Supplementary Note S9: Rationale behind the definition/creation of functional networks of co-expressed genes in this study

In this study, the data analyses follow several steps. First, we searched for differentially expressed genes, *i.e.* genes for which average logFC values were significantly different from 0. This is a classical statistical analysis performed in any transcriptomics project. Dedicated R packages exist (e.g. LIMMA) and we directly applied them following standard protocols. As a result, we obtained a set of 637 genes, which are referred to in the main text as "iron responsive genes". To be considered good candidates for being involved in the iron homeostasis processes, the biological relevance of this list of genes had to be assed. For that, we browsed the functional descriptions associated to the genes (available in CGD and GRYC databases) and searched for significantly enriched GO terms. Again, these are very common procedures in transcriptomics projects. We were satisfied to find genes and functions consistent with current knowledge of yeast cell adaptations to iron deficiency or iron overload. But to go further, we needed to disconnect our exploration of the iron responsive genes from the biological knowledges we already had. In this context, the data visualization methods are very useful. They helped us to change our perception of the data and imagine new analyses.

We decided to visualize our entire list of iron responsive genes in the form of a network, searching for unexpected properties in the relationships between genes, or in the transcriptome overall dynamics between low and high iron conditions. We quantified the co-expression relationships between gene expression profiles and obtained the result shown here https://thomasdenecker.github.io/iHKG/all.html. In this network, we observe up- and down-regulated genes, respectively in low or high iron conditions. This is the global picture of the transcriptome changes that occurs in the cell, but not only. Indeed, it is important to emphasize that in this network, co-expressed genes are represented by neighbouring nodes. This is an interesting property since co-expressed genes are generally considered as good candidates to be involved in common biological processes or to be regulated by the same factors. A relevant way to take advantage of our network is therefore to search for undescribed (or functionally unannotated) genes, which are in the immediate neighbourhood of a well-known iron responsive gene.

In that respect, we decided to split the global network of co-expressed genes (which comprises the 636 iron responsive genes) into six functional sub-networks, referred to as "Metabolism", "Regulation", "Redox Signaling", "Transport Trafficking", "ISC synthesis and assembly" and "Others" (https://thomasdenecker.github.io/iHKG/). Our aims here were to (*i*) limit the number of genes in each network, therefore facilitating the data visualization and exploration and (*ii*) provide a global picture of the cellular functions in which substantial gene expression rewiring was observed. We therefore had to define a limited number of functional categories, including a maximum number of the 636 genes. To avoid confusing interpretations of the functional networks of genes, we decided to assign each gene to only one function.

Functional categories were created as explained in the main text. Of course, different choices for the functional assignment of genes could have been made. But we succeed in clustering a large majority of the 637 iron responsive genes into only 5 functions. The reader must keep in mind that our purpose was (only) to separate the iron responsive genes into coherent sub-networks of genes to simplify their exploration. Other choices could have been made, but they are, in any case, totally disconnected from the definition and the identification of the iron homeostasis key genes (iHKG) presented in this article.