salt, D.E. (2009). A high-throughput method for Saccharomyces cerevisiae (yeast) ionomics. J. Anal. At. Spectrom. *24*, 103–107. https://doi.org/10.1039/B803529F.
27. Messner, C.B., Demichev, V., Muenzner, J., Aulakh, S.K., Barthel, N., Röhl, A., Herrera-Domínguez, L., Egger, A.-S., Kamrad, S., Hou, J., et al. (2023). The proteomic landscape of genome-wide genetic perturbations. Cell *186*, 2018–2034.e21. https://doi.org/10.1016/j.cell.2023.03.026.
28. Demichev, V., Messner, C.B., Vernardis, S.I., Lilley, K.S., and Ralser, M. (2020). DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. Nat. Methods *17*, 41–44. https://doi.org/10.1038/s41592-019-0638-x.
29. Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K., and Weissman, J.S. (2003). Global analysis of protein expression in yeast. Nature *425*, 737–741. https://doi.org/10.1038/nature02046.
30. Creamer, T.P. (2020). Calcineurin. Cell Commun. Signal. *18*, 137. https://doi.org/10.1186/s12964-020-00636-4.
31. de Nadal, E., and Posas, F. (2022). The HOG pathway and the regulation of osmoadaptive responses in yeast. FEMS Yeast Res. *22*, foac013. https://doi.org/10.1093/femsyr/foac013.
32. Zou, R., Wang, X., Li, S., Chan, H.C.S., Vogel, H., and Yuan, S. (2022). The role of metal ions in G protein-coupled receptor signalling and drug discovery. WIREs Comput. Mol. Sci. *12*. https://doi.org/10.1002/wcms.1565.
33. Knape, M.J., Ballez, M., Burghardt, N.C., Zimmermann, B., Bertinetti, D., Kornev, A.P., and Herberg, F.W. (2017). Divalent metal ions control

activity and inhibition of protein kinases. Metallomics *9*, 1576–1584. https://doi.org/10.1039/c7mt00204a.

34. Li, L., Bertram, S., Kaplan, J., Jia, X., and Ward, D.M. (2020). The mitochondrial iron exporter genes MMT1 and MMT2 in yeast are transcriptionally regulated by Aft1 and Yap1. J. Biol. Chem. *295*, 1716–1726. https://doi.org/10.1074/jbc.RA119.011154.
35. Lesuisse, E., and Labbe, P. (1995). Effects of cadmium and of YAP1 and CAD1/YAP2 genes on iron metabolism in the yeast Saccharomyces cerevisiae. Microbiology (Reading) *141*, 2937–2943. https://doi.org/10.1099/13500872-141-11-2937.
36. Mülleder, M., Calvani, E., Alam, M.T., Wang, R.K., Eckerstorfer, F., Zelezniak, A., and Ralser, M. (2016). Functional Metabolomics Describes the Yeast Biosynthetic Regulome. Cell *167*, 553–565.e12. https://doi.org/10.1016/j.cell.2016.09.007.
37. Ronan, T., Qi, Z., and Naegle, K.M. (2016). Avoiding common pitfalls when clustering biological data. Sci. Signal. *9*, re6. https://doi.org/10.1126/scisignal.aad1932.
38. Kapp-Joswig, J.-O., and Keller, B.G. (2023). CommonNNClustering—A Python Package for Generic Common-Nearest-Neighbor Clustering. J. Chem. Inf. Model. *63*, 1093–1098. https://doi.org/10.1021/acs.jcim.2c01493.
39. Lemke, O., and Keller, B.G. (2018). Common Nearest Neighbor Clustering—A Benchmark. Algorithms *11*, 19. https://doi.org/10.3390/a11020019.
40. Jin, X., and Han, J. (2010). K-Means Clustering. In Encyclopedia of Machine Learning, C. Sammut and G.I. Webb, eds. (Springer), pp. 563–564.
41. Arthur, D., and Vassilvitskii, S. (2007). K-means++: The advantages of careful seeding. Symposium on Discrete Algorithms, 1027–1035.
42. Traag, V.A., Waltman, L., and van Eck, N.J. (2019). From Louvain to Leiden: guaranteeing well-connected communities. Sci. Rep. *9*, 5233. https://doi.org/10.1038/s41598-019-41695-z.
43. Ward, J.H., Jr. (1963). Hierarchical Grouping to Optimize an Objective Function. J. Am. Stat. Assoc. *58*, 236–244. https://doi.org/10.1080/01621459.1963.10500845.
44. Yu, D., Danku, J.M.C., Baxter, I., Kim, S., Vatamaniuk, O.K., Vitek, O., Ouzzani, M., and Salt, D.E. (2012). High-resolution genome-wide scan of genes, gene-networks and cellular systems impacting the yeast ionome. BMC Genomics *13*, 623. https://doi.org/10.1186/1471-2164-13-623.
45. Iacovacci, J., Peluso, A., Ebbels, T., Ralser, M., and Glen, R.C. (2020). Extraction and Integration of Genetic Networks from Short-Profile Omic Data Sets. Metabolites *10*, 435. https://doi.org/10.3390/metabo10110435.
46. Huh, W.-K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., and O'Shea, E.K. (2003). Global analysis of protein localization in budding yeast. Nature *425*, 686–691. https://doi.org/10.1038/nature02026.
47. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2.
48. UniProt ybr287w https://www.uniprot.org/uniprotkb/P38355/entry.
49. Philpott, C.C., and Protchenko, O. (2008). Response to iron deprivation in Saccharomyces cerevisiae. Eukaryot. Cell *7*, 20–27. https://doi.org/10.1128/EC.00354-07.
50. YNB Media Sigma https://www.sigmaaldrich.com/GB/en/product/sigma/y0626.
51. Verduyn, C., Postma, E., Scheffers, W.A., and van Dijken, J.P. (1990). Energetics of Saccharomyces cerevisiae in anaerobic glucose-limited chemostat cultures. J. Gen. Microbiol. *136*, 405–412. https://doi.org/10.1099/00221287-136-3-405.
52. Dann, S.G., and Thomas, G. (2006). The amino acid sensitive TOR pathway from yeast to mammals. FEBS Lett. *580*, 2821–2829. https://doi.org/10.1016/j.febslet.2006.04.068.
53. Kustatscher, G., Collins, T., Gingras, A.-C., Guo, T., Hermjakob, H., Ideker, T., Lilley, K.S., Lundberg, E., Marcotte, E.M., Ralser, M., et al. (2022). Understudied proteins: opportunities and challenges for functional proteomics. Nat. Methods *19*, 774–779. https://doi.org/10.1038/s41592-022-01454-x.
54. Busigny, V., Planavsky, N.J., Jézéquel, D., Crowe, S., Louvat, P., Moureau, J., Viollier, E., and Lyons, T.W. (2014). Iron isotopes in an Archean ocean analogue. Geochim. Cosmochim. Acta *133*, 443–462. https://doi.org/10.1016/j.gca.2014.03.004.
55. Rouxel, O.J., Bekker, A., and Edwards, K.J. (2005). Iron isotope constraints on the Archean and Paleoproterozoic ocean redox state. Science *307*, 1088–1091. https://doi.org/10.1126/science.1105692.
56. Keller, M.A., Turchyn, A.V., and Ralser, M. (2014). Non-enzymatic glycolysis and pentose phosphate pathway-like reactions in a plausible Archean ocean. Mol. Syst. Biol. *10*, 725. https://doi.org/10.1002/msb.20145228.
57. Ralser, M. (2014). The RNA world and the origin of metabolic enzymes. Biochem. Soc. Trans. *42*, 985–988. https://doi.org/10.1042/BST20140132.
58. Keller, M.A., Kampjut, D., Harrison, S.A., and Ralser, M. (2017). Sulfate radicals enable a non-enzymatic Krebs cycle precursor. Nat. Ecol. Evol. *1*, 83. https://doi.org/10.1038/s41559-017-0083.
59. Keller, M.A., Zylstra, A., Castro, C., Turchyn, A.V., Griffin, J.L., and Ralser, M. (2016). Conditional iron and pH-dependent activity of a non-enzymatic glycolysis and pentose phosphate pathway. Sci. Adv. *2*, e1501235. https://doi.org/10.1126/sciadv.1501235.
60. Muchowska, K.B., Varma, S.J., Chevallot-Beroux, E., Lethuillier-Karl, L., Li, G., and Moran, J. (2017). Metals promote sequences of the reverse Krebs cycle. Nat. Ecol. Evol. *1*, 1716–1721. https://doi.org/10.1038/s41559-017-0311-7.
61. Mayer, R., and Moran, J. (2023). Metal Ions Turn on a Stereoselective Nonenzymatic Reduction of Keto Acids by the Coenzyme NADH. Preprint at ChemRxiv. https://doi.org/10.26434/chemrxiv-2023-jdhxk.
62. Dupont, C.L., Butcher, A., Valas, R.E., Bourne, P.E., and Caetano-Anollés, G. (2010). History of biological metal utilization inferred through phylogenomic analysis of protein structures. Proc. Natl. Acad. Sci. USA *107*, 10567–10572. https://doi.org/10.1073/pnas.0912491107.
63. Kamrad, S., Rodríguez-López, M., Cotobal, C., Correia-Melo, C., Ralser, M., and Bähler, J. (2020). Pyphe, a python toolbox for assessing microbial growth and cell viability in high-throughput colony screens. Elife *9*, e55160. https://doi.org/10.7554/eLife.55160.
64. Demichev, V. DIA-NN GitHub (Github) .
65. Darzi, Y., Letunic, I., Bork, P., and Yamada, T. (2018). iPath3.0: interactive pathways explorer v3. Nucleic Acids Res. *46*, W510–W513. https://doi.org/10.1093/nar/gky299.
66. Sprouffske, K., and Wagner, A. (2016). Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. BMC Bioinformatics *17*, 172. https://doi.org/10.1186/s12859-016-1016-7.
67. Väremo, L., Nielsen, J., and Nookaew, I. (2013). Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. Nucleic Acids Res. *41*, 4378–4391. https://doi.org/10.1093/nar/gkt111.
68. Tenenbaum, D., and Maintainer, B. KEGGREST: Client-side REST access to the Kyoto Encyclopedia of Genes and Genomes (KEGG). R package version. https://www.bioconductor.org/packages/devel/bioc/html/KEGGREST.html.
69. Kerseviciute, I., and Gordevicius, J. (2023). aPEAR: an R package for autonomous visualization of pathway enrichment networks. Bioinformatics *39*, btad672. https://doi.org/10.1093/bioinformatics/btad672.
70. viridis: Colorblind-Friendly Color Maps for R https://sjmgarnier.github.io/viridis/.

