Ant Colony Optimization for predicting gene interactions from expression data

Ester Moncho Escrivà

Master's Degree in Bioinformatics and Biostatistics.

Programming for Bioinformatics.

Tutor UOC: Brian Jimenez

PRAs: Alexandre Sánchez Pla, Antoni Pérez Navarro, Carles Ventura Royo, David Merino Arranz, Jose Antonio Morán Moreno, Maria Jesús Marco Galindo.

24th May, 2017

FINAL WORKSHEET

Project Title:	*Ant Colony Optimization for predicting gene interactions from expression data*
Author's name:	*Ester Moncho Escrivà*
Tutor's name:	*Brian Jimenez*
PRA's name:	*Alexandre Sánchez Pla, Antoni Pérez Navarro, Carles Ventura Royo, David Merino Arranz, Jose Antonio Morán Moreno, Maria Jesús Marco Galindo*
Deadline:	*05/2017*
Degree's name:	*Master's Degree in Bioinformatics and Biostatistics*
Area of Final Project:	*Programming for Bioinformatics* *Swarm intelligence – Algorithms inspired in nature*
Language:	*English*
Keywords:	*Ant colony optimization, gene interaction network, gene expression data*

Abstract:

Inferring gene interaction networks from gene expression data is an important task in systems biology research. Gene interaction networks play an important role to understand how cells work internally, identifying biomarkers for diseases which in turn, can help in drug design.

Ant Colony Optimization is an algorithm from the Swarm Intelligence family that has been used for optimization problems including prediction of key gene interactions from expression data. In this Final Project, ACO has been modified and used for predicting more than key interactions in gene interaction networks from expression data.

Some modifications have been applied to ACO along with the approach of how to use it. It has been modified to accept correlation matrix from expression data as input, and as output it returns a list of gene interactions. Also, the approach has been modified, instead of resolving Travelling Salesman Problem and going through each node once, it has been modified in order to choose how many times we want the agent (ant) to visit each node (gene).

The modified ACO (ACOGeneInteraction) has been tested in two datasets and the results show that the modified ACO allows us to predict more than key interactions and also, allows us to identify central genes in the studied networks.

Resum:

Predir interaccions a la de xarxa d'interacció de gens a partir de dades d'expressió gènica és una tasca important en la investigació de la biologia de sistemes. Les xarxes d'interacció de gens juguen un paper important per entendre com funcionen les cèl·lules internament, així com per a la identificació de biomarcadors de malalties que poden ajudar, a la seua vegada, en el disseny de fàrmacs.

Ant Colony Optimization és un algoritme basat en la intel·ligència d'eixam que s'ha utilitzat per a resoldre problemes d'optimització de grafs, incloent la predicció d'interaccions clau a partir de dades d'expressió gènica.

En aquest Treball Final de Màster, ACO s'ha modificat i utilitzat per a la predicció d'interaccions clau en xarxes d'interacció gènica a partir de les dades d'expressió.

Algunes modificacions s'han aplicat a l'ACO, així com l'enfocament en com ha estat utilitzat. L'algoritme s'ha modificat per acceptar matrius de correlació provinents de dades d'expressió gènica com a *input* i com a *output* retorna una llista de les interaccions entre gens. A més, l'enfocament ha estat modificat, en lloc d'utilitzar l'estratègia del *Traveling Saleman Problem* i passar per cada node una vegada, s'ha modificat per tal de triar el nombre de vegades que es desitja que l'agent (*ant*) visiti cada node (gen).

El ACO modificat (ACOGeneInteraction) ha estat provat en dos conjunts de dades i els resultats mostren que l'ACO modificat ens permet predir mes interaccions que amb l'enfocament del Traveling Salesman Problem, i al mateix temps, ens permet identificar els gens centrals a les xarxes estudiades.

Index

1. Introduction 8
    1.1 Context and justification of the topic proposed 8
    1.2 Objectives 9
    1.3 Approach and methodology 10
    1.4 Work plan 11
    1.5 Brief summary of products obtained 13
    1.6 Brief summary of the chapters covered in the study 14
2. Ant Colony Optimization algorithm 16
    2.1 Introduction to Swarm Intelligence 16
    2.2 Ant Colony Optimization Algorithm 18
        2.2.1 Ant System 19
    2.3 Ant Colony Optimization modifications 21
        2.3.1 Elitist ant system 21
        2.3.2 Rank-based ant system 21
        2.3.3 Min-Max ant system 21
        2.3.4 Ant colony system 22
    2.4 Graph Optimization 24
3. Gene interaction networks 27
    3.1 Gene interaction networks 27
    3.2 Raw data: DNA microarrays 29
        3.2.1 Conversion of gene expression data to gene correlation matrix 29
4. ACO modification for predicting gene interactions 32
    4.1 Design modifications of ACO 32
    4.2 Modifications 34
    4.3 Implementation of modified ACO algorithm 35
    4.4 Test modified ACO algorithm (ACOGeneInteraction) 37
5. Study of specific gene regulatory networks 39

5.1 Study network gene interactions Dataset1 41
    5.1.1 Correlation matrix 41
    5.1.2 ACOGeneInteraction 42
    5.1.3 Table of predicted interactions 44
    5.1.4 Network predicted using Cytoscape 46
    5.1.5 Central genes in network 48
    5.1.6 Key genes 50
5.2 Study network gene interactions Dataset2 51
    5.2.1 Correlation matrix 51
    5.2.2 ACOGeneInteraction 52
    5.2.3 Table of predicted interactions 54
    5.2.4 Network predicted using Cytoscape 56
    5.2.5 Central genes in network 57
    5.2.6 Key genes 58
6. Conclusion 59
7. Glossary 61
8. Bibliography 62
9. Annexes 64
Annex 1. Table of predicted interactions Dataset1 65
Annex 2. Table of central genes in the network Dataset1 82
Annex 3. Table of predicted interactions Dataset2 85
Annex 4. Table of central genes in the network Dataset2 88

List of figures

Figure 1. Bridge experiments 18

Figure 2. DNA microarray steps 29

Figure 3. Number possible interactions as number of nodes increases 32

Figure 4. Fraction of interactions possible predicted by ACO 33

Figure 5. ACO-TSP 37

Figure 6. ACO-NewApproach (‘times’ = 2) 38

Figure 7. ACO-NewApproach (‘times’ = 3) 38

Figure 8. ACO-NewApproach (‘times’ = 4) 38

Figure 9. Number of interactions predicted in Dataset1 42

Figure 10. Number of different interactions predicted Dataset1 43

Figure11. Image network predicted by ACO using Dataset1 46

Figure 12. Number of interactions predicted in Dataset2 52

Figure 13. Number of different interactions predicted Dataset2 53

Figure 14. Image network predicted by ACO using Dataset2 56

1. INTRODUCTION

1.1. Context and justification of the topic proposed

Nowadays, technology and artificial intelligence allows us to process the large amount of information that is produced from a variety of sectors (biology, economics, etc.). In the field of artificial intelligence there are a lot of algorithms whose operations seems to imitate, or at least are inspired in some of the processes that nature uses on daily basis (bio-inspired algorithms).

The collective intelligence algorithms (Swarm Intelligence, SI) are artificial intelligence techniques based on the study of collective behavior present in nature, such as ant colonies, flocks of birds, fish bank, etc. These systems are typically composed of a population of simple agents that interact with each other and with their environment. The agents follow very simple rules, and although there is no centralized control structure dictating how individual, local and, to some extent, random agents should behave, the interactions between these agents result in the emergence of "intelligent" global behavior. They have showed good results when applied to problems that are spatially distributed and change over time. [1]

Swarm intelligence techniques have been applied to many different types of problems: routing problems, task assignment for robots in a factory, gene interaction prediction, etc.

This final project (TFM) focuses on the collective intelligence algorithm Ant Colony Optimization (ACO), which is an optimization algorithm inspired by ant colonies. In nature, ants of some species initially wander randomly until they find a food source and return to their colony laying down a pheromone trail. If other ants find this trail, they are more likely not to keep travelling at random, but instead to follow the trail, and reinforcing it if they eventually find food. Using this approach ACO tries to solve problems of finding optimal or suboptimal paths in graphs, such as the traveling salesman problem (TSP).

In computational biology, gene regulatory networks (GRNs) have attracted a lot of interest and many methods have been introduced for their statistical inference from gene expression data. GRN is a collection of molecular regulators that interact with each other and with other substances in the cell to govern the gene expression levels of mRNA and proteins. For this reason, the study of these networks provides valuable insight about how the cells work internally and how the biological processes are regulated. [2]

The network of these interactions called Gene Regulatory Network (GRN), which has a structure similar to a graph, is one of the optimization problems recently studied. GRN information can be used for grouping clusters of proteins, predicting protein complexes or protein functions, investigating topological properties of protein networks and helping to study biological processes.

This paper aims to propose and implement an improved Ant Colony Optimization algorithm that allows us to predict gene interactions in a gene interaction network.

1.2. Objectives

1. General objectives
    - ✓ Study the Ant Colony Optimization algorithm.
    - ✓ Study Gene Regulatory Networks.
    - ✓ Use Ant Colony Optimization algorithm to make predictions on gene regulatory networks.
2. Specific objectives
    - ✓ Study Ant Colony Optimization algorithm and its modifications and / or improvements.
    - ✓ Propose a change to try to improve the algorithm Ant Colony Optimization.
    - ✓ Implement Ant Colony Optimization algorithm modified.
    - ✓ Study specific gene regulatory networks.
    - ✓ Perform predictions about the interactions in the gene regulatory network using the modified algorithm.
    - ✓ Improve knowledge of the programming language Python.

1.3. Approach and methodology

This study focuses on developing a new method based on Swarm Intelligence (Ant Colony Optimization) for predicting gene interactions in gene regulatory networks.

In computer science and mathematical optimization, metaheuristic algorithm refers to a procedure intended to generate a partial search algorithm (heuristic) that may provide an acceptable solution to an optimization problem, especially when working with large amounts of data or limited computation capacity. [3]

Metaheuristics sample a fraction of solutions from the total possible solutions, which turns out to be too large to be completely sampled. Metaheuristics make few assumptions about the optimization problem being solved, they are problem-independent techniques that do not consider any specificities of a particular problem and therefore, they may be used for a variety of problems. [4.]

The algorithm ACO is a metaheuristic approach inspired by ant colonies, which is able to solve problems of finding optimal or near-optimal paths in graphs.

The biological problem to be solved or improved in this study is to predict interactions between genes in a GRN. GRN can be expressed as graphs: genes are represented as vertices and if there is interaction between two genes appears an edge that connects the two vertices (genes).

Since gene regulatory networks can be expressed as a graph, it was decided to use the ACO algorithm to predict possible interactions among the present genes. The idea of this study is to develop a new algorithm based on ACO that allows us to predict the interactions in the network.

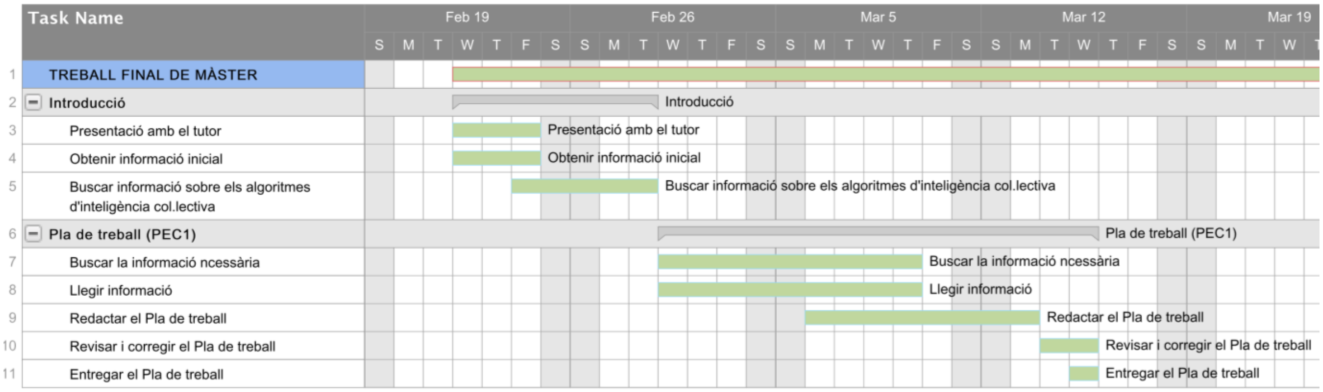
For the implementation of the new algorithm Python was used. Unlike R which is a language specific for statistics and / or data analysis, Python is a high-level programming language widely used for programming in general. Python is well-known for its productivity and code readability, it is also a very flexible language that can create a wide range of products.

Besides, for graphical representation of the predicted gene interaction networks Cytoscape software will be used.