71. Neuwirth, E. (2022). ColorBrewer Palettes [R package RColorBrewer version 1.1-3].
72. Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). circlize Implements and enhances circular visualization in R. Bioinformatics *30*, 2811–2812. https://doi.org/10.1093/bioinformatics/btu393.
73. Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (Springer International Publishing).
74. Gehlenborg, N. (2019). A More Scalable Alternative to Venn and Euler Diagrams for Visualizing Intersecting Sets [R package UpSetR version 1.4.0].
75. Wagih, O., and Parts, L. (2014). gitter: a robust and accurate method for quantification of colony sizes from plate images. G3 (Bethesda) *4*, 547–552. https://doi.org/10.1534/g3.113.009431.
76. Pedregosa, F. (2011). Scikit-learn: Machine learning in python Fabian. J. Mach. Learn. Res. *12*, 2825.
77. Waskom, M. (2021). seaborn: statistical data visualization. J. Open Source Softw. *6*, 3021. https://doi.org/10.21105/joss.03021.
78. Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2.
79. Harris, C.R., Millman, K.J., van der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., et al. (2020). Array programming with NumPy. Nature *585*, 357–362. https://doi.org/10.1038/s41586-020-2649-2.
80. Csárdi, G., Nepusz, T., Horvát, S., Traag, V., Zanini, F., and Noom, D. (2023). Igraph (Zenodo). https://doi.org/10.5281/ZENODO.3630268.
81. Heirendt, L., Arreckx, S., Pfau, T., Mendoza, S.N., Richelle, A., Heinken, A., Haraldsdóttir, H.S., Wachowiak, J., Keating, S.M., Vlasov, V., et al. (2019). Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. Nat. Protoc. *14*, 639–702. https://doi.org/10.1038/s41596-018-0098-2.
82. Cherry, J.M., Hong, E.L., Amundsen, C., Balakrishnan, R., Binkley, G., Chan, E.T., Christie, K.R., Costanzo, M.C., Dwight, S.S., Engel, S.R., et al. (2012). Saccharomyces Genome Database: the genomics resource of budding yeast. Nucleic Acids Res. *40*, D700–D705. https://doi.org/10.1093/nar/gkr1029.
83. Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., and Tanabe, M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res. *40*, D109–D114. https://doi.org/10.1093/nar/gkr988.
84. Schomburg, I., Chang, A., Hofmann, O., Ebeling, C., Ehrentreich, F., and Schomburg, D. (2002). BRENDA: a resource for enzyme data and metabolic information. Trends Biochem. Sci. *27*, 54–56. https://doi.org/10.1016/s0968-0004(01)02027-8.
85. Kamrad, S., Bähler, J., and Ralser, M. (2022). High-Throughput, High-Precision Colony Phenotyping with Pyphe. Methods Mol. Biol. *2477*, 381–397. https://doi.org/10.1007/978-1-0716-2257-5_21.
86. Eide, D.J., Clark, S., Nair, T.M., Gehl, M., Gribskov, M., Guerinot, M.L., and Harper, J.F. (2005). Characterization of the yeast ionome: a genome-wide analysis of nutrient mineral and trace element homeostasis in Saccharomyces cerevisiae. Genome Biol. *6*, R77. https://doi.org/10.1186/gb-2005-6-9-r77.
87. Hanner, A.S., Dunworth, M., Casero, R.A., Jr., MacDiarmid, C.W., and Park, M.H. (2019). Elevation of cellular Mg2+ levels by the Mg2+ transporter, Alr1, supports growth of polyamine-deficient Saccharomyces cerevisiae cells. J. Biol. Chem. *294*, 17131–17142. https://doi.org/10.1074/jbc.RA119.009705.
88. Holmes-Hampton, G.P., Jhurry, N.D., McCormick, S.P., and Lindahl, P.A. (2013). Iron content of Saccharomyces cerevisiae cells grown under iron-deficient and iron-overload conditions. Biochemistry *52*, 105–114. https://doi.org/10.1021/bi3015339.
89. Demichev, V. diann-rpackage: Report processing and protein quantification for MS-based proteomics (GitHub). https://github.com/vdemichev/diann-rpackage.
90. Carlson, M., and Id, G. AnnotationDbi: Introduction To Bioconductor Annotation Packages. http://bioconductor.statistik.tu-dortmund.de/packages/2.11/bioc/vignettes/AnnotationDbi/inst/doc/IntroToAnnotationPackages.pdf.
91. AnnotationDbi Bioconductor. https://bioconductor.org/packages/release/bioc/html/AnnotationDbi.html.
92. org.Sc.sgd.db Bioconductor. https://bioconductor.org/packages/release/data/annotation/html/org.Sc.sgd.db.html.
93. GO.db Bioconductor. https://bioconductor.org/packages/release/data/annotation/html/GO.db.html.
94. Lemke, O., and Keller, B.G. (2016). Density-based cluster algorithms for the identification of core sets. J. Chem. Phys. *145*, 164104. https://doi.org/10.1063/1.4965440.
95. Keller, B., Daura, X., and van Gunsteren, W.F. (2010). Comparing geometric and kinetic cluster algorithms for molecular simulation data. J. Chem. Phys. *132*, 074110. https://doi.org/10.1063/1.3301140.
96. R Core Team (2025). Contributors https://www.r-project.org/contributors.html.
97. Plotly https://plotly.com/r/.
98. Conway, J.R., Lex, A., and Gehlenborg, N. (2017). UpSetR: an R package for the visualization of intersecting sets and their properties. Bioinformatics *33*, 2938–2940. https://doi.org/10.1093/bioinformatics/btx364.
99. Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: an information aesthetic for comparative genomics. Genome Res. *19*, 1639–1645. https://doi.org/10.1101/gr.092759.109.
100. Vlassis, N., Pacheco, M.P., and Sauter, T. (2014). Fast reconstruction of compact context-specific metabolic network models. PLoS Comput. Biol. *10*, e1003424. https://doi.org/10.1371/journal.pcbi.1003424.
101. Lu, H., Li, F., Sánchez, B.J., Zhu, Z., Li, G., Domenzain, I., Marcišauskas, S., Anton, P.M., Lappa, D., Lieven, C., et al. (2019). A consensus S. cerevisiae metabolic model Yeast8 and its ecosystem for comprehensively probing cellular metabolism. Nat. Commun. *10*, 3586. https://doi.org/10.1038/s41467-019-11581-3.
102. yeast-GEM: The consensus GEM for Saccharomyces cerevisiae (GitHub).
103. KEGG: Kyoto Encyclopedia of Genes and Genomes https://www.genome.jp/kegg/.
104. EC explorer - BRENDA Enzyme Database https://www.brenda-enzymes.org/ecexplorer.php?browser=1.
105. UniProt Consortium (2025). UniProt: The universal protein knowledgebase in 2025. Nucleic Acids Res. *53*, D609–D617. https://doi.org/10.1093/nar/gkae1010.
106. Yamada, T., Letunic, I., Okuda, S., Kanehisa, M., and Bork, P. (2011). iPATH interactive pathways Explorer. Nucleic Acids Res. *39*, W412–W415. https://doi.org/10.1093/nar/gkr313.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
Ammonium sulphate	Sigma Aldrich	204501-250G
Biotin	Sigma Aldrich	B4501-100MG
Calcium pantothenate	MilliporeSigma	C8731-25G
Folic Acid	MilliporeSigma	F8798-5G
Inositol	MilliporeSigma	I5125-50G
Nicotinic Acid	MilliporeSigma	73763-5G
p-Aminobenzoic acid	Sigma Aldrich	429767-1G
Pyridoxine HCl	MilliporeSigma	P9755-25G
Riboflavin	MilliporeSigma	R7649-25G
Thiamine HCl	MilliporeSigma	T4562-25G
Boric Acid	Sigma Aldrich	202878-10G
Copper Sulfate	Sigma Aldrich	451657-10G
Postassium Iodide	Sigma Aldrich	204102-10G
Ferric Chloride	Sigma Aldrich	701122-1G
Manganese sulfate	Sigma Aldrich	229784-25G
Sodium molybdate	Sigma Aldrich	737860-5G
Zinc sulphate	Sigma Aldrich	204986-10G
Potassium phosphate monobasic	Sigma Aldrich	229806-25G
Magnesium sulphate	Sigma Aldrich	203726-25G
Sodium Sulfate	Sigma Aldrich	204447-10G
Sodium phosphate monobasic	Sigma Aldrich	229903-10G
Sodium chloride	Sigma Aldrich	204439-20G
Calcium chloride	Sigma Aldrich	499609-10G
Glucose	Sigma Aldrich	G8270
DiPyridyl	Sigma Aldrich	D216305-10G
UltraPure™ Agarose	Invitrogen Thermo Fisher Scientific	16500500
HNO3 (Optima™, for Ultra Trace Elemental Analysis)	Fisher Chemical™	A467-1
Internal metal standards	Agilent	5188-6525
Metal standard - CGP10	Inorganic Ventures (EssLab)	CGP10
Metal standard - K-CGK10	Inorganic Ventures (EssLab)	K-CGK10
Metal standard - Ca-MSCA100PPM	Inorganic Ventures (EssLab)	Ca-MSCA100PPM
Metal standard - S-CGS10	Inorganic Ventures (EssLab)	S-CGS10
Metal standard - S-MS100PM	Inorganic Ventures (EssLab)	S-MS100PM
Metal standard - Zn-MSZN100PPM	Inorganic Ventures (EssLab)	Zn-MSZN100PPM
Metal standard - P-MSP-100PPM	Inorganic Ventures (EssLab)	P-MSP-100PPM
Metal standard - Mo-MSM010PPM	Inorganic Ventures (EssLab)	Mo-MSM010PPM
Metal standard - Mg-CGMG1	Inorganic Ventures (EssLab)	Mg-CGMG1
Metal standard - Na- CGNA1		N/A
Tuning solution	Agilent	5185-5959
PA solution	Agilent	5188-6524
8M urea	Honeywell Research Chemicals	33247H
Di-thio threitol (DTT)	Sigma Aldrich	43815
Iodacetamide (IAA)	Sigma Aldrich	I1149