1.4. Work Plan

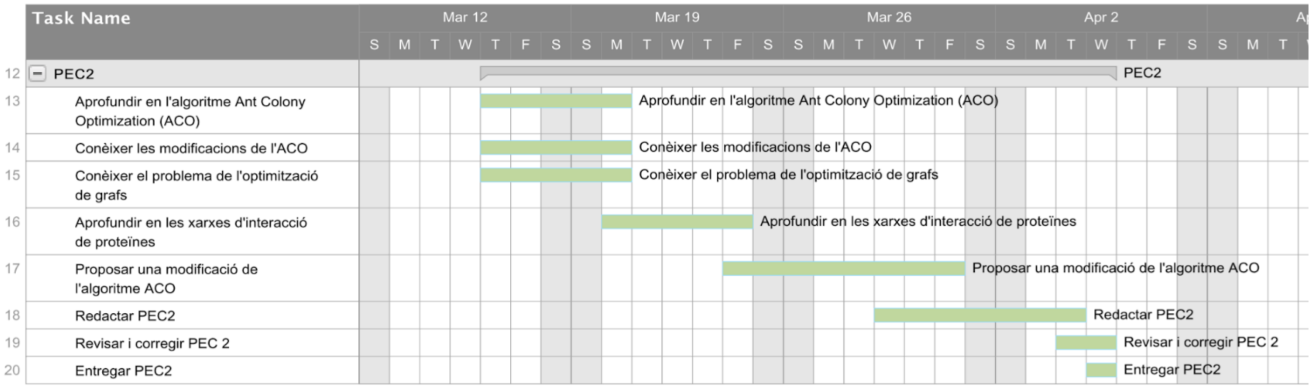
Work plan is showed in Gantt diagram format.

- Introduction and Work Plan (PEC1):



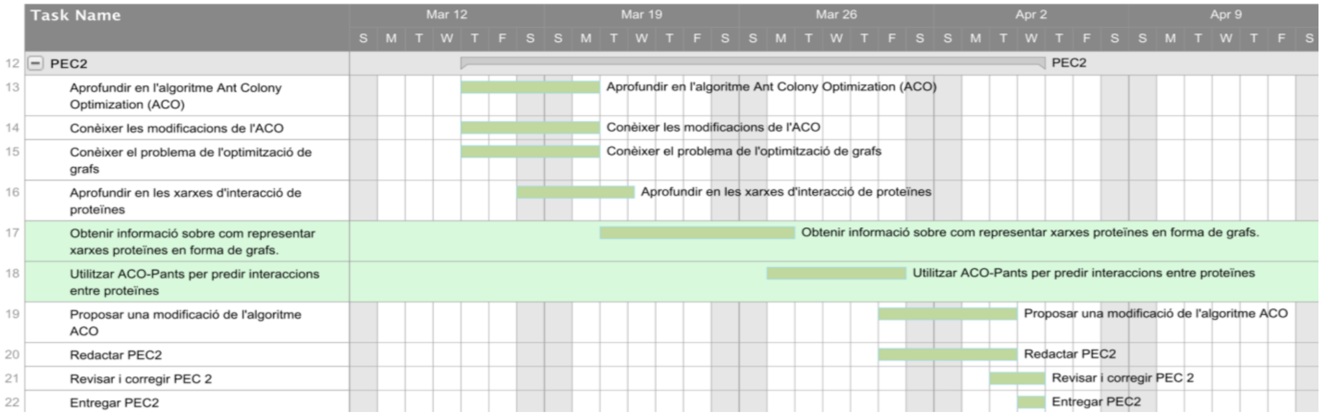
	Task Name
1	**TREBALL FINAL DE MÀSTER**
2	**Introducció**
3	Presentació amb el tutor
4	Obtenir informació inicial
5	Buscar informació sobre els algoritmes d'intel·ligència col.lectiva
6	**Pla de treball (PEC1)**
7	Buscar la informació ncessària
8	Llegir informació
9	Redactar el Pla de treball
10	Revisar i corregir el Pla de treball
11	Entregar el Pla de treball

- PEC2:



	Task Name
12	**PEC2**
13	Aprofundir en l'algoritme Ant Colony Optimization (ACO)
14	Conèixer les modificacions de l'ACO
15	Conèixer el problema de l'optimització de grafs
16	Aprofundir en les xarxes d'interacció de proteïnes
17	Proposar una modificació de l'algoritme ACO
18	Redactar PEC2
19	Revisar i corregir PEC 2
20	Entregar PEC2

Some modifications were applied (highlighted in green):

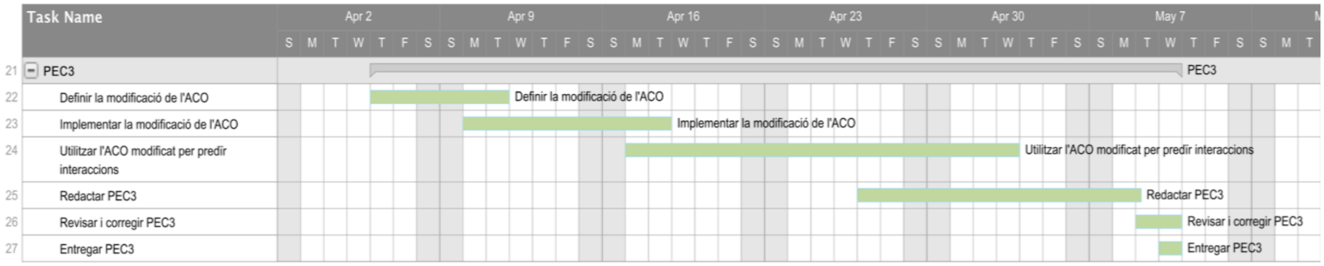


	Task Name
12	**PEC2**
13	Aprofundir en l'algoritme Ant Colony Optimization (ACO)
14	Conèixer les modificacions de l'ACO
15	Conèixer el problema de l'optimització de grafs
16	Aprofundir en les xarxes d'interacció de proteïnes
17	Obtenir informació sobre com representar xarxes proteïnes en forma de grafs.
18	Utilitzar ACO-Pants per predir interaccions entre proteïnes
19	Proposar una modificació de l'algoritme ACO
20	Redactar PEC2
21	Revisar i corregir PEC 2
22	Entregar PEC2

When we were working on getting information about interaction networks, we realized one important step was how to represent the interaction network as a graph, and it was decided to stablish a specific task for this issue.

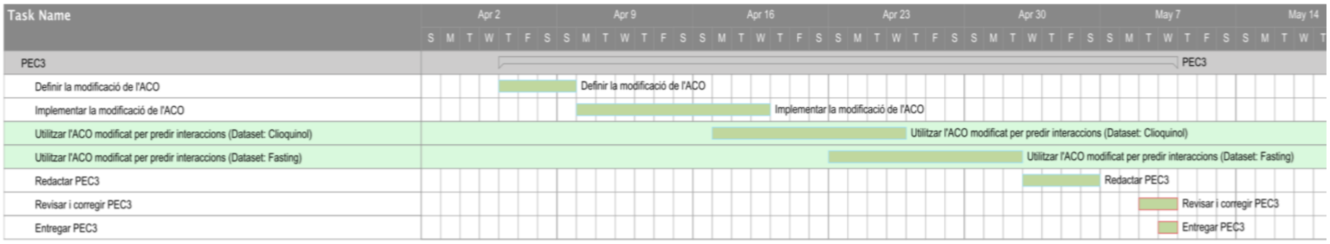
Also, we decided it was necessary to reproduced the results from previous studies which used ACO-TSP to predict gene interactions to use it as a benchmark for our study.

- **PEC3:**



	Task Name
21	PEC3
22	Definir la modificació de l'ACO
23	Implementar la modificació de l'ACO
24	Utilitzar l'ACO modificat per predir interaccions
25	Redactar PEC3
26	Revisar i corregir PEC3
27	Entregar PEC3

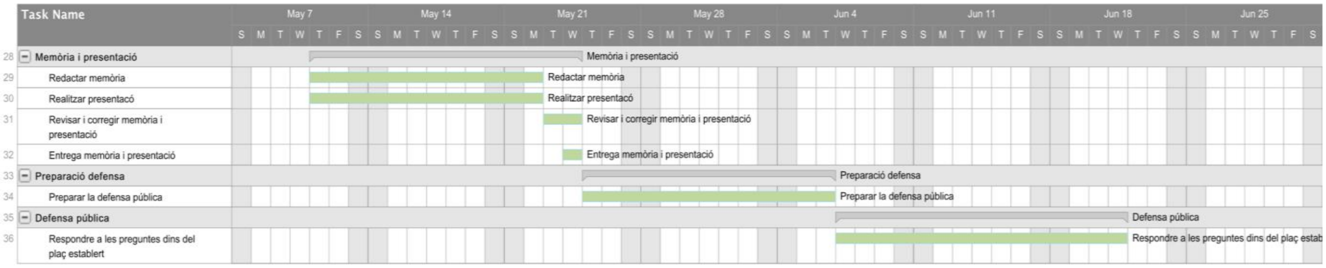
Some minor modifications were applied:



Task Name
PEC3
Definir la modificació de l'ACO
Implementar la modificació de l'ACO
Utilitzar l'ACO modificat per predir interaccions (Dataset: Clioquinol)
Utilitzar l'ACO modificat per predir interaccions (Dataset: Fasting)
Redactar PEC3
Revisar i corregir PEC3
Entregar PEC3

Two different datasets were used to predict interactions, so the task of predicting interactions was divided into two to represent this minor modification.

- **Final project and presentation, preparation and public defense of the project:**



	Task Name
28	Memòria i presentació
29	Redactar memòria
30	Realitzar presentació
31	Revisar i corregir memòria i presentació
32	Entrega memòria i presentació
33	Preparació defensa
34	Preparar la defensa pública
35	Defensa pública
36	Respondre a les preguntes dins del plaç establert

1.5. Brief summary of products obtained

Official documents for the Universitat Oberta de Catalunya:

- Continuous assessment:

1. Work plan (PEC1).

The Work Plan Report establishes all the necessary information to carry out the Final Project: context, objectives, work plan, methodology and structure of the work to be performed. Also, it organizes all the tasks (set the deadlines and the content for PEC2 and PEC3) and establishes the work schedule considering the potential risks and estimating the time needed for each task.

2. Follow-up report Phase I (PEC2).

PEC2 covers all the following topics:

1. Study the collective intelligence algorithm Ant Colony Optimization.
2. Search about optimization graphs.
3. Investigate about protein interaction networks.
4. Propose a change to try to improve ACO algorithm.

3. Follow-up report Phase II (PEC3).

PEC3 covers the following points:

1. Define and specify the modification of the ACO algorithm.
2. Implement the proposed modifications to ACO algorithm using the programming language Python.
3. Use the modified algorithm to predict the interactions in the protein network.

4. Final Report.

The final report covers all the topics studied through the TFM subject, from the introduction (Work plan) to the results of the study (Conclusions).

5. Presentation.

Presentation of the study performed, along with the results found.

- Results of the study:

1. Online repository of scripts in Python, in GitHub.
2. Research on biological gene interaction on specific networks. For each dataset, it has been attached:
    - Table of predicted interactions.
    - Graph of the interactions predicted.
    - Table of central key genes in the network.

1.6. Brief summary of the chapters

The final Project is composed by the five following chapters:

The **first chapter** is a general introduction to understand the topic and the problem to be solved in this study. It covers the context of the topic and introduces the biological problem, also the approach and methodology selected to carry out the study.

The **second chapter** introduces the different algorithms based on Swarm intelligent, furthermore provides a deep overview about Ant Colony Optimization since it is the algorithm that is going the be used later on in this study. To understand ACO, first it is explained how ants in a colony work and communicate in nature and how ACO imitates this process. After that, it is described the first algorithm based on ants (Ant System) and the following modifications and improvements that have appeared through time. To finish this chapter optimization of graphs is covered along with the Traveling Salesman Problem approach.

The **third chapter** describes gene interaction networks and provides a review of some previous studies related to this subject. Also, there is a description of how row data is produced (DNAmicroarrays) and how correlation matrices are calculated from these data.

The **forth chapter** is about the modifications applied to ACO. In this part, it is explained how the algorithm has been modified and which problems is these modifications trying to solve. Also, it has been added the pseudocode and an example of how the algorithm works.

The **fifth chapter** covers the study of two specific datasets (‘Clioquinol effect on yeast’ and ‘Short-term fasting on skeletal muscle’) using ACOGeneInteraction and showing the results obtained.

And finally, there is a **conclusion** section with the more relevant results derived from the findings of the study.

2. ANT COLONY OPTIMIZATION ALGORITHM

2.1. Introduction: Swarm Intelligence

In the field of artificial intelligence there is a large number of algorithms, whose operations are attempted to resemble, or at least are inspired, in some of the processes that nature carries out day by day.

These bioinspired algorithms are models, designs and techniques that rely on biological patterns that help the machine understand and emulate the intelligence of species (such as ant colonies or fish banks), or biological processes (such as evolution or neural networks).

These bioinspired algorithms can be divided into large groups:

- Swarm Intelligence Algorithms

Set of techniques based on emerging social and cooperative behavior of organisms grouped in colonies, swarms, etc.