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Ammonium bicarbonate	Sigma Aldrich	40867
Trypsin (Sequence grade)	Promega	Cat# V511X
Deposited data		
Growth rates in perturbed metal conditions	This study	Zenodo: https://zenodo.org/records/10708992
Measured metal concentration of cells	This study	Zenodo: https://zenodo.org/records/10708992
Raw proteome data	This study	http://proteomecentral.proteomexchange.org/ dataset ID = PXD057956.
Processed proteome data	This study	Zenodo: https://zenodo.org/records/10708992
Absolute protein abundance data	Ghaemmaghami et al.[29]	https://www.nature.com/articles/nature02046
Raw growth screen of knockouts	This study	Zenodo: https://zenodo.org/records/10708992
Processed growth screen of knockouts	This study	Zenodo: https://zenodo.org/records/10708992
Processed proteome data from Messner et al.	Messner et al.[27]	ProteomeXchange: PXD036062
Processed metallomics data	Iacovacci et al.[45]	https://www.mdpi.com/2218-1989/10/11/435
Experimental models: Organisms/strains		
BY4741+pHLUM library	Mülleder et al.[19]	https://pubmed.ncbi.nlm.nih.gov/23222782/
Prototrophic Saccharomyces cerevisiae deletion collection (MATa, restored prototrophy)	Winzeler et al.[25] and Mülleder et al.[19]	https://www.science.org/doi/10.1126/science.285.5429.901, https://pubmed.ncbi.nlm.nih.gov/23222782/
Software and algorithms		
Pyphe	Kamrad et al.[63]	https://github.com/Bahler-Lab/pyphe
DIA-NN	Demichev et al.[64]	https://github.com/vdemichev/DiaNN
iPath3	Darzi et al.[65]	https://pathways.embl.de/
TECAN		N/A
Agilent ICP-MS MassHunter Version C.01.05 (SI 588.15)	Agilent	https://www.agilent.com/en/product/atomic-spectroscopy/inductively-coupled-plasma-mass-spectrometry-icp-ms/icp-ms-software/icp-ms-masshunter-software
Scanners	pyphe	https://elifesciences.org/articles/55160
R Statistical Computing Software	The R Foundation	https://www.r-project.org/
RStudio	https://posit.co/download/rstudio-desktop/	https://posit.co/download/rstudio-desktop/
python	https://www.python.org/	https://www.python.org/
Anaconda	https://www.anaconda.com/	https://www.anaconda.com/
jupyterlab	https://jupyter.org/	https://jupyter.org/
MATLAB	https://www.mathworks.com/products/matlab.html	https://www.mathworks.com/products/matlab.html
growthcurver	Sprouffske and Wagner[66]	https://cran.r-project.org/web/packages/growthcurver/index.html
AnnotationDbi	Carlson & Maintainer, Bioconductor	https://bioconductor.org/packages/release/bioc/html/AnnotationDbi.html
org.Sc.sgd.db	Bioconductor	https://bioconductor.org/packages/release/data/annotation/html/org.Sc.sgd.db.html
GO.db	Bioconductor	https://bioconductor.org/packages/release/data/annotation/html/GO.db.html

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
piano	Väremo et al.[67]	https://bioconductor.org/packages/release/bioc/html/piano.html
KEGGREST	Tenenbaum and Maintainer[68]	https://bioconductor.org/packages/release/bioc/html/KEGGREST.html
plotly	Plotly	https://plotly.com/r/
aPEAR	Kerseviciute and Gordevicius[69]	https://github.com/kerseviciute/aPEAR
viridis	Garnier et al.[70]	https://sjmgarnier.github.io/viridis/
RColorBrewer	Neuwirth[71]	https://cran.r-project.org/web/packages/RColorBrewer/index.html
circlize	Gu et al.[72]	https://jokergoo.github.io/circlize_book/book/
ggplot2	Wickham et al.[73]	https://ggplot2.tidyverse.org/
UpSetR	Gehlenborg et al.[74]	https://cran.r-project.org/web/packages/UpSetR/index.html
gitter	Wagih and Parts[75]	https://github.com/omarwagih/gitter
scikit-learn	Pedregosa et al.[76]	https://scikit-learn.org/
seaborn	Waskom[77]	https://seaborn.pydata.org/
scipy	Virtanen et al.[78]	https://www.scipy.org/
numpy	Harris et al.[79]	https://numpy.org/
igraph	Csárdi et al.[80]	https://igraph.org/python/
leidenalg	Traag et al.[42]	https://github.com/vtraag/leidenalg
pyphe	Kamrad et al.[63]	https://github.com/Bahler-Lab/pyphe
COBRA Toolbox	Heirendt et al.[81]	https://opencobra.github.io/cobratoolbox/stable/
CofactorYeast	Chen et al.[20]	https://github.com/SysBioChalmers/CofactorYeast
Github repository containing all code related to this manuscript	https://github.com/Ralser-lab/metallica	https://doi.org/10.5281/zenodo.15047222
Other		
BreathEasy breathable seals	Sigma Aldrich	Z380059
Adhesive PCR plate foil	Thermo Scientific	AB0626
Short-well 96-well plates	Corning Costar Fischer Scientific	10695951
PVDF membrane plates (pore size 0.45um)	Agilent	200931–100
Spherical glass beads	Sigma Aldrich	G8772
Silicone mats	Spex SamplePrep	UY-04575-65
Solid Phase Extraction plates – BioPureSPN PROTO C18 MACRO 50–450 μL	The Nest Group	Cat#SNS SS18VL
Collection plate (for SPE)	Agilent	204355–100
Collection plate (for peptide extracts)	Waters	186005837
Lunatic plates	Unchained Labs	7012000
Cap mats	Spex	Cat#2201
Corning multiwell plates, plate lids and sealing mats	Sigma Aldrich	Cat#CLS3098
ABgene storage plates	Thermo Scientific	Cat#AB-0661
Spark-Stacker	TECAN	https://lifesciences.tecan.com/spark-stack
Biomek NXP	Beckman Coulter	https://www.beckman.com/liquid-handlers/biomek-nxp
Agilent 7900 ICP-MS	Agilent	https://www.agilent.com/en/product/atomic-spectroscopy/inductively-coupled-plasma-mass-spectrometry-icp-ms/icp-ms-instruments/7900-icp-ms

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
SPS 4 Autosampler for ICP-MS	Agilent	https://www.agilent.com/en/product/atomic-spectroscopy/inductively-coupled-plasma-optical-emission-spectroscopy-icp-oes/icp-oes-accessories/sps-4-autosampler
Spex Sample Prep MiniG 1600	Thermo Fisher Scientific	https://www.fishersci.co.uk/shop/products/minig-1600-homogenizer/15340422
Speedvac		N/A
Lunatic	Unchained Labs	https://www.unchainedlabs.com/lunatic/?gad_source=1&gclid=CjwKCAiAlcyuBhBnEiwAOGZ2S_-j7WznkNUZFZaxtEa4weCiao6_LuQQWrBoUOdksa1Im9NTYlmE8xoCACUQAvD_BwE
nanoAcquity UPLC System	Waters	https://www.waters.com/webassets/cms/library/docs/720001083en.pdf
nanoEasy M/Z HSS T3 column	Waters	https://www.waters.com/nextgen/en/shop/columns/186009249-nanoease-m-z-hss-t3-column-100a-18–m-300-mm-x-150-mm-1-pk.html
TriopleTOF 6600	Sciex	https://sciex.com/content/dam/SCIEX/pdf/customer-docs/user-guide/6600-system-user-guide-en.pdf
V800 PHOTO scanner	Epson	https://www.epson.co.uk/en_GB/products/scanners/consumer/perfection-v800-photo/p/13567?srsltid=AfmBOooJGN0qT27PGHb4gWbJjrdjpXdMc8a5Mi8irignGjUpS9deFp4n
Yeast reference proteome databases	Uniprot	https://www.uniprot.org
Full GO term annotation	Gene Ontology Consortium	http://current.geneontology.org/products/pages/downloads.html
GO slim terms	Cherry et al.[82]	https://www.yeastgenome.org/
KEGG	Kanehisa et al.[83]	https://www.genome.jp/kegg/
Summary of all S. cerevisiae growth screens	Turco et al.[9]	https://www.science.org/doi/10.1126/sciadv.adg5702
Brenda Enzyme Commission Numbers	Schomburg et al. (2002))[84]	https://www.brenda-enzymes.org/

EXPERIMENTAL MODEL DETAILS

Strains and mutant libraries

Saccharomyces cerevisiae (S288C) haploid (*MAT*a) was used as the experimental model system. Specifically, the *his3Δ* deletion strain from the prototrophic BY4741 + pHLUM library[19] was used as the wild type (WT) strain for the growth rate, cellular metallomics and proteomics experiments in metal perturbation media conditions. The library contains a centromeric plasmid containing the HIS3, LEU2, URA5 and MET17 genes to restore prototrophy in the BY4741 strain. The *his3Δ* deletion strain contains a kanamycin resistance cassette at the HIS3 locus, which is complemented by the pHLUM plasmid, rendering the strain metabolically prototrophic like the parent S299c strain. It was chosen as the WT strain for consistency with previous genome scale work.[27,36] The same[19] deletion mutant library was employed for the growth screen under metal depletion conditions on agarose media and the *his3Δ* deletion strain was used as the control strain reference to assess growth alterations for this screen.

METHOD DETAILS

Cultivation of wild type *S. cerevisiae* cells

Wild type (WT) *S. cerevisiae* cells were revived from cryostocks on YPD (Yeast Peptone Dextrose) agar plates and incubated at 30°C for ~24 hrs (until colonies appeared). A single colony was picked and streaked onto synthetic minimal (SM) media agarose plates and incubated at 30°C for ~36 hrs (until colonies appeared). Then, a single colony from the SM plate was used to inoculate a 5mL starter liquid SM culture, which was then incubated on a shaker at 30°C for ~36 hrs. The pellet from this culture was washed three times with

water and resuspended in SM media without the addition of metal salts (SM0 media) except for KH_2PO_4, Mg_2SO_4 and $ZnSO_4.7H_2O$ (Methods Table S1). The resuspended pellet was used to inoculate 300mL of SM0 media such that the OD_{600} of the culture at inoculation was 0.05. After incubation in a shaker at 30°C, the pellet from the SM0 culture was washed three times with water and used to inoculate deep-well 96-well plates (Eppendorf, 10052143) filled with each of the 91 cultivation media containing perturbed metal concentrations (Methods Table S1) such that the starting OD_{600} of the culture was 0.05. Four replicates of each cultivation condition were distributed across at least two different 96-well plate layouts (Methods Table S2). Each 96-well plate contained six control media (known as ''Allele media'') which were identical to SM but created in the same manner as all other media. The borders of each plate were filled with water instead of cultures to avoid edge effects that were observed in preliminary tests. Three wells in each plate were emptied and filled with technical controls for mass-spectrometry (MS) measurements post-cultivation. All media were prepared in plastic, all reagents used for preparation were of ICP-MS grade (except for glucose), and only deionized water that had previously been checked for metal contamination by ICP-MS was used for preparing media or washing the cells. All deep-well 96-well plates were covered with Breathe-Easy seals during cultivation.

ICP-MS measurements of cultivation media

All cultivation media bearing variations in metal concentrations were analysed by ICP-MS to quantify the concentration of each metal. A 17-point calibration series was freshly prepared (up to 24 hrs before measurement) using certified metal standards (key resources table; Methods Table S3). Plate layouts for ICP-MS measurements of media are in Methods Table S4. All samples were measured on an Agilent 7900 ICP-MS coupled to an SPS-4 auto-sampler and an Agilent MicroMist nebulizer. The instrument was operated with Nickel (Ni) cones and the measurement parameters were optimised using Tuning and PA solutions. The following gas modes were used for different metals: Helium (^{24}Mg, ^{59}Co, ^{63}Cu, ^{66}Zn), High Energy Helium (^{32}S, ^{31}P, ^{55}Mn), and hydrogen (^{39}K, ^{40}Ca, ^{56}Fe) mode. Details of all peristaltic pump settings, tune parameters and all raw data from the ICP-MS can be found in Methods Table S12.