- Genetic algorithms

Set of heuristics that base their behavior in modeling evolutionary processes based on the survival of the most apt individuals of the population.

- Others: Artificial neural networks, Decision trees, etc.

Set of algorithms based on different biological processes that are able to predict or classify.

Swarm Intelligence Algorithms

The algorithms of Swarm Intelligence (SI) are techniques of artificial intelligence based on the study of collective behaviors present in nature, such as ant colonies, flocks of birds, fish banks, etc.

Bonabeau defined swarm intelligence as "emerging collective intelligence from a group of simple agents." [5]

The two main characteristics of these systems are:

- the division of labor and,
- self-organization.

Agents follow very simple rules **without centralized control**, leading in many cases to the **division of tasks or specialization** to coordinate their actions. Different agents perform simple and workable tasks. This division allows the whole swarm to perform more complex tasks that requires all individuals working together.

Self-organization requires **interactions between individuals**. These types of interactions start from a communication between individuals that can be given directly, when communication is made from individual to individual, or indirectly, when communication arises from the changes that each individual performs in the environment.

Moreover, self-organization relies on four basic properties:

- positive feedback,
- negative feedback,
- fluctuations and
- multiple interactions.

Positive feedback is useful for amplification while negative feedback is useful for stabilization. Fluctuations are useful for randomness, and multiple interactions take place when the agents share information among themselves within their searching area. [5]

Such intelligence is decentralized, self-organizing and distributed throughout an environment. The information is typically stored throughout the participating homogeneous agents, or is stored or communicated in the environment itself such as using pheromones in ants, dancing in bees, and proximity in fish and birds. [6]

Such social behavior defines the movements of decision variables in the search space and guides them towards optimal solutions.

2.2. ACO Algorithm

Experiments with the species Iridomyrmex humilis, Linepithema humile and Lasius niger showed that there is an indirect communication between individuals through pheromones. The ants, on their way from the nest to the food source and vice versa, deposit pheromones in the soil forming a trail, which are able to smell the rest of the components of the colony. The greater the concentration on a path the greater the probability that an ant will follow that path.

To test this fact, double-bridge experiments were designed, changing the relationship $(r = ll/ls)$ between the lengths of branches, being ll and ls are the lengths of the long and short branches, respectively. [5]

In the first experiment the nest was connected with the food source using two branches of same length, while in the second experiment the two branches were of different length.

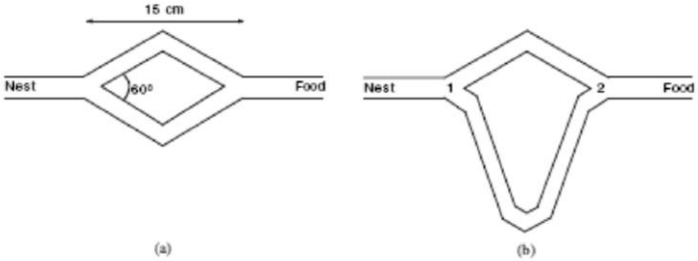


Figure 1.Bridge experiments: a) Branches of same length, r=1. b) Branches of different length, r= 2. [7]

In the first case the results showed that ants move in one way or another indifferently, but eventually more pheromone accumulates in the most traveled path, reason why the ants choose it with greater probability. After several repetitions of the experiment was observed that the branch was chosen indistinctly.

In the second experiment, they used two branches of different length and after letting the colony move freely it was observed that eventually the ants chose a single branch. But in this case, in most trials the ants chose the shortest path. This phenomenon is explained by the fact that the shorter path allows the ants to travel more times between the nest and the source and, therefore, more pheromone is deposited. [5]

2.2.1. Ant System (AS)

ACO is a metaheuristic approach inspired by Ant System (AS) initially proposed by Marco Dorigo in 1992 in his doctoral thesis.

This algorithm consists of four main components:

- **Ant:**

Ants are the agents of the algorithm that explore and exploit the searching area.

- **Pheromone:**

The agents (ants) drop pheromone in the search space and the quantities of these pheromones indicate the probability of that path to be chosen by other agents. The intensity of the pheromone on the paths may be considered as a memory of the system.

- **Daemon action:**

It is used to gather global information of the system that cannot be done by a single ant. If the convergence is slowly the daemon action adds an extra pheromone to accelerate the process.

- **Decentralized control:**

The decentralized control operated by the algorithm makes it robust and flexible.

To find the best solution the agents choose the next node using the equation known as the *transition rule*, that represents the probability for ant k to go from city i to city j on the tth tour, also known as *random proportional transition rule*:

$$p_{ij}^{k}(t) = \frac{[\tau_{ij}(t)]^{\alpha} \cdot [\eta_{ij}]^{\beta}}{\sum_{l \epsilon J_i^k} [\tau_{il}(t)]^{\alpha} \cdot [\eta_{il}]^{\beta}}$$

In the equation, τ_{ij} represents the pheromone trail and η_{ij} the visibility between the two cities, while α and β are adjustable parameters that control the relative weight of trail intensity and visibility. [5]

Pheromones are a very important component in the algorithm, since it contributes to the memory of the system. Agents leave a trail of pheromones that make paths to be chosen with higher probability by the next agents.

Another very important component is the *pheromone evaporation rate* or decay factor. Too high evaporation rate makes the agents to explore more, since there will be less pheromone on the searching area. While low evaporation rate results in exploitation behavior of the agents.

Pheromone evaporation has the advantage of avoiding convergence to a locally optimal solution. If there were no evaporation the paths chosen by the first ants would tend to be excessively attractive to the following ants, and hence the exploration of space and solutions would be limited. The influence of pheromone evaporation on real ant systems is unclear, but it is very important in artificial systems. [5]

Pheromone decay factor is implemented by using a decay coefficient ρ ($0 \leq \rho < 1$), and the pheromone update rule that is applied to all edges is as follows:

$$\tau_{ij}(t) \longleftarrow (1-\rho) \cdot \tau_{ij}(t) + \Delta\tau_{ij}(t)$$

where $\Delta\tau_{ij}(t) = \sum_{k=1}^{m} \Delta\,\tau_{ij}^{k}(t)$, being m = number of ants.

Another important parameter is the *number of ants* (m), it keeps constant over time but it has to be chosen carefully since too many ants lead to suboptimal path, whereas too few ants are not able to cooperate between them due to pheromone decay.

This algorithm has been used to solve TSP, the quadratic assignment problem, routing and image processing, among others.

2.3. Ant Colony Optimization modifications

Some modifications have been created in order to improve AS performance:

2.3.1. Elitist ant system (EAS)

This variant was proposed by Dorigo and is based on an elitist strategy, in which the best solution found so far deposits an additional quantity of pheromone in each iteration. This is controlled by a parameter (e) which defines the weight of each given solution.

The computational results showed that the elitist strategy allows to find better solutions with a smaller number of iterations than the AS. [5]

2.3.2. Rank-based ant system

AS rank-based was proposed by Bullnheimer, Hartl and Straus. The proposal is to choose the ants with the best solutions of each iteration to make the deposit of pheromones. Prior to the update of pheromone, the r best ants so far are sorted based on performance in a list, and each of them deposits an amount of pheromone depending on the range assigned (w - r), being the best solution so far that adds more (w), even if it does not belong to the current iteration. [5]

2.3.3. Max-Min Ant System

The approach was introduced by Stützle and Hoos in 2000 and limits pheromone values within a range. [5]

Several aspects regarding the original algorithm are modified:

- At the beginning, pheromone traces are set at the maximum value to encourage exploration of the environment.
- Pheromone values are limited to one interval to avoid stagnation.

- Only one ant is allowed to add pheromone, to favor the exploitation of the best solutions found during the execution of the algorithm.

2.3.4. Ant Colony System

This algorithm was introduced by Dorigo and Gambardella to improve the performance of the Ant System algorithm.

Uses an approach based on a centralized (global) update for the pheromone update, so that it concentrates the search close to the best solution found so far in order to improve the convergence time. [5]

ACS is based on four modifications of the AS:

1. Different transition rule

The transition rule is modified to allow explicitly for exploration. It allows to cut the exploration by tuning a parameter, what make the algorithm to concentrate on the best solutions find so far instead of exploring constantly.

This is done following the rule:

$$j = \begin{cases} \arg max_{u \in J_i^k} \left\{ [\tau_{iJ}(t)] \cdot [\eta_{iJ}]^{\beta} \right\} & if\ q \leq q_0 \\ J & if\ q > q_0 \end{cases}$$

where q is a random variable uniformly distributed over [0,1], and q_0 is an adjustable parameter (0 ≤ q_0 ≤ 1), and $J \in J_i^k$ is a city randomly chosen according to probability:

$$p_{ij}^k(t) = \frac{[\tau_{ij}(t)] \cdot [\eta_{ij}]^{\beta}}{\sum_{l \epsilon J_i^k} [\tau_{ij}(t)] \cdot [\eta_{il}]^{\beta}}$$

2. Different pheromone trail update rule

Only the agent (ant) that generates the best tour is allowed to globally update the concentration of the pheromone. The update of pheromone is applied only on the edges of the best tour found so far. In this way, the agents are encouraged to explore paths in the vicinity of the best tour found until that moment.

$$\tau_{ij}(t) \longleftarrow (1-\rho)\cdot \tau_{ij}(t) + \rho\cdot \Delta\tau_{ij}(t)$$

Where (i,j) represents the edges of the best tour in the trial find so far (T^{+}), and $\Delta\tau_{ij}(t) = 1/L^{+}$, where L^{+} is the length T^{+}.

3. Use of local updates of pheromone trail to favor exploration

In order to avoid stagnation local updates are performed. The local updated rule makes the level of pheromone diminish on the edges visited. So, the edges visited are less attractive allowing the agents to explore more.

The local update: when the ant k is in the city i and goes to city j the local pheromone concentration is updated following the formula:

$$\tau_{ij}(t) \longleftarrow (1-\rho)\cdot \tau_{ij}(t) + \rho\cdot \tau_0$$

4. Use of candidate list to restrict the choice of the next city to visit

Cities in the candidate list are ordered by increasing distance, and the list is scanned sequentially. The agents first explore the cities on the list, and consider the other cities only if the ones in the list have already been visited.

2.4. Graph Optimization

Travelling Salesman Problem is one of the most famous problems in computer science for studying optimization, the objective is to find a complete route that connects all the nodes of a network, visiting them only once and returning to the starting point while minimizing the total distance of the route.

The problem of the traveling agent has an important variation, and this depends on whether the distances between one node and another are symmetric or not, that is, that the distance between A and B is equal to the distance between B and A, since in practice is very unlikely to be so. The number of possible routes in a network is determined by the equation:

$$(\boldsymbol{n}-\mathbf{1})!$$

This means that in a network of 6 nodes the number of probable routes is equal to (6-1)! = 120, and as the number of nodes increases, the number of possible routes grows factorially.

In the case that the problem is symmetrical the number of possible routes is reduced to half:

$$(\,(\boldsymbol{n}-\mathbf{1})!\,)\ /\ \mathbf{2}$$

The complexity of calculating the traveler's problem has sparked multiple initiatives to improve efficiency in route calculation. The most basic method is the one known by the name of ‘brute force’, which consists of the calculation of all possible routes, which becomes extremely inefficient and almost impossible in large networks. [8]

There are also heuristics that have been developed by the complexity in the calculation of optimal solutions in robust networks, that is why there are methods such as the nearest neighbor and the cheapest insertion.

Finally, we find algorithms that provide near optimal solutions such as ACO algorithm. The basic idea underlying all the ant-based algorithm is to use a positive feedback mechanism based on the laying pheromone. The pheromone component allows the best solutions found to be kept in memory, which can be used to get better solutions. To avoid stagnation of the algorithm a form of negative feedback is implemented through pheromone evaporation, but it must not evaporate too fast in order to make cooperation behavior emerge.

In TSP, the goal is to find the tour with minimal length connecting n given cities, and each city must be visited only once.

The distance between cities can be defined by Euclidean distance or other distance functions. In the graph, the cities would be the nodes and the connections between the cities are the edges of the graph.

The graph does not need to be fully connected, all the nodes may not be connected to all the other nodes. And the distances may not be symmetric, distance ij may be different that distance ji.

In order to solve TSP using ACO the transitions of the ants from city to city depends on the following premises:

- Whether or not the city has already been visited. Each ant has a memory or tabu list to make sure each city is visited once per tour.
- The inverse of the distance between two nodes (visibility). Visibility is based on local information and represents the heuristic desirability of choosing city j when in city i.
- The amount of virtual pheromone on the edges. It is a global type of information, represents the learned desirability of choosing city j when in city i.