Growth rate measurement of WT *S. cerevisiae* cells

WT *S. cerevisiae* cells were prepared as described above and inoculated into short-well 96-well plates with 180 μL of media in each well. The plates were then placed in a Spark-Stacker (TECAN) plate reader operating in kinetic mode for sequential acquisition of multiple 96-well plates. Absorbance at OD_{600} was calculated from the mean values of five multi-well reads obtained every 30 min for 48 hrs for each position.

ICP-MS measurements of WT *S. cerevisiae* cells

Immediately before inoculation, the media for cultivating cells for intracellular metal quantification with ICP-MS was filtered through a PVDF membrane plate (Agilent 200931–100) into fresh, deep-well (2 mL) 96-well plates to ensure no insoluble precipitates were present that could interfere with cell washing on PVDF membranes post-cultivation or the ICP-MS measurements. WT *S. cerevisiae* cells were prepared, inoculated into deep-well 96-well plates (Eppendorf, 10052143) containing metal perturbation media and cultivated on a shaker at 30°C for 24 hrs as described above. After cultivation, the cells were collected by filtering the cultures in each deep-well 96-well plate through a 96-well PVDF membrane plate. The yeast cells on the 96-well PVDF membranes were then washed three times with a solution composed of 10μM EDTA and 3μM TrisHCl. Centrifugation speeds and durations were modified as needed to ensure that the entire volume of each culture passed through the PVDF membrane. Then, the PVDF membrane plates bearing the washed cells were incubated on a hot plate at 70°C until completely dry. Internal metal standards were added to the membranes and the dried cell pellet was then digested by adding 60μL of HNO_3 and heating at 94°C for 40 min. Deionised water was added to each digested pellet to achieve a final HNO_3 concentration of 10% (v/v). Due to minor differences in evaporation rates, slightly different volumes remained in each plate after incubation at 94°C; therefore, the total amount of deionized water to be added and the final volume available for ICP-MS varied (Methods Table S5). The diluted cell extracts from one batch (96-well plate) at a time, together with the fresh calibrants, were measured on an Agilent 7900 ICP-MS coupled to an SPS-4 auto-sampler and an Agilent MicroMist nebulizer using the same methodology as described for '**ICP-MS measurements of cultivation media**'.

Proteomics sample preparation

WT *S. cerevisiae* cells were prepared, inoculated into deep-well 96-well plates (Eppendorf, 10052143) (Eppendorf, 10052143) containing metal perturbation media and cultivated at 30°C with 1000 rpm shaking (Heidolph Titramax incubator) for 24 hrs. Then, peptide extract preparation from cell pellets and measurement of mass spectrometry data was performed as previously described.[27] Briefly, after cultivation, 50μL culture was removed from each well and transferred into a transparent, short-well, 96-well plate pre-filled with 50μL H_2O for OD_{600} measurements. Each deep-well plate was centrifuged at 3220 rcf (Eppendorf Centrifuge 5810R) to pellet the cells. Then, the Breathe-Easy seals and the supernatant were removed, and the plates were sealed with adhesive PCR plate foil and a plastic lid and frozen at -80°C until further processing. Protein extraction and digestion was carried out in batches of four, 96-well plates (384 samples). To reduce batch effects, stock solutions (120 mM iodoacetamide, 55 mM DL-dithiothreitol, 9 μl 0.1 mg/ml trypsin, 2 μl 4x iRT) were prepared one time and stored at –80°C until required. Stock solutions of 7 M urea, 0.1 M ammonium bicarbonate, 10% formic acid were stored at 4°C. All pipetting steps were carried out with a Beckman Coulter Biomek NX^P liquid-handling robot.

To lyse the cells, 200 μl 7 M urea / 100 mM ammonium bicarbonate and glass beads (~100 mg/well, 425–600 μm) were added to the frozen pellet. Then, each plate was sealed with a silicone mat [Cap mats, (Spex) 2201] and the cells were lysed using a Geno-Grinder

(Spex) bead beater for 5 min at 2300 g. After centrifuging the plates for 1 min at 2300 g, 20 μl 55 mM dithiothreitol (DTT) was added and mixed to achieve a final concentration of 5 mM. The samples were then incubated for 1 h at 30°C before 20 μl 120 mM iodoacetamide was added (final concentration 10 mM) to each well. The plates were incubated for 30 min in the dark at room temperature before adding 1 mL 100 mM ammonium bicarbonate. Each plate was centrifuged for 3 min at 2300 g and then 230 μL of the supernatant was transferred to prefilled trypsin plates. The samples were incubated for 17 h at 37°C for trypsin digestion, after which 24 μl 10% formic acid was added to each well. The digestion mixtures were cleaned up using BioPureSPN PROTO C18 MACRO 96-well plates.

For solid-phase extraction, the samples were centrifuged for 1 min at various speeds (listed below) using an Eppendorf 5810R centrifuge 5810R. For the solid-phase extraction, each plate was conditioned with methanol (200 μl, centrifuged at 50 *g*), washed twice with 50% acetonitrile (ACN) (200 μl, centrifuged at 50 *g* and flow-through discarded), and equilibrated three times with 3% ACN, 0.1% formic acid (FA) (200 μl, centrifuged at 50 *g*, 80 *g*, 100 *g*, respectively, flow-through discarded). Finally, 200 μl of each trypsin-digested sample was loaded onto solid phase extraction column plates, centrifuged at 100 *g* and washed three times with 200uL of a solution of 3% ACN and 0.1% FA (centrifuged at 100 *g*). After the last washing step, the plates were centrifuged once more at 180 *g* before the peptides were eluted in three steps (eluted twice with 120 μL and once with 130 μL 50% ACN; centrifugation at 180 *g*) into a collection plate (1.1 mL, V-bottom). The collected material was completely dried in a vacuum concentrator (Concentrator Plus (Eppendorf)) and redissolved in 40 μL of the 3% ACN plus 0.1% FA solution before being transferred into a 96-well plate (700 μL round, Waters, 186005837). Quality control samples for repeat injections were prepared by pooling the digested and cleaned-up control samples from all the 96-well plates. To quantify the total peptide concentration, 2 μl of each sample were loaded onto Lunatic microfluidic 96-well plates (Unchained Labs) and measured with the Lunatic instrument (Unchained Labs). The total peptide concentration in each peptide extract was calculated using the Lunatic's native software from the absorbance value at 280 nm and the protein-specific extinction coefficient.

Liquid chromatography–mass spectrometry

The digested peptides were analysed on a nanoAcquity (Waters) running as microflow LC (5 μl/min), coupled to a TripleTOF 6600 (SCIEX). The yeast digests (2 μg; the injection volume was adjusted for each sample based on the measured peptide concentration) were injected and the peptides were separated in a 19-min nonlinear gradient (Methods Table S8) ramping from 3% B to 40% B (solvent A: 1% acetonitrile/0.1% formic acid; solvent B: acetonitrile/0.1% formic acid). A HSS T3 column (Waters, 150 mm × 300 μm, 1.8 μm particles) was used with a column temperature of 35°C. The DIA acquisition method consisted of an MS1 scan from 400 to 1,250 m/z (50 ms accumulation time) and 40 MS2 scans (35 ms accumulation time) with variable precursor isolation width covering the mass range from 400 to 1250 m/z. Rolling collision energy (default slope and intercept) with a collision energy spread of 15 V was used. A DuoSpray ion source was used with ion source gas 1 (nebuliser gas), ion source gas 2 (heater gas), and curtain gas set to 15 psi, 20 psi, and 25 psi, respectively. The source temperature was set to 0°C and the ion-spray voltage to 5,500 V. The measurements were conducted over a period of 2 months on the same instrument.

Growth screen of deletion mutants

Growth assays were performed to explore the contribution of each non-essential gene to fitness in the context of media depleted for each metal. The prototrophic *S. cerevisiae* haploid deletion collection, abbreviated as PHKo (Ko for knock-out)[29] was grown on 20 different media (corresponding to depletion of Ca, Cu, Fe, K, Mg, Mn, Mo, Na and Zn and the control; Methods Table S9) chosen after pre-tests in which the WT prototrophic *S. cerevisiae* BY4741+pHLUM strain was cultivated on agarose media containing various concentrations of metal salts. These media included two types of depletion for Ca and Fe [Ca omission, Ca omission with chelator EGTA (Ethylene glycol tetraacetic acid), Fe omission and Fe omission combined with the chelator dipyridyl (DiP)], three concentrations of K, Mg and Zn and three controls (SM media, SM with DiP and SM with EGTA). The PHKo library was revived from Yeast extract Peptone Dextrose (YPD) and glycerol stocks in 96-well plates frozen at -80°C on YPD-agar and then combined with a grid of the control strain for the library (BY4741+pHLUM *his3Δ*) into a 1,536-spot layout on SM-agarose in four different arrangements. The resulting assay contained four biological replicates of each strain in the PHKo collection.

As the growth of neighbouring strains may affect the colony size of a strain, the re-arrangement strategy allowed for neighbourhood effects on colony size to be considered before making inter-strain comparisons. For the reference grid, the BY4741+pHLUM *his3Δ* control strain was streaked out on SM agar and grown at 30°C for 2 days. A 24-hr culture of a single colony was made in 40 mL of liquid YPD media and pinned from a bath on YPD agar in 96-spot format and incubated at 30°C for 2 days. The PHKo library was revived from cryostocks in 384-spot format on YPD agar and incubated for 2 days at 30°C. A custom Singer ROTOR HD™ program was used to reshuffle the library (using 96 short pins) into four random arrangements, consisting of five plates each, on standard SM media. At this stage, the reference grid was combined with the library by pinning the reference strain colonies onto the combined plates into the A1, D4 and C2 sub-positions in 1,536-spot format. These combined plates, containing all the library strains and the reference grid, were cultivated at 30°C for 2 days and then copied onto fresh SM agarose plates to obtain a clean source plate with evenly spaced and sized colonies. Three copies of each combined plate were made, yielding 60 source plates in total, which were incubated at 30°C for 2 days. Finally, colonies from the source plates were transferred onto the assay agarose plates bearing different concentrations of metal salts, using the 'Replicate Many' program of the Singer ROTOR HD™ with the following settings: *recycle = Yes, revisit = Yes, source_pressure = 40%, source_pin_speed = 15 mm/s, source_overshoot = 1.5 mm, target_pressure = 25%, target_pin_speed = 13mm/s, target_overshoot = 1.2 mm, target_max = No, source_mix = No*. After 2 days of

incubation at 30°C, all plates were scanned on an Epson V800 PHOTO scanner in grey and transmission scanning mode at 600 dpi using *pyphe*.[85]

QUANTIFICATION AND STATISTICAL ANALYSIS

Analysis of growth curves of WT cells

Data from the TECAN Spark-Stacker were processed in R[64]. OD_{600} values of blank wells were subtracted from all sample wells before fitting sigmoidal growth curves using the *growthCurver*[66] R package.

Analysis of ICP-MS data

Raw data from the Agilent 7700 ICP-MS were processed using Agilent MassHunter™. The metal concentration [in parts per billion (PPB)] in each cultivation media and cell digest was calculated in MassHunter™ using measurements of the calibrants (a 17-point dilution series of certified element standards) and scandium (Sc) as the internal standard to correct for minor deviations of the instrument during measurement. These values were processed further in R. For the cultivation media measurements, data from each media were compared with values in the control media and visualised. For the cell digests, minor deviations in the total volume after the dried pellets on the PVDF membrane from each batch (Methods Table S5) were first corrected for; then, PPB values for each element in the blank samples was used to determine the Limit of Quantification (LOQ). The LOQ was defined as follows: mean (PPB) + 5*sd (PPB) of the signal of each metal in the blanks of each individual batch. For all elements other than Na in batches 2-8 and Cu in batch 4, the quality control sample had a mean (PPB) > LOQ; in total, 23/342 samples were filtered out from the cellular metallomics data (Figure S1F).

To correct for varying cell numbers in the cell digests, phosphorus (P) and OD_{600} normalisation strategies were compared. The P-normalised data was less variable across biological replicates. Therefore, after filtering based on LOQ, the P signal (which was observed to be stable and dependent on cell count) was used to normalise all other metals and the PPB values were scaled up to the original scale using the mean PPB values of the control samples, as follows:

$$Normalised\ PPB_{(m,x)} = \frac{PPB_{(m,x)}}{PPB_{(P,x)}} \times mean\left(PPB_{(P,belongs\ to\ SM\ control)}\right)$$

Batch correction was carried out such that the median value of each metal in the control samples was the same across batches. Data from samples with OD_{600} values > 0.1 were discarded. Nanograms per well values (1 ng/ml = 1 PPB) after P normalisation were used to compare metal quantities across samples. For buffering capacity calculations, the measured metal concentrations in cultivation media were combined with those measured in cell digests and each set was normalised to the metal quantities in control samples (SM media and cells cultivated in SM media, respectively). To estimate the atoms per cell, the cell number was estimated using OD_{600} values as previously described[80]:

$$Atoms\ per\ cell = \frac{6.022 \times 10^{23}}{atomic\ mass} \cdot pg\ per\ cell_{BC} \times 10^{-12}$$

where $pg\ per\ cell_{BC} = \frac{pg\ per\ cell}{median\ (pg\ per\ cell\ in\ controls\ of\ batch)} \times median\ (pg\ per\ cell\ across\ all\ batches)$

and

$$pg\ per\ cell = \frac{PPB\ of\ metal\ (ng/ml) \times 1000}{OD_{600} \times 1.8 \times 10^{7} \times volume\ of\ culture\ transferred\ to\ each\ filter\ plate}$$

The average coefficient of variation (CoV) across biological replicates of the ng/mL of digested cell cultures was 0.042, while the CoV across biological replicates of control samples was 0.032 (Methods Table S6) and the CoV of picogram per cell estimations across biological replicates of control samples was 0.13 (Methods Table S7). These values indicate that the OD_{600} estimates are noisier than the ICP-MS measurements and better cell counting methods are required to obtain reliable estimations of atoms of each metal per cell. For most metals, the obtained cellular concentrations were consistent with previous studies, except for Ca and Mn that were slightly higher than in previous reports[20,86–88] (Figure S1F)

Processing of LC-MS data using DIA-NN

All raw proteomics data (.wiff files) were processed using DIA-NN (Data-Independent Acquisition by Neural Networks[28,64]; version 1.8, compiled on 28 June 2021). The DIA spectral library (available at http://proteomecentral.proteomexchange.org/) dataset ID PXD036062[27](resubmitted to proteomeXchange with the dataset from this study, ID = PXD057956) and FASTA file for the UniProt yeast canonical proteome (downloadable from https://ftp.uniprot.org/pub/databases/uniprot/current_release/knowledgebase/pan_proteomes/UP000002311.fasta.gz) used were identical to those generated and described previously.[27] The DIA-NN parameters used to process data are described in Methods Table S13.

Normalisation, batch correction, filtering, and protein quantification

Quality control metrics exported by DIA-NN [including the number of identified precursors > 0.4*max(number of precursors identified in any individual file), number of proteins identified >1000, total signal quantity > 1000000, MS1 signal quantity > 1000000, MS2 signal

quantity > 10000000, normalisation instability < 0.5, Proteotypic ==1, Q value < 0.01, GG Q Value <= 0.01 and PG Q value < 0.01], the optical density measured at the end of cultivation (OD_{600} units sampled > 0.75) and manual inspection of the total ion chromatograms of certain problematic files (26 files) were used to filter data processed by DIA-NN. In addition, 11 conditions (Cu 50, Cu 100, Fe 100, K 10, Mg 20, Mg 50, Mo 20, Mo 50, Zn 2e-04, Zn 0.001, Zn 0.002) were deemed as unsuitable for inclusion based on growth rate and metallomics measurements of media. Samples corresponding to perturbations of H_3BO_4 that were acquired and processed with the dataset were excluded to focus the study on metals. In total, 266/437 proteome data files were retained.

Protein quantities were estimated from peptide quantities using maxLFQ [using the DIA-NN R function diann_maxlfq()[89]]. Batch correction (median scaling) was carried out using the median protein quantities of all control samples (WT yeast cells cultivated in SM media). No imputation was carried out before statistical analysis unless otherwise stated. After all processing steps, the average replicate CoVs for perturbation condition samples was ~15.7% with an average of 1,837 proteins quantified per sample. For the control samples alone, the replicate CoV was also ~15.7% with an average of 1,871 proteins quantified per sample.

Protein mass values were downloaded from the UniProt database (on 4th February 2024) and previously reported protein copy number values[29] were combined with these to calculate what fraction of the protein mass of the 3,841 proteins (for which protein copy number data was available) was measured or significant along any metal perturbation or cellular metal concentration series.

Identification of proteins differentially abundant along environmental metal concentration

Linear models with 0 (null model), 1, 2, and 3 degrees of freedom (dof) were fitted for each protein, modelling protein abundance as a function of the measured metal concentration in cultivation media i.e. protein abundance ~ poly (metal concentration, dof), where dof = 0, 1, 2 or 3. Only protein–metal combinations with at least four distinct points along the concentration gradient (rounded to 3 decimal spaces) and at least eight individual protein abundance measurements were used to fit the linear models. A series of F-tests was conducted between all combinations of fitted models, using the anova() R function. P values from each F-test were adjusted using the Benjamini-Hochberg correction for multiple testing. To choose the simplest model that explains the data, the following logic was used:

i. If the cubic (dof=3) model significantly outperformed all other models (dof=2, dof=1, and dof=0) i.e., adjusted p-value < 0.05, the cubic polynomial model was chosen.
ii. If the cubic model did not outperform the linear model (dof=1), and the quadratic model (dof=2) did not surpass the linear model, but the linear model was better than the null model, the linear model was preferred.
iii. If the cubic model was not better than the linear model (dof=1), and the quadratic model (dof=2) was not better than the linear model, but the linear model was better than the null model, the linear model was selected.
iv. If the cubic model was not better than the quadratic model, and the quadratic model was better than both the linear and null models, the quadratic model was chosen.
v. If none of the cubic, quadratic and linear models performed better than the null model, the null model was selected as the simplest model.

After selecting the least complex model, an additional threshold was applied for determining significant differential abundance along the metal perturbation series: proteins with a magnitude of fold difference (relative to control sample) change along metal perturbation series of at least 50% [i.e. $abs(max(\textit{fold difference along metal}) - min(\textit{fold difference along metal})) > log_2(1.5)$] and P value of the simplest model to explain the expression pattern < 0.05 were deemed significantly affected.

Identification of proteins differentially abundant along cellular metal concentration

To identify differentially abundant proteins along measured cellular metal concentrations, relative metal quantification values from all data (e.g., Fe values from Fe perturbation as well as along Mg, Zn, Ca etc perturbations) were binned into bins of size 0.01 (the metal concentration in each sample was normalised to that in control samples and rounded to two decimal places). For each measured metal and each protein, the median protein abundance values across the entire dataset corresponding to each bin along the measured cellular metal concentration was modelled as a function of the measured cellular metal concentration (as described above) to identify significantly differentially abundant proteins along environmental (media) metal concentration.

Correlation analysis using proteomics and metallomics profiles

The Spearman's rank-based correlation coefficient between each pair of samples within the metallomics dataset was computed in python 3 using the *scikit-learn* library.[76] The correlation coefficients were visualised as a heat map using the *seaborn* library.[77] The same methodology was followed for computing correlation coefficients between each pair of samples using proteomics data. The correlation coefficients computed based on proteomes and metallomes were then compared using the *pearsonr* and *spearmanr* functions from *scipy.stats*.[78]

Focused analysis of metal-related proteins

Gene function annotations in the gene ontology–molecular function (GO–MF) database (annotations were fetched using the *AnnotationDbi*,[90,91] *org.Sc.sgd.db*[92] and *GO.db*[93] R libraries) were used to annotate proteins associated with open reading frames (ORFs) in the databases as "metal binding proteins", "metal transport proteins" or "other metal related proteins".

Metallochaperones classified both as metal-binders and metal-transporters were considered "metal transport proteins" because their main function is to facilitate the incorporation of a metal into other metal-requiring proteins. Proteins for which a metal-specific annotation existed in the GO–MF were annotated with that metal, while those annotated for more than one metal were included in metal-binding or metal-transport lists for both metals and those that bore "metal binding" or "metal transport" annotations without any specific annotations for a metal were labelled "orphan".

Focused analysis of proteins involved in proton motive force generation

Gene function annotations in the gene ontology biological process (GO–BP) database (annotations were fetched using the *AnnotationDbi*,[90,91] *org.Sc.sgd.db*[92] and *GO.db*[93] R libraries) were used to shortlist proteins annotated with terms "proton transmembrane transport". These annotations were merged with subcellular localisation annotations from the gene ontology cellular compartment (GOCC) database. The subset of these proteins that showed significant alterations in protein abundance along environmental or cellular metal concentrations were visualised in Figures S2J and S2K.

Ensemble clustering analysis

An ensemble clustering framework was set up in *Python 3.9.13* (*numpy* 1.22.4,[79] *scikit-learn* 1.1.1,[76] *igraph* 0.9.9,[80] *leidenalg* 0.8.9,[42] *seaborn 0.12.0*[77] and *scipy* 1.8.1[78]), as previously described.[37] Briefly, the proteomics data were clustered into two parallel branches. The first, called *allmetal-clustering*, included proteins that were detected in at least 85% of all samples. The missing values in this dataset were imputed using the following imputation strategy:

i. If measured quantities for a protein were missing in all samples in a metal perturbation condition [e.g., Fe depletion (all samples cultivated in media containing lower Fe concentration than control SM media)], then the missing values for the protein in each sample corresponding to this perturbation were replaced with the minimum quantity of the protein detected in the entire dataset.
ii. If the protein was measured in at least one sample in the metal perturbation but missing in all replicates of a specific condition (e.g., protein detected in Fe 0.5, but absent in all replicates of Fe 0.1), the missing values were replaced with the median quantity measured for the protein in all the control samples (cells cultivated in synthetic minimal media).
iii. If the quantity of a protein was missing in only some replicates of a cultivation condition (e.g., two replicates of Fe 0.5 have missing values for a protein while the remaining two do not), the missing values were replaced by the median of the protein quantity of the replicates for which protein quantities were available.

For the second branch of ensemble clustering, called *metalwise-clustering*, a completeness filter of detection in at least 60% of all samples along each metal perturbation series was used before the imputation was carried out as described above. Of note, an additional filter was applied before performing the ensemble clustering analysis—only those proteins that were detected as being differentially abundant along either the environmental (media) metal concentration or along the measured cellular metal concentration were retained.

Three separate clustering algorithms were then used to cluster the proteomics data within each parallel clustering pipeline:

i. CommonNN[38,39,94,95] — the parameter R was varied equal-distant between 0.5**R_cut* and *R_cut*, where R_cut was defined as the distance at which around 5% (2.5% for Mg and Mn due to smaller data set size) of the distances are smaller than *R_cut* (10 Rs), the number of shared nearest neighbours *N* was varied between 2 and 10 (step size 1; 9 Ns), the minimal cluster size *M* was set to 5. In total 90 R, N- combinations were used.
ii. kMeans++[40,41,76] — the cluster number *k* was varied between 10 and 98 with a step size of 2 (between 10 and 55 with a step size of 1 for Mg and Mn). In total, there were 45 clustering steps.
iii. Leiden[42] — the graph was set up using the 10 closest neighbours for each data point, with edge weights of 1-scaled(distance), and all distances scaled between 0 and 1.

For reproducibility, the seed was fixed to 42 for all clustering analyses. For each clustering algorithm a co-clustering matrix was calculated, where each element denotes the probability that two data points were clustered together. The co-clustering matrices from each clustering algorithm were combined into a single matrix, using equal weights. The final clusters were obtained by hierarchical clustering of the combined co-clustering matrix (Ward clustering[43,78]), using a linkage-based cutoff for cluster extraction. Results from the clustering analysis were exported to R([96] for gene-set enrichment analysis and visualisation as described above. The functions are available at: https://github.com/OliverLemke/ensemble_clustering and within the code folder associated with this publication.

Gene set enrichment analysis

All gene set enrichment analyses (except for the network plots used to visualise clusters resulting from the ensemble clustering analysis shown in Figures 3D and 3E) were performed in R[64] using the *piano*[67] library and gene sets defined using the GO database (terms fetched as described above), KEGG database (fetched using the *KEGGREST*[68] library) and GOslim annotations downloaded from the Saccharomyces Genome Database (http://sgd-archive.yeastgenome.org/curation/literature/go_slim_mapping.tab). The *piano::runGSAhyper()* function was used to carry out the enrichment analysis with gene set size limits of 3 (lower limit) and 400 (upper

limit), using all proteins or ORFs (in the case of genes) quantified as the background. The Benjamini-Hochberg method was used to correct P values of enrichment for multiple testing. Gene set terms with a P value (adjusted) < 0.05 were considered significant. The results were visualised as Sankey plots (Figures 2K, 3C, S2G, and S4B) using the *plotly*[97] R library. Gene set enrichments and visualisation for Figures 3A and 3B were conducted using the aPEAR R library.[69]

Analysis of growth screen of deletion mutants

Images of agarose plates from the Epson V800 PHOTO scanner were processed using the *Gitter* R library to extract the colony size. The.dat output from *gitter*[75] were combined with the experiment design table and analysed further (grid normalisation, data aggregation, quality control checks and statistical analysis to obtain effect sizes and P values) using *pyphe*.[85] Only 1% of the negative control positions (footprints) (49/4,875 empty spots in total across all agarose plates) were contaminated and no systematic contamination was observed. One plate (corresponding to the reduction of K to 1/50th the level in synthetic minimal agarose media) contained 13 contaminated footprints and was therefore excluded from further analysis. Correlation between replicates of the control strain within a single plate was 0.78 for the raw colony sizes and increased to 0.95 after correction for surface effects using the control strain grid. Next, *pyphe-interpret* was used to obtain effect sizes of each mutant relative to the control strain (*Δhis3* from the haploid prototrophic library we used). The P values derived from Welch's t-test for samples with unequal variance were corrected for multiple testing using the Benjamini-Hochberg method. In total, 357,972 colony size measurements, corresponding to 4,759 unique deletion (and control) strains across 17 cultivation media conditions, remained after data processing. A P value (adjusted) threshold of < 0.10 and an abs($\log_2$(effect size)) > $\log_2$(1.2) was chosen to determine which mutants showed significantly altered growth in each condition. Data from different levels of depletion of K, Mg and Zn and from Fe depletion and Fe depletion combined with the dipyridyl chelator, were combined at this stage to determine the final list of ORF deletions that were affected to enable a metal-wise comparison with all other datasets.

Defining metal-related genes

Molecular Function gene ontology (GO) annotations[93] were used to determine a set of metal-binding proteins, metal-transport proteins and other metal-related genes that do not fall under the first two categories (e.g., "calcium-dependent protein kinase C activity"). Proteins encoded by ORFs mapping to GO terms containing the word "binding" and any word reflecting involvement of Ca, Cu, Fe, Mg, Mn, Mo, Na, Zn, heme or protoheme were included in the metal-binding annotation set with specific annotations. Proteins encoded by ORFs mapping to terms containing the words "metal binding" but no indication of which metal or metal-containing group is bound were annotated as "orphan".

Metal transporters and other metal-related proteins were filtered from the GO molecular function database using a manual creation process; text parsing using regular expressions was not sufficient to include only metal-related transporters or the metal-dependent enzyme activities included in the "other metal related" set. A list of these terms is available within the code repository associated with this publication.

Comparison and integration with published datasets

The metallomics data we used for integration with our results were the Z-scores of the cellular metal concentration in each deletion and overexpression mutant reported by Iacovacci et al.,[45] which were calculated using metal concentration measurements collected by Danku et al.[26] (cellular metal concentration in each haploid deletion mutant of *S. cerevisiae*) and Yu et al.[44] (overexpression mutants). Any Z-score with an absolute value > 1.959 (corresponding to *P* value < 0.05) was annotated as a significant change in metal concentration in a mutant. Protein abundance data in haploid deletion mutants were sourced from Messner et al.[27] This dataset was filtered to retain only deletion mutants of genes known to be connected to specific metals in the GO molecular function database-based mapping described above, and any protein quantified in these mutants with absolute $\log_2$(fold difference vs. control) > $\log_2$(1.5) and a P value < 0.05 were considered as being significantly altered. The upset plot (Figure 4E) to visualise commonalities between genes identified as being connected to metals were created using the upSet[98] R package. UniProt annotation status annotations were merged with all the datasets to assess how many poorly characterised genes were identified as significantly affected in each dataset. The circular plot to summarise current annotation status, the metal-binding and metal transporter annotations and all the datasets were created using the *circos*[99] R library. Gene set enrichments shown in Figure S4B were conducted and visualised as described above.

Simulations of metabolic flux using CofactorYeast

Simulations of excess and depletion of each metal were carried out in MATLAB using the *cobratoolbox*[81,100] and the CofactorYeast framework.[20] CofactorYeast incorporates metal ion cofactors and import and export reactions for each metal ion into the Yeast8 genome-scale metabolic model.[101] Growth rates were fixed to the lowest experimentally measured growth rate upon each metal perturbation (e.g., Fe depletion or Zn excess). The protein abundances of each enzyme were allowed to vary to minimise or maximise metal uptake. A flux balance analysis was carried out to simulate the fluxes required to achieve the objective function (minimisation or maximisation of the metal transport) under the growth rate constraint. The flux results (Supplementary Information – Results) were then processed in R([96] to calculate the flux change values (flux in perturbation condition / flux in control condition). This processing resulted in several infinite values due to flux = 0 of either the control condition simulation or the perturbation condition simulation. Therefore, infinite flux changes values of conditions for which both the control condition and simulation condition flux was 0, was

set to 1. For conditions where the control condition flux was not zero, but the test condition flux was zero were set to a low value with the correct direction (i.e., $log_2(fold\ change\ flux) = -sign(control\ condition\ flux)$ and for those where the control condition flux was zero, but the perturbation flux was nonzero, it was set to $log_2(fold\ change\ flux) = sign(perturbation\ flux)$. Reactions through which $abs(log_2(fold\ change\ flux)) > log_2(1.5)$ were considered significantly affected.

Metabolic network analysis using *igraph*

The Yeast8 metabolic model[101,102] was used as the input to create a directed, bipartite graph using the *igraph*[80] python library. Reaction IDs and metabolite IDs were used as the two types of nodes. Directed edges between the nodes were from each substrate metabolite node to each reaction node and from each reaction node to each product node. As 1,670 of the 4,131 reactions in the Yeast8 model were reversible, directed edges were added to the graph in both directions for these cases. A slight imbalance was noted in the fraction of reversible reaction nodes mapping to at least one metal-linked enzyme versus those that did not: only 21.38% of the metal-linked reaction nodes were reversible while 43.71% of those without metal-linked annotations were reversible. Therefore, calculations on the *igraph* that were grouped using the metal-linked annotation have a slight bias of counting nodes without metal annotations more often than metal-requiring nodes. The final graph was not a fully connected graph as 99.55% of nodes would be retained if the graph was filtered for its largest weakly connected component and 88.47% nodes would be retained in the largest strongly connected component. Weakly or strongly connected components were determined by the *clusters (mode = "weak")* and *clusters (mode = "strong")* functions from the *igraph* python library. The graph was visualised, and all centrality metrics were calculated using *igraph* functions. For the control for network based metrics, an equal number of nodes as metal-binding nodes in the test graph were picked randomly 1000 times and centrality metrics were calculated on each iteration. We also tested the same analysis after removing H+ and H_2O from the metabolic network as connectivity through these nodes is not biologically meaningful for most metabolic pathways. We obtained similar results for all centrality metric calculations except for closeness which could not be calculated due to the graph becoming more disconnected as well as hub score which showed no difference between the metal nodes and the randomly selected nodes. To calculate distances to and from nearest metal requiring nodes (Figures 5G and 5H), we used the H^+ and H_2O filtered network (graphical representation in Figure S5B) to have a more robust analysis as shortest distance calculations are more susceptible to being skewed by hyperconnected metabolite nodes than overall graph centrality metrics.

Publicly available databases used for analysis

Gene Ontology database[93]
Kyoto Encyclopedia of Genes and Genomes[103]
Saccharomyces Genome Database[92]
Brenda Enzyme Commission Numbers[104]
UniProt[105]

Data visualisation

All data visualisations except Figures 5B, 5C, and 5G were created in R using the *ggplot2*,[73] *plotly*,[97] *viridis*,[70] *aPEAR*,[69] *RColorBrewer*[71] and *circlize*.[72] The bipartite, directed network in Figure 5C was visualised using the *igraph*[80] python library and Figures 5B and 5G were created using iPATH.[106] Adobe illustrator and Biorender were used to assemble some figures.