Other things to take into account are:

- The transition rule: the probability for an ant to go from city i to city j while building a tour.
- Pheromone decay: Without pheromone decay the algorithm would lead to amplification of the initial random fluctuation, that will produce not optimal solutions.
- Total number of ants: It is an important parameter since too many ants will reinforce not optimal solutions, while too few ants would not produce cooperative effect due to pheromone decay. It is suggested to use a number of ants equal to the number of cities in the graph. [5]

3. GENE INTERACTION NETWORKS

3.1. Gene interaction networks

As defined by Gurkan Bebek in his article “Identifying gene interaction networks”: ‘A gene interaction network is a set of genes (nodes) connected by edges representing functional relationships among these genes. These edges are named interactions, since the two given genes are thought to have either a physical interaction through their gene products, e.g., proteins, or one of the genes alters or affects the activity of other gene of interest’. [9]

In previous studies has been used ACO with the TSP approach for predicting interactions in gene interaction networks.

- *Ant Colony Optimization for Inferring Key Gene Interaction* Khalid Raza and Mahish Kholi [10]

In this study is was applied ACO for inferring the key gene interactions from gene expression data.

Genes are not independent, and their expression is affected by other genes that can modify with their expression. In previous articles researchers have used different type of algorithms and approaches for predicting gene interaction network such as Boolean networks, differential equations, machine learning, etc., these techniques give acceptable results but since they are complex algorithms, when the number of genes increases the complexity of the algorithm also increases.

On the other hand, nature based algorithm such as ACO are simpler in nature and have been applied in some biological problems like drug design and 2D protein folding. Since, ACO is an optimization algorithm the authors used it in the study for the optimization of the gene interaction matrix. The researchers considered gene correlation matrix as a network where genes are the nodes and the pair-wise correlation values are assigned to the edges between nodes (genes). And the algorithm used had to find the maximum correlation Hamiltonian circuit (graph that visits each vertex exactly once) in the gene interaction network.

The results were that out of 9 interactions, 3 interactions were predicted correctly. But as stated in the article, ACO-TSP has one limitation: **due to his nature, it can find out a total number of interactions equal to the number of genes.**

ACO based algorithm does not produce very accurate result, but they are robust because same result was produced after each iteration.

To fix these problems in the future the authors suggest trying to merge ACO based algorithms with some other algorithm.

- *Ant colony Optimization for Optimal path finding in protein interaction network* by Jamaludin Sallim, Rosni Abdullah y Ahamad Tajudir Khader. [11]

In this study ACO algorithm was used along with the TSP approach to cluster proteins in protein interaction network.

Understanding the protein interaction is one important problem in biology, and understanding the networks of protein interactions will help to understand the internal functions of the cells.

In the study, the distance in the distance matrix actually represents how similar the set of the different proteins interaction partners are. The optimal path is based on the shortest path constructed by ACO.

The experimental results showed that ACO has the efficiency to find the optimal path for the chosen protein interaction network (PIN).

The important thing that needs to be considered in this work is the transforming process to obtain the distances for the PIN based on similarity.

3.2. Raw data: DNA microarrays

DNA microarrays, also called DNA chips, are a technique used to compare differential gene expression under two different conditions, such as comparing gene expression in normal cells with cancer cells.

It is based on the ability of hybridization between 2 single-stranded DNA molecules complementary to each other. It is formed by a solid surface, which can be of glass, silica or plastic, that contains spots of about 200µm in diameter, where thousands of copies of a nucleotide sequence are deposited. Each spot contains different sequences which are perfectly attached to the surface.

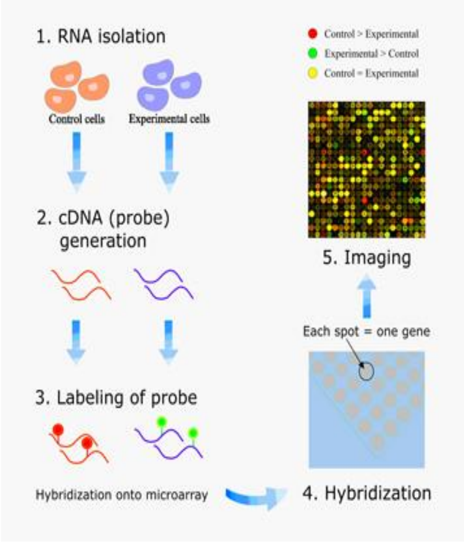


Figure 2. DNA microarray steps. [12]

As said earlier, the principle of this technique is based on the ability of hybridization between two strands of DNA complementary to each other by means of hydrogen bonds. The less complementary the sequences to each other, the lower affinity will have between the two single-stranded DNA molecules. And hence, after washing the chip only the most complementary hybridizations will remain attached to the surface while the partially complementary sequences will be removed.

The signal strength of a given probe depends on the amount of complementary sequences present in the sample that have been hybridized with the probe. In this way, the amount of a given gene expressed in a cell or in a tissue as a function of the signal intensity can be quantified in a relative way. As the specific situation of each probe in a microarray is known, the expression of each gene can be compared under different conditions.

3.2.1. Conversion of gene expression to gene correlation matrix

In 2000 on the article '*Gene expression data analysis*', the authors Brazma A, Vilo J. stated that: *'Microarrays are one of the latest breakthroughs in experimental molecular biology, which allow* *monitoring of gene expression for tens of thousands of genes in parallel and are already producing huge amounts of valuable data. Analysis and handling of such data is becoming one of the major bottlenecks in the utilization of the technology. The raw microarray data are images, which have to be transformed into gene expression matrices-tables where rows represent genes, columns represent various samples such as tissues or experimental conditions, and numbers in each cell characterize the expression level of the particular gene in the particular sample. These matrices have to be analyzed further, if any knowledge about the underlying biological processes is to be extracted.'* [13]

Therefore, in order to further study the data produced by the arrays, some modifications had to be done. The expression values of the gens from the raw data were converted to a gene correlation matrix. The correlation matrix is a table showing correlation coefficients between gens. Each gene in the table is correlated with each of the other genes present in the table, which shows pair-wise correlation among all of them.

The expression values were transformed using log2 and a normal standardization of the data and then converted to correlation matrix using Pearson correlation coefficient.

Pearson correlation coefficient was applied to find out regulatory relationship between pairs of genes. The correlation is +1 if there is a perfect positive linear relationship, −1 if there is a perfect negative linear relationship and values between −1 and 1 indicates the degree of linear dependence between the variables. Closer the coefficient to either −1 or +1, stronger the correlation between the variables. When the correlation coefficient is close to zero there is no evidence of any relationship.

If we have n samples (conditions) of x and y genes, written as x_i and y_i where i = 1, 2, ..., n, the correlation coefficient between x and y (r_{xy}) can be estimated as:

$$r_{xy} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{(n\sum x_i^2 - (\sum x_i)^2)}\,(n\sum y_i^2 - (\sum y_i)^2)}$$

Confidence in a relationship is formally determined not just by the correlation coefficient but also by the number of pairs in the data. If there are very few pairs then the coefficient needs to be very close to 1 or –1 for it to be deemed 'statistically significant', but if there are many pairs then a coefficient closer to 0 can still be considered 'highly significant'.

The standard statistically method to measure the 'significance' of the analyses is the p-value. The p-value is a number between 0 and 1 representing the probability that this data would have arisen if the null hypothesis were true.

When we make contrast of hypothesis we establish a null hypothesis (H0) that says that the observed difference between the comparison groups is due to chance. Then, we calculate the probability that the difference is due to chance and, if it is less than a given value (usually 0.05 or 0.01), we reject H0 and state that it is highly unlikely that the difference is due to chance. But highly improbable does not mean to be certain, there is always a 5% or 1% chance that, being true H0. This is what is called making an error type I error (false positive).

Applying p-values to microarrays is complicated by the large number of multiple comparisons (genes) involved. For example, a p-value of 0.05 is typically thought to indicate significance, since it estimates a 5% probability of observing the data by chance. But with 10,000 genes on a microarray, 500 genes would be identified as significant at $p < 0.05$ even if there were no difference between the experimental groups. One obvious solution is to consider significant only those genes meeting a much more stringent p-value criteria. For that reason, in his study has been used a Bonferroni correction.

The Bonferroni method is the best known and most widely used, allowing simple and without additional assumptions to be tested against each of the individual (null) hypotheses involved in the study.

Bonferroni uses α / k as the cut-off point, where alpha is the level of significance and k the number of contrasts. However, this method is very conservative and increases the probability of Type II error (false negative).

4. ACO MODIFICATION FOR PREDICTING GENE INTERACTIONS

4.1. Design modifications of ACO

As said earlier, ACO has been used to predict gene interactions using the approach of the Traveling Salesman Problem (TSP).

Using the TSP approach ACO is able to predict gene interactions with one limitation: it only can find out a total number of interaction equal to the total number of genes. In a small network could be acceptable to use the TSP approach to predict interactions but as the networks gets bigger using TSP approach predicts just a small fraction of the possible interactions.

The number of possible interactions increases exponentially as the number of nodes (genes) increases:

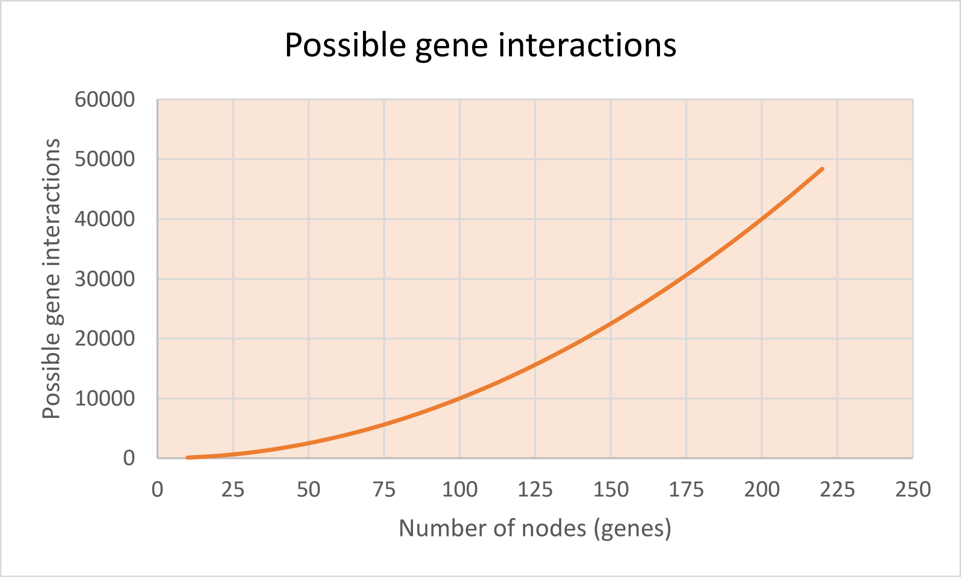


Figure 3. Number possible interactions as number of nodes increases.

The number of possible interactions increases exponentially as the number of nodes increases, but the ACO - TSP approach only predicts as many interactions as nodes there are present in the graph, which means that the **fraction of interactions predicted by the algorithm decreases as the number of nodes increases:**

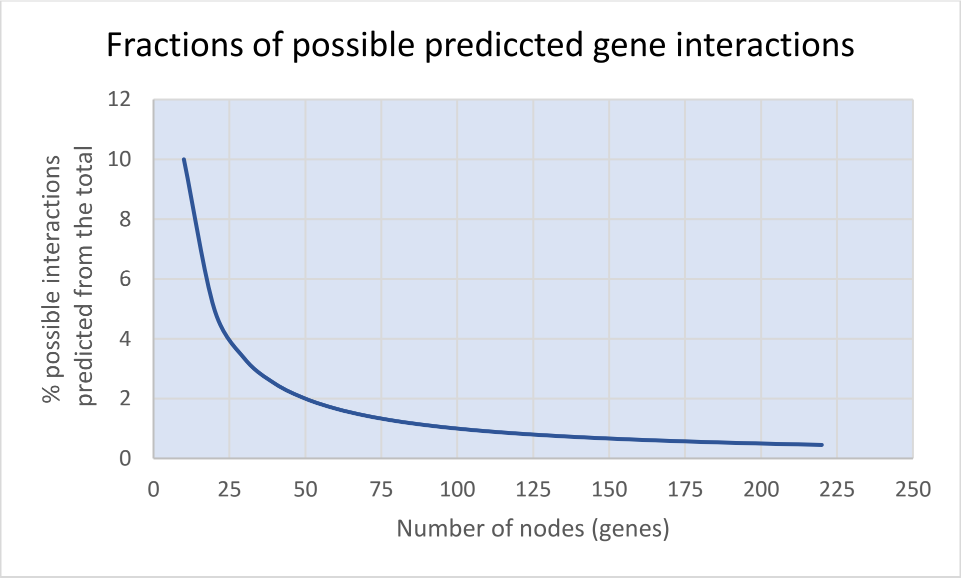


Figure 4. Fraction of interactions possible predicted by ACO

For instance, if the network has 10 nodes, there are 100 (10 x 10) possible interactions, the algorithm using TSP approach would be able to predict 10 interactions which represents 10% of the total possible interactions.

But, if we are working with a network that has 200 nodes, there are 40.000 (200 x 200) possible interactions and the algorithm would be able to predict as maximum 200 interactions, which in this case represents only 0.5% of the total possible interactions.

In order to use ACO to predict gene interactions in bigger networks, the algorithm has been modified to try to overcome this limitation. Instead of using the approach of going through the nodes only once (TSP approach), the algorithm will be forced to go through each node different times. This will lead ACO to go through unvisited routes increasing the number of possible interactions.

4.2. Modifications

In this study ACOGeneInteractions was implemented based on the previous ACOPants already implemented in Python.

In order to use ACOPants for predicting gene interactions some modifications were done:

1. The input for ACOPants is a list of coordinates (x, y) of the nodes, and providing a length function to the algorithm this is able to calculate the distances from node i to j.

In ACOGeneInteraction the **input is the correlation matrix from raw data** (DNAmicroarrays), where the column's names represent the names of the genes (nodes), and the distances between two nodes are given by the pair-wise correlation value in the matrix. In this case, nodes have no coordinates so we have to modify the algorithm to accept this data format.

2. ACOPants is implemented using the TSP approach and each agent (ant) visits each node (gene) once.

In the modified version, **instead of going once through each node (gene) the algorithm will be forced to visit each node a specific number of times which is set by a parameter called 'times'**. The purpose of this modification is to overcome the limitation of the algorithm found in other studies: ACO is able to find as many interactions as number of nodes are present in the graph. The idea is to force the algorithm to visit alternatives routes in order to find more gene interactions.

3. The outputs when using ACOPants is the list of nodes visited sequentially during the tour and the distance of the complete optimal tour.

ACOGeneInteraction has been modified so also the **output shows a list of interactions**, pairs of genes that interact between them.

4.3. Implement modified ACO algorithm

Steps in algorithm ACOGeneInteraction

1. Create WorldGI:

- ✓ Select correlation matrix and how many ‘times’ nodes will be visited
- ✓ Create nodes and edges
- ✓ Set pheromone trail to 0.01

2. Solve WorldGI:

- ✓ Adjust all the parameters needed: number of ants, independent trials, number of tour in each trial, etc.
- ✓ Place all ants at different genes randomly (number of ants selected will be placed randomly, if no number of ants has been selected as may ants as nodes will be place in a round robin way)
- ✓ Nodes unvisited = {nodes · times}
- ✓ Tabu list initially = {*i*} (Present gene)
- ✓ Each ant moves to next node by a probability function =

$$p_{ij}^{k}(t) = \begin{cases} \dfrac{[\tau_{ij}(t)]^{\alpha} \cdot [\eta_{ij}]^{\beta}}{\sum_{l \epsilon J_i^k} [\tau_{il}(t)]^{\alpha} \cdot [\eta_{il}]^{\beta}} & if\ j\ \in allowed_k \\ 0 & otherwise \end{cases}$$

Where,

$p_{ij}^{k}(t)$ = Probability of moving from gene i to j at time t.

k ϵ {1 to m} where m = number of ants

η_{ij} = Correlation between the gene i and j or visibility

α = Relative importance place on trail intensity

β = Relative importance place on distance

- ✓ After N iteration all ants complete tour, the best tour is chosen and pheromone trail is updated.
- ✓ If termination condition is met, tour ends, otherwise nodes unvisited list is emptied and cycle continues from this step.

Product in Github

The scripts created during the development of the Final project have been publicly deposited in an open-source online repository called GitHub. This collaborative development platform allows hosting projects using a version control system.

This repository was created with the purpose of an effective supervision of the scripts that were elaborated for the processing of data. This repository con be access in the following link: https://github.com/emoncho/MasterFinalProject

It contains:

- ACOGeneInteraction.py
- ExampleACOGI.py
- ClioquinolDataset.py
- FastingDataset.py
- README.md
- .cvs files obtained during the study

4.4. Test modified ACO algorithm

Example: network with 8 nodes from the article *'Ant Colony Optimization for Inferring Key Gene Interactions', authors: k Raza and M Kholi.*

Correlation matrix from the article:

	uvrD	lexA	umuDC	recA	uvrA	uvrY	ruvA	polB
uvrD	1	0.7647	0.1982	0.8013	0.8018	0.3838	0.1912	0.4326
lexA	0.7647	1	0.5101	0.9538	0.9603	0.2135	0.6497	0.6267
umuDC	0.1982	0.5101	1	0.5962	0.4584	0.2996	0.4551	0.427
recA	0.8013	0.9538	0.5962	1	0.9779	0.4535	0.5668	0.6465
uvrA	0.8018	0.9603	0.4584	0.9779	1	0.4009	0.5796	0.5855
uvrY	0.3838	0.2135	0.2996	0.4535	0.4009	1	-0.0175	0.3807
ruvA	0.1912	0.6497	0.4551	0.5668	0.5796	-0.0175	1	0.5159
polB	0.4326	0.6267	0.427	0.6465	0.5855	0.3807	0.5159	1

Predicted interactions using TSP approach (each node visited once):

```
8 interactions:

{('lexA', 'recA'), ('polB', 'uvrY'),

('recA', 'umuDC'), ('ruvA', 'polB'),

('umuDC', 'ruvA'), ('uvrA', 'lexA'),

('uvrD', 'uvrA'), ('uvrY', 'uvrD')}
```

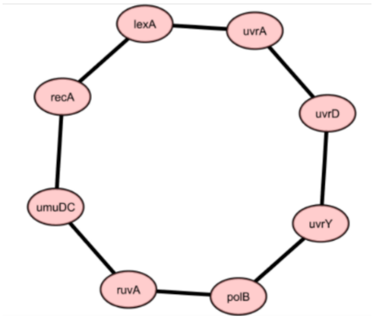


Figure 5. ACO-TSP

Predicted interactions using new approach (each node visited times = 2):

```
10 interactions: {('lexA', 'uvrA'), ('lexA', 'uvrD'), ('polB', 'ruvA'), ('polB', 'uvrD'),
('recA', 'lexA'), ('ruvA', 'umuDC'), ('umuDC', 'uvrY'), ('uvrA', 'recA'), ('uvrD', 'uvrA'),
('uvrY', 'polB')}
```

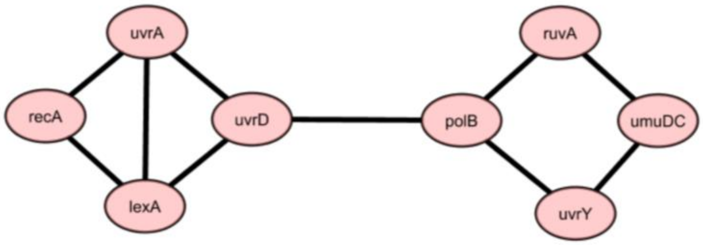


Figure 6. ACO-NewApproach ('times' = 2)

Predicted interactions using new approach (each node visited times = 3):

```
12 interactions: {('lexA', 'uvrD'), ('polB', 'ruvA'), ('polB', 'umuDC'), ('polB', 'uvrY'),
('recA', 'lexA'), ('recA', 'umuDC'), ('recA', 'uvrD'), ('ruvA', 'umuDC'),('uvrA', 'lexA'),
('uvrA', 'recA'), ('uvrD', 'uvrA'), ('uvrD', 'uvrY')}
```

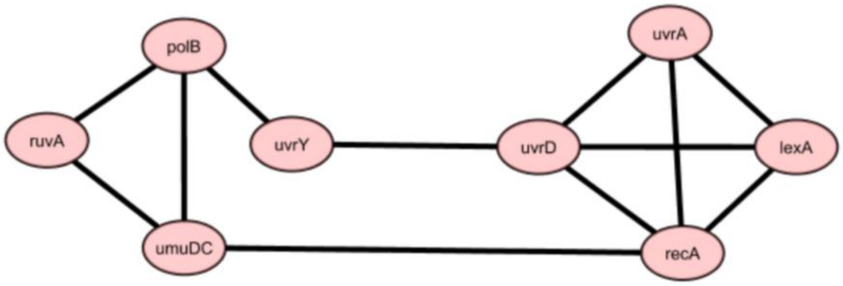


Figure 7. ACO-NewApproach ('times' = 3)

Predicted interactions using new approach (each node visited times = 4):

```
13 interactions: [('lexA', 'uvrA'), ('polB', 'ruvA'), ('polB', 'uvrD'), ('umuDC', 'uvrD'),
('lexA', 'recA'), ('ruvA', 'uvrD'), ('recA', 'uvrA'), ('uvrD', 'uvrY'), ('ruvA', 'umuDC'),
('lexA', 'ruvA'), ('lexA', 'uvrD'), ('polB', 'umuDC'), ('uvrA', 'uvrD')]
```

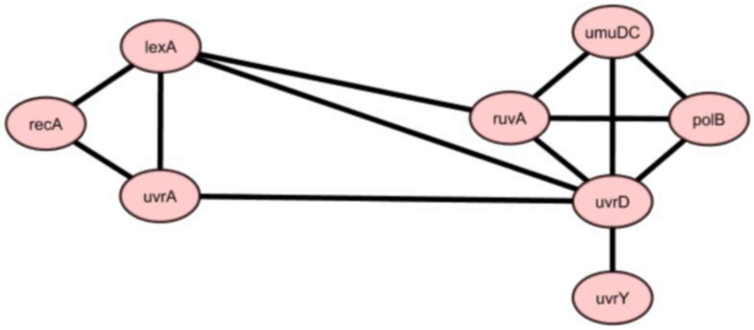


Figure 8. ACO-NewApproach ('times' = 4)

Changing the time that each node has to be visited the algorithm shows us more interactions, increasing the fraction of predicted interactions through the network.

5. STUDY OF SPECIFIC GENE REGULATORY NETWORKS

All the expression data has been collected from the National Center for Biotechnology Information. NCBI is a U.S. government-funded national resource for molecular biology information, which provides access to many public databases.

(https://www.ncbi.nlm.nih.gov/sites/GDSbrowser).

- **Dataset1: Clioquinol effect on yeast.**

Analysis of wildtype BY4742 yeast cultures treated with clioquinol (CQ), a hydrophobic chelator of copper, zinc, and iron in vitro. CQ significantly decreased amyloid-beta levels in a mouse model of Alzheimer's disease (AD). Results provide insight into molecular mechanisms underlying CQ action.

This dataset was created during the study carried by Li C., Wang J., Zhou B. in 2010 that lead to the article: *'The metal chelating and chaperoning effects of clioquinol: insights from yeast studies'* published at J Alzheimers Dis., 21(4):1249-62.

In the past Clioquinol (CQ) was a popular antibiotic, used to treat fungal and protozoal infections but was found to be neurotoxic when used at high doses. However, CQ had been investigated as a treatment for Alzheimer's disease since it is a metal protein attenuating compound, and according to previous studies, increasing bioactive metal levels in the aging brain accelerates formation of amyloid plaques, as well as neurotoxic oxidative processes. The rationale of evaluating clioquinol was that it would prevent accumulation of the amyloid-β peptide while restoring copper and zinc ion homoeostasis in cells.

Since, CQ had shown encouraging effects for the treatment of Alzheimer's disease (AD), but the mechanism was unclear, the authors of this study decided to use yeast Saccharomyces cerevisiae as a model to investigate how CQ affects molecular and cellular functions and particularly, copper, iron, and zinc homeostasis.

The conclusion was: 'the results show that CQ can regulate metal homeostasis by binding metal ions, resulting the cell sensing a state of deficiency of bioavailable metal ions while simultaneously increasing available metals to SOD1 (via Ccs1) and possibly some other metalloproteins that can access CQ-bound metals.' [14]

The array produced will be used to predict gene interactions among the first 100 genes that were expressed differentially in the array. Since they are the genes more differentially expressed they should be the genes related to the copper, iron, and zinc homeostasis.

- Dataset2: Short-term fasting effect on skeletal muscle: time course.

Analysis of skeletal muscle biopsies collected at 1.5, 4, 10, and 24hr post-meal time from healthy young men subjected to a 24hr fast. Results provide insight into the molecular mechanisms underlying metabolic adaptation to fasting in healthy humans.

In the article '*Regulation of skeletal muscle energy/nutrient-sensing pathways during metabolic adaptation to fasting in healthy humans'* in 2014 by the authors Marjolein A. Wijngaarden, Leontine E. H. Bakker, Gerard C. van der Zon, Peter A. C. Hoen, Ko Willems van Dijk, Ingrid M. Jazet, Hanno Pijl and Bruno Guigas, the metabolic adaptations during 24 hours fasting was studied.

Under abstinence of food, rapid metabolic adaptations are required to maintain energy homeostasis in the body. The adaptations to the new situation occur in a well-coordinated process that regulates energy/nutrient-sensing pathways, which in turn leads to the activation and repression of specific sets of genes.

The aim of this study was to achieve further information about the pathways that take place during fasting. As stated by the authors in the article: 'As expected, fasting induced a time-dependent decrease in plasma insulin and leptin levels, whereas levels of ketone bodies and free fatty acids increased. This was associated with a metabolic shift from glucose toward lipid oxidation. At the molecular level, activation of the protein kinase B (PKB/Akt) and mammalian target of rapamycin pathways was time-dependently reduced in skeletal muscle during fasting, whereas the AMP-activated protein kinase activity remained unaffected. Furthermore, we report some changes in the phosphorylation and/or content of forkhead protein 1, sirtuin 1, and class IIa histone deacetylase 4, suggesting that these pathways might be involved in the transcriptional adaptation to fasting.’ [15]

The array produced for this study will be used to predict gene interactions among the first 100 genes that were expressed differentially in the array. Since they are the genes more differentially expressed they should be the genes related fasting homeostasis.

5.1. Study network gene interactions Dataset1 (Clioquinol effect on yeast)

5.1.1. Correlation matrix

To study gene expression profiles of yeast treated with CQ, yeast cells were treated with CQ or DMSO for 5 hours and RNAs were extracted for microarray analysis. There were 3 replicates of each, with a total of 6 samples.

From the expression data 100 gens were selected to predict interactions among them, and the matrix correlation was calculated.

Sample from the 7 first genes in the correlation matrix:

	GDE1	ANB1	HEM13	CTR1	ILV3	1773710_at	VMR1
GDE1	1.000	-0.971	-0.999	0.981	-0.998	-0.998	0.997
ANB1	-0.971	1.000	0.979	-0.994	0.979	0.978	-0.978
HEM13	-0.999	0.979	1.000	-0.986	0.999	0.999	-0.998
CTR1	0.981	-0.994	-0.986	1.000	-0.984	-0.982	0.987
ILV3	-0.998	0.979	0.999	-0.984	1.000	0.999	-0.999
1773710_at	-0.998	0.978	0.999	-0.982	0.999	1.000	-0.996
VMR1	0.997	-0.978	-0.998	0.987	-0.999	-0.996	1.000

5.1.2. ACOGeneInteraction

ACOGI was applied to predict interactions adjusting the following parameters:

a) Ant System: Ant colony System

b) Alpha (relative importance placed on pheromones): 1

c) Beta (relative importance placed on distances): 3

d) Number iterations: 100

e) Rho (ratio of evaporated pheromone): 0.8

f) Number ants: 50

g) Times (times is increased sequentially until no more interactions are predicted): 1 to 20

Number of interactions predicted increasing ‘times’ sequentially:

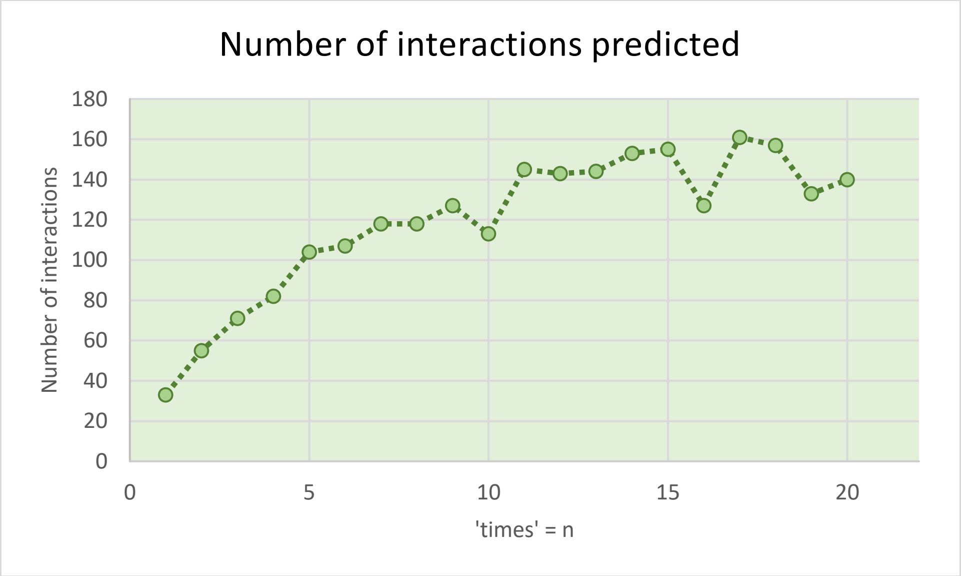


Figure 9. Number of interactions predicted in Dataset1.

As we increase the number of ‘times’ that the algorithm has to go through each node the interactions increase until it turns out to be stable.

Interactions predicted are not the same when using different 'times':

The interactions predicted when ‘times’ = x does not necessarily appear when ‘times’ = x + 1, but if we run the algorithm increasing ‘times’ sequentially and take all the interactions predicted so far into account, at some point ACO does not predict more interactions:

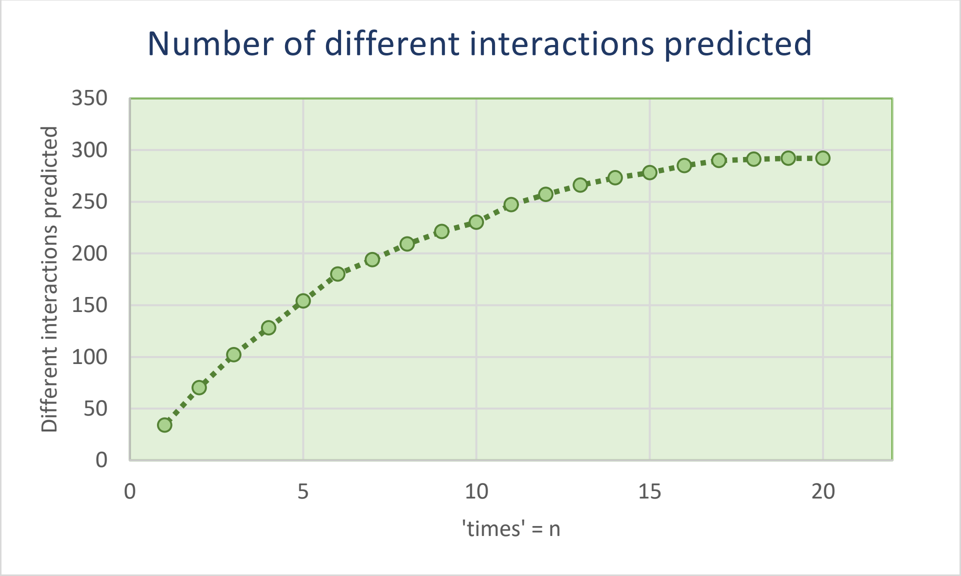


Figure 10. Number of different interactions predicted Dataset1.

As the variable ‘times’ increases the number of different interactions predicted increases exponentially until no more interactions are predicted.

In this case, the total of different interactions predicted is 292, while using the TSP approach (going once through all the nodes) only 33 interactions were predicted.

5.1.3. Table of predicted interactions

Some interactions appear always when using different 'times' while others only appear few times. These interactions can be sorted by the times they have appeared, that can be a way to represent how important is each interaction in the network studied.

Only the 17 first interactions predicted are shown in the table below:

	N	Gene1	Gene2	Description1	Description2
1	20	PCA1	PET10	Cadmium transporting P-type ATPase; may also have a role in copper and iron homeostasis; stabilized by Cd binding, which prevents ubiquitination; S288C and other lab strains contain a G970R mutation which eliminates Cd transport function	Protein of unknown function that co-purifies with lipid particles; expression pattern suggests a role in respiratory growth; computational analysis of large-scale protein-protein interaction data suggests a role in ATP/ADP exchange
2	20	DFG5	HEM13	Putative mannosidase, essential glycosylphosphatidylinositol (GPI)-anchored membrane protein required for cell wall biogenesis in bud formation, involved in filamentous growth, homologous to Dcw1p	Coproporphyrinogen III oxidase, an oxygen requiring enzyme that catalyzes the sixth step in the heme biosynthetic pathway; localizes to the mitochondrial inner membrane; transcription is repressed by oxygen and heme (via Rox1p and Hap1p)
3	20	EGD1	URA1	Subunit beta1 of the nascent polypeptide-associated complex (NAC) involved in protein targeting, associated with cytoplasmic ribosomes; enhances DNA binding of the Gal4p activator; homolog of human BTF3b	Dihydroorotate dehydrogenase, catalyzes the fourth enzymatic step in the de novo biosynthesis of pyrimidines, converting dihydroorotic acid into orotic acid
4	20	ERG1	FAA4	Squalene epoxidase, catalyzes the epoxidation of squalene to 2,3-oxidosqualene; plays an essential role in the ergosterol-biosynthesis pathway and is the specific target of the antifungal drug terbinafine	Long chain fatty acyl-CoA synthetase, regulates protein modification during growth in the presence of ethanol, functions to incorporate palmitic acid into phospholipids and neutral lipids
5	20	INO1	TRP5	Inositol 1-phosphate synthase, involved in synthesis of inositol phosphates and inositol-containing phospholipids; transcription is coregulated with other phospholipid biosynthetic genes by Ino2p and Ino4p, which bind the UASINO DNA element	Tryptophan synthase involved in tryptophan biosynthesis, regulated by the general control system of amino acid biosynthesis
6	20	1779659_at	GDB1	0	Glycogen debranching enzyme containing glucanotranferase and alpha-1,6-amyloglucosidase activities, required for glycogen degradation; phosphorylated in mitochondria
7	20	1778446_at	MCH2	0	Protein with similarity to mammalian monocarboxylate permeases, which are involved in transport of monocarboxylic acids across the plasma membrane; mutant is not deficient in monocarboxylate transport
8	20	RNR2	SSB2	Ribonucleotide-diphosphate reductase (RNR), small subunit; the RNR complex catalyzes the rate-limiting step in dNTP synthesis and is regulated by DNA replication and DNA damage	Cytoplasmic ATPase that is a ribosome-associated molecular chaperone, functions with J-protein partner Zuo1p; may be involved in the folding of

				checkpoint pathways via localization of the small subunits	newly-synthesized polypeptide chains; member of the HSP70 family; homolog of SSB1
9	20	OLE1	ZPS1	Delta(9) fatty acid desaturase, required for monounsaturated fatty acid synthesis and for normal distribution of mitochondria	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH
10	19	TIS11	VBA3 /// VBA5	mRNA-binding protein expressed during iron starvation; binds to a sequence element in the 3'-untranslated regions of specific mRNAs to mediate their degradation; involved in iron homeostasis	Permease of basic amino acids in the vacuolar membrane /// Putative transporter of the Major Facilitator Superfamily (MFS); proposed role as a basic amino acid permease based on phylogeny
11	19	DAN1	TIR3	Cell wall mannoprotein with similarity to Tir1p, Tir2p, Tir3p, and Tir4p; expressed under anaerobic conditions, completely repressed during aerobic growth	Cell wall mannoprotein of the Srp1p/Tip1p family of serine-alanine-rich proteins; expressed under anaerobic conditions and required for anaerobic growth
12	19	AGX1	PDR5	Alanine:glyoxylate aminotransferase (AGT), catalyzes the synthesis of glycine from glyoxylate, which is one of three pathways for glycine biosynthesis in yeast; has similarity to mammalian and plant alanine:glyoxylate aminotransferases	Plasma membrane ATP-binding cassette (ABC) transporter, multidrug transporter actively regulated by Pdr1p; also involved in steroid transport, cation resistance, and cellular detoxification during exponential growth
13	19	FAA4	FIT2	Long chain fatty acyl-CoA synthetase, regulates protein modification during growth in the presence of ethanol, functions to incorporate palmitic acid into phospholipids and neutral lipids	Mannoprotein that is incorporated into the cell wall via a glycosylphosphatidylinositol (GPI) anchor, involved in the retention of siderophore-iron in the cell wall
14	19	AMS1	HTB1	Vacuolar alpha mannosidase, involved in free oligosaccharide (fOS) degradation; delivered to the vacuole in a novel pathway separate from the secretory pathway	Histone H2B, core histone protein required for chromatin assembly and chromosome function; nearly identical to HTB2; Rad6p-Bre1p-Lge1p mediated ubiquitination regulates transcriptional activation, meiotic DSB formation and H3 methylation
15	19	ILV3	TIS11	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids	mRNA-binding protein expressed during iron starvation; binds to a sequence element in the 3'-untranslated regions of specific mRNAs to mediate their degradation; involved in iron homeostasis
16	19	ARN1	PDR10	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to ferrirubin, ferrirhodin, and related siderophores	ATP-binding cassette (ABC) transporter, multidrug transporter involved in the pleiotropic drug resistance network; regulated by Pdr1p and Pdr3p
17	19	CTT1	ICY1	Cytosolic catalase T, has a role in protection from oxidative damage by hydrogen peroxide	Protein of unknown function, required for viability in rich media of cells lacking mitochondrial DNA; mutants have an invasive growth defect with elongated morphology; induced by amino acid starvation

5.1.4. Network predicted using Cytoscape

Representation of the gene interaction network will be done using the Cytoscape. In order to create the graph, an input file containing the direct interactions given in the network as a pair of genes is necessary. The format of the file is simple since it consists of two columns with a first gene that interacts with the second. After that, once the file is imported into Cytoscape, the software shows the interactions as a graph.

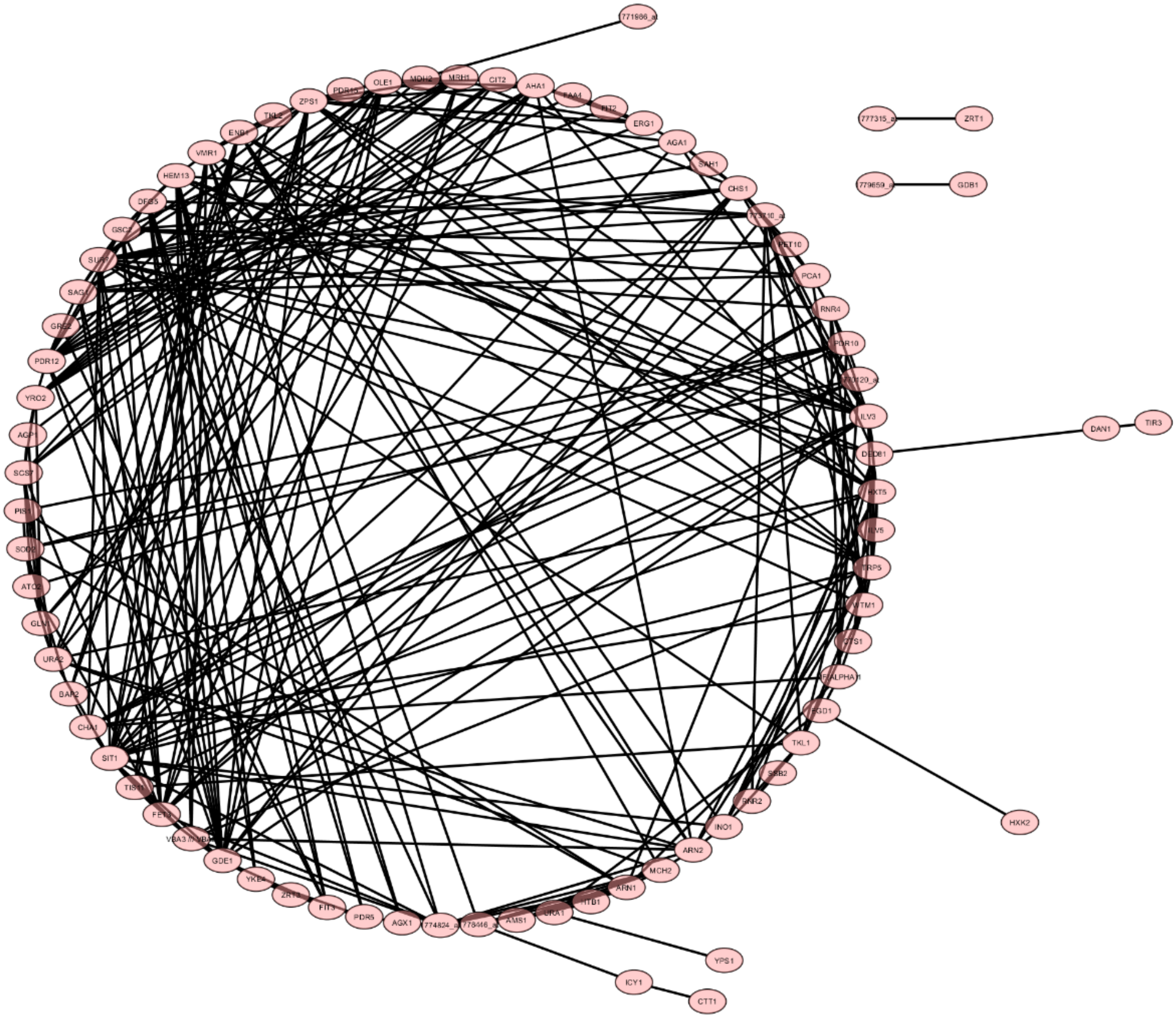


Figure 11. Image network predicted by ACOGeneInteraction using Dataset1.

Image in **Figure 11** represents all the predicted gene interactions, also we can see that some genes have more interactions than others, meaning that these genes have a more central role in the network studied.

5.1.5. Central gens in the network

Unlike ACO-TSP, ACOGeneInteractions allows to predict more than just two interactions per gene. Sorting the gens by the number of interactions with other genes give us an idea how central is each gene in the network.

Genes sorted by the number of interactions with other genes, 20 first genes are shown:

	Gene	N inter	Description
1	ZPS1	17	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH
2	SUR7	14	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants
3	SIT1	14	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p
4	TRP5	12	Tryptophan synthase involved in tryptophan biosynthesis, regulated by the general control system of amino acid biosynthesis
5	URA2	10	Bifunctional carbamoylphosphate synthetase (CPSase)-aspartate transcarbamylase (ATCase), catalyzes the first two enzymatic steps in the de novo biosynthesis of pyrimidines; both activities are subject to feedback inhibition by UTP
6	WTM1	10	Transcriptional modulator involved in regulation of meiosis, silencing, and expression of RNR genes; required for nuclear localization of the ribonucleotide reductase small subunit Rnr2p and Rnr4p; contains WD repeats
7	VMR1	10	Protein of unknown function that may interact with ribosomes, based on co-purification experiments; member of the ATP-binding cassette (ABC) family; potential Cdc28p substrate; detected in purified mitochondria in high-throughput studies
8	PDR12	10	Plasma membrane ATP-binding cassette (ABC) transporter, weak-acid-inducible multidrug transporter required for weak organic acid resistance; induced by sorbate and benzoate and regulated by War1p; mutants exhibit sorbate hypersensitivity
9	VBA3 /// VBA5	9	Permease of basic amino acids in the vacuolar membrane /// Putative transporter of the Major Facilitator Superfamily (MFS); proposed role as a basic amino acid permease based on phylogeny
10	YRO2	9	Putative protein of unknown function; the authentic, non-tagged protein is detected in a phosphorylated state in highly purified mitochondria in high-throughput studies; transcriptionally regulated by Haa1p

11	SAG1	8	Alpha-agglutinin of alpha-cells, binds to Aga1p during agglutination, N-terminal half is homologous to the immunoglobulin superfamily and contains binding site for a-agglutinin, C-terminal half is highly glycosylated and contains GPI anchor
12	OLE1	8	Delta(9) fatty acid desaturase, required for monounsaturated fatty acid synthesis and for normal distribution of mitochondria
13	ILV3	8	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids
14	HEM13	7	Coproporphyrinogen III oxidase, an oxygen requiring enzyme that catalyzes the sixth step in the heme biosynthetic pathway; localizes to the mitochondrial inner membrane; transcription is repressed by oxygen and heme (via Rox1p and Hap1p)
15	SCS7	7	Sphingolipid alpha-hydroxylase, functions in the alpha-hydroxylation of sphingolipid-associated very long chain fatty acids, has both cytochrome b5-like and hydroxylase/desaturase domains, not essential for growth
16	TKL1	7	Transketolase, similar to Tkl2p; catalyzes conversion of xylulose-5-phosphate and ribose-5-phosphate to sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate in the pentose phosphate pathway; needed for synthesis of aromatic amino acids
17	MCH2	6	Protein with similarity to mammalian monocarboxylate permeases, which are involved in transport of monocarboxylic acids across the plasma membrane; mutant is not deficient in monocarboxylate transport
18	SOD2	6	Mitochondrial superoxide dismutase, protects cells against oxygen toxicity; phosphorylated
19	PDR10	5	ATP-binding cassette (ABC) transporter, multidrug transporter involved in the pleiotropic drug resistance network; regulated by Pdr1p and Pdr3p
20	YKE4	5	Zinc transporter; localizes to the ER; null mutant is sensitive to calcofluor white, leads to zinc accumulation in cytosol; ortholog of the mouse KE4 and member of the ZIP (ZRT, IRT-like Protein) family

5.1.6. Key genes

Although the number of total interactions using this new approach increases, the number of key interactions decreases, since here the conditions to be considered ‘key interactions’ are more restrictive.

In previous studies using ACO-TSP (or ‘times’ = 1), all the interactions predicted were considered ‘Key interactions’ based on the statement that ACO looks for the optimal path.

Using this approach, we consider ‘key interactions’ those interactions that appear always when using different ‘times’.

Comparing ‘key interactions’ between both approaches we see that 43 gene interactions were found when using ACO-TSP (‘times’ = 1), but only 20 were considered key interactions after running the ACOGeneInteraction algorithm (using ‘times’ from 1 to 20).

Which means that only 46% of those that appear in ‘key interactions-ACOTSP’ appear in ‘Key interactions-ACOGI’, meaning that using TSP approach are considered ‘key interactions’ some of them that with this new approach are not.

5.2. Study network gene interactions Dataset2 (Short-term fasting effect on skeletal muscle: time course)

5.2.1. Correlation matrix

For the research 12 young healthy men were studied during a 24-h fast. Whole body glucose/lipid oxidation rates were determined by indirect calorimetry, and blood and skeletal muscle biopsies were collected and analyzed at baseline and after 10 and 24 h of fasting.

From the expression data 100 gens were selected to predict interactions among them, and the matrix correlation was calculated.

Sample from the 7 first genes in the correlation matrix:

	ILMN_1697499	ILMN_1663640	ILMN_1715169	ILMN_2165753	ILMN_1779241	ILMN_2246083	ILMN_1665691
ILMN_1697499	1.000	0.089	-0.623	0.281	0.433	-0.177	0.292
ILMN_1663640	0.089	1.000	-0.066	0.060	0.059	-0.391	0.108
ILMN_1715169	-0.623	-0.066	1.000	-0.217	-0.329	0.000	-0.503
ILMN_2165753	0.281	0.060	-0.217	1.000	0.113	-0.038	0.062
ILMN_1779241	0.433	0.059	-0.329	0.113	1.000	0.039	0.324
ILMN_2246083	-0.177	-0.391	0.000	-0.038	0.039	1.000	0.060
ILMN_1665691	0.292	0.108	-0.503	0.062	0.324	0.060	1.000

5.2.2. ACOGeneInteraction

ACOGI was applied to predict interactions adjusting the following parameters:

a) Ant System: Ant colony System
b) Alpha (relative importance placed on pheromones): 1
c) Beta (relative importance placed on distances): 3
d) Number iterations: 100
e) Rho (ratio of evaporated pheromone): 0.8
f) Number ants: 50
g) Times (times is increased sequentially until no more interactions are predicted): 1 to 20

ACO was applied to predict interactions using different 'times':

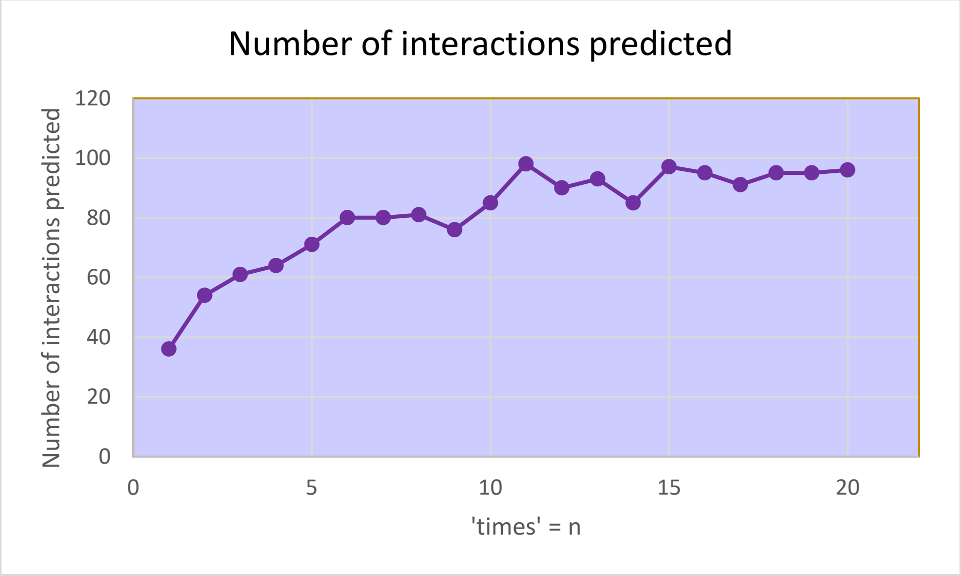


Figure 12. Number of interactions predicted in Dataset2.

As seen in the previous dataset, as the variable 'times' increases the number of different interactions predicted increases exponentially until no more interactions are predicted.

Interactions predicted are not the same when using different 'times':

The interactions predicted when ‘times’ = x does not necessarily appear when ‘times’ = x + 1, but if we run the algorithm increasing ‘times’ sequentially and take all the interactions predicted so far into account, at some point ACO does not predict more interactions:

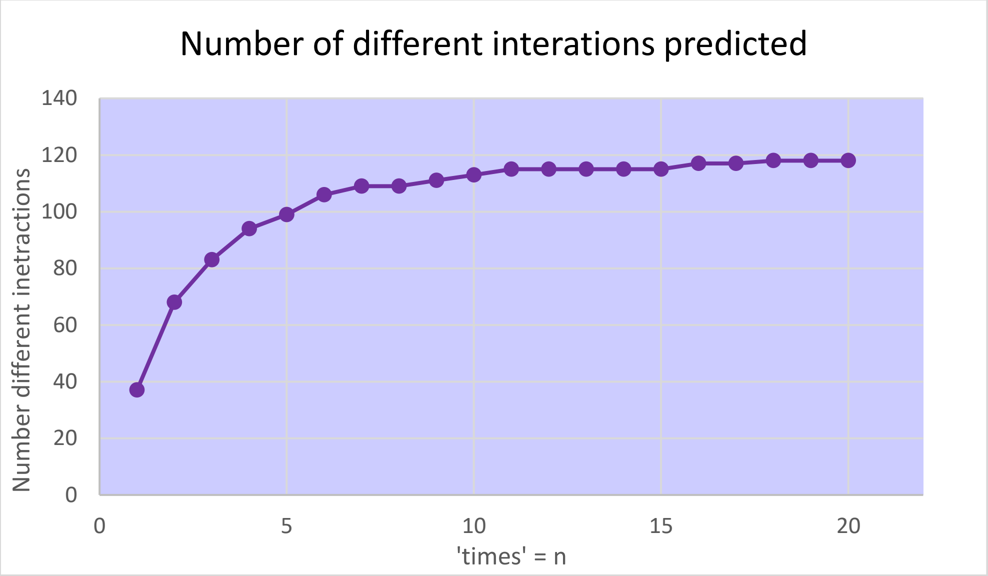


Figure 13. Number of different interactions predicted Dataset2.

The same pattern as in Dataset1 is repeated: as the variable ‘times’ increases the number of different interactions found increases exponentially until no more interactions are predicted. In the figure, we can see that there is a big amount of different interactions found when using ‘times’ = 1 and ‘times’ = 2.

In this case, the total of different interactions predicted is 118, while using the TSP approach (going once through all the nodes) only 37 interactions were predicted.

5.2.3. Table of predicted interactions

Some interactions appear always when using different ‘times’ while others only appear few times. These interactions can be sorted by the ‘times’ they have appeared, that can be a way to represent how important is each interaction in the network studied.

Only the 13 first interactions predicted are shown in the table below:

	N	Gene1	Gene2	Description1	Description2
1	20	SHISA2	SLITRK4	Homo sapiens shisa homolog 2 (Xenopus laevis) (SHISA2), mRNA.	Homo sapiens SLIT and NTRK-like family, member 4 (SLITRK4), mRNA.
2	20	AURKA	FRZB	['Homo sapiens aurora kinase A (AURKA), transcript variant 5, mRNA.'	Homo sapiens frizzled-related protein (FRZB), mRNA.
3	20	IL32	LOC100008589	['Homo sapiens interleukin 32 (IL32), transcript variant 7, mRNA.'	['Homo sapiens 28S ribosomal RNA (LOC100008589), non-coding RNA.'
4	20	THBS4	IL8	Homo sapiens thrombospondin 4 (THBS4), mRNA.	Homo sapiens interleukin 8 (IL8), mRNA.
5	20	SPINK1	ALG1L	Homo sapiens serine peptidase inhibitor, Kazal type 1 (SPINK1), mRNA.	Homo sapiens asparagine-linked glycosylation 1-like (ALG1L), mRNA.
6	20	CCL2	GSTM1	Homo sapiens chemokine (C-C motif) ligand 2 (CCL2), mRNA.	['Homo sapiens glutathione S-transferase M1 (GSTM1), transcript variant 1, mRNA.'
7	20	SLC25A25	SLITRK4	Homo sapiens solute carrier family 25 (mitochondrial carrier; phosphate carrier),	Homo sapiens SLIT and NTRK-like family, member 4 (SLITRK4), mRNA.

				member 25 (SLC25A25), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA.	
8	20	ATP2A1	LOC649143	['Homo sapiens ATPase, Ca++ transporting, cardiac muscle, fast twitch 1 (ATP2A1), transcript variant a, mRNA.'	PREDICTED: Homo sapiens similar to HLA class II histocompatibility antigen, DRB1-9 beta chain precursor (MHC class I antigen DRB1*9) (DR-9) (DR9), transcript variant 2 (LOC649143), mRNA.
9	20	MT1E	AURKA	Homo sapiens metallothionein 1E (MT1E), mRNA.	['Homo sapiens aurora kinase A (AURKA), transcript variant 5, mRNA.'
10	20	PTER	ZP3	Homo sapiens phosphotriesterase related (PTER), transcript variant 2, mRNA.	Homo sapiens zona pellucida glycoprotein 3 (sperm receptor) (ZP3), mRNA.
11	20	GSTT1	FAM174A	Homo sapiens glutathione S-transferase theta 1 (GSTT1), mRNA.	Homo sapiens family with sequence similarity 174, member A (FAM174A), mRNA.
12	20	OLFM1	CD79B	['Homo sapiens olfactomedin 1 (OLFM1), transcript variant 2, mRNA.'	Homo sapiens CD79b molecule, immunoglobulin-associated beta (CD79B), transcript variant 3, mRNA.
13	20	TMEM70	LOC100134364	['Homo sapiens transmembrane protein 70 (TMEM70), transcript variant 1, mRNA.'	PREDICTED: Homo sapiens hypothetical protein LOC100134364 (LOC100134364), mRNA.

5.2.4. Network predicted using Cytoscape

Representation of the gene interaction network will be done using the Cytoscape as in the previous dataset.

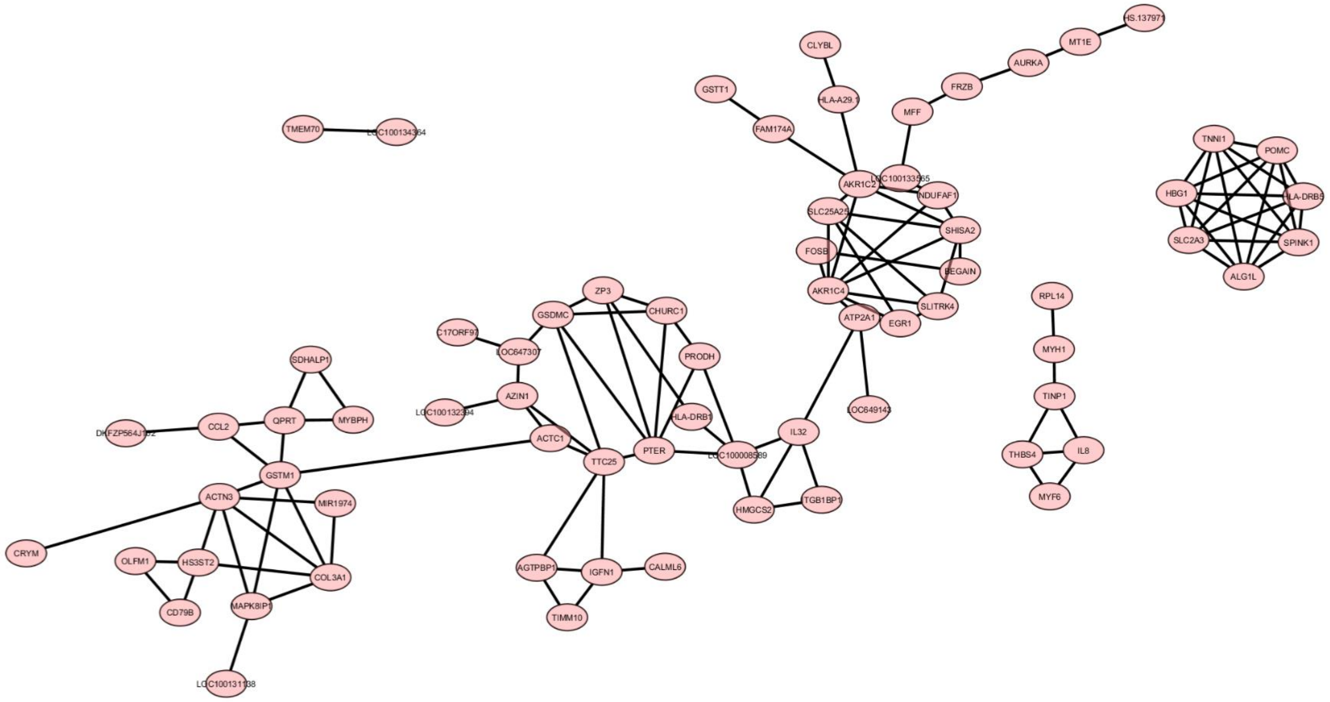


Figure 14. Image network predicted by ACOGeneInteraction using Dataset2.

This image represents all the predicted gene interactions, also we can see that some genes have more interactions than others, meaning that these genes have a more central role in the network studied.

Moreover, we see that there are some groups of genes that are more related between them forming clusters. We have to take into account that in both datasets has been used the same layout for creating the graph while the graphs obtained are very different.

5.2.5. Central genes in the network

Genes sorted by the number of interactions with other genes:

	Gene	N times	Description
1	ALG1L	6	Homo sapiens asparagine-linked glycosylation 1-like (ALG1L), mRNA.
2	AKR1C2	6	Homo sapiens aldo-keto reductase family 1, member C2 (dihydrodiol dehydrogenase 2; bile acid binding protein; 3-alpha hydroxysteroid dehydrogenase, type III) (AKR1C2), transcript variant 1, mRNA. XM_943424 XM_943425 XM_943427
3	HBG1	5	Homo sapiens hemoglobin, gamma A (HBG1), mRNA.
4	SLITRK4	4	Homo sapiens SLIT and NTRK-like family, member 4 (SLITRK4), mRNA.
5	LOC100008589	4	Homo sapiens 28S ribosomal RNA (LOC100008589), non-coding RNA.
6	GSTM1	4	Homo sapiens glutathione S-transferase M1 (GSTM1), transcript variant 1, mRNA.
7	ZP3	4	Homo sapiens zona pellucida glycoprotein 3 (sperm receptor) (ZP3), mRNA.
8	COL3A1	4	Homo sapiens collagen, type III, alpha 1 (COL3A1), mRNA.
9	SPINK1	4	Homo sapiens serine peptidase inhibitor, Kazal type 1 (SPINK1), mRNA.
10	PTER	4	Homo sapiens phosphotriesterase related (PTER), transcript variant 2, mRNA.
11	IL8	3	Homo sapiens interleukin 8 (IL8), mRNA.
12	SLC2A3	3	Homo sapiens solute carrier family 2 (facilitated glucose transporter), member 3 (SLC2A3), mRNA.
13	LOC100133565	3	PREDICTED: Homo sapiens similar to hCG23738 (LOC100133565), mRNA.
14	SLC25A25	3	Homo sapiens solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 25 (SLC25A25), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA.
15	AGTPBP1	3	Homo sapiens ATP/GTP binding protein 1 (AGTPBP1), mRNA.
16	CHURC1	3	Homo sapiens churchill domain containing 1 (CHURC1), mRNA.
17	LOC647307	3	PREDICTED: Homo sapiens misc_RNA (LOC647307), miscRNA.
18	SHISA2	3	Homo sapiens shisa homolog 2 (Xenopus laevis) (SHISA2), mRNA.
19	CD79B	2	Homo sapiens CD79b molecule, immunoglobulin-associated beta (CD79B), transcript variant 3, mRNA.

5.2.6. Key genes

As said in the previous dataset, using this new approach, we consider 'key interactions' those interaction that appear always when using different 'times'.

Hence, although the number of total interactions using this new approach increases, the number of key interactions decreases, since here the conditions to be considered 'key interactions' are more restrictive.

In previous studies using ACO-TSP (or 'times' = 1), all the interactions predicted were considered 'Key interactions' based on the statement that ACO looks for the optimal path.

Comparing 'key interactions' between both approaches we see that 48 gene interactions were found when using ACO-TSP ('times' = 1), but only 24 were considered key interactions after running the ACOGeneInteractions algorithm (using 'times' from 1 to 20).

Which means that only 50% of those that appear in 'key interactions-ACOTSP' appear in 'Key interactions-ACOGI', meaning that using TSP approach are considered 'key interactions' some of them that with this new approach are not.

6. CONCLUSIONS

- The modification (ACOGeneInteraction) allows us to predict more interactions than using ACO-TSP approach, meaning that the limitation of only being able to predict as many interactions as number of nodes present in the graph has been improved.

- Running the algorithm sequentially increasing the number of 'times' allows us to sort the interactions giving an additional confidence that the predicted interaction may exist. Those that appear always (or several times) when using different 'times' are more probably to exist than those that appear just one time.

- Using this new approach, we consider 'key interactions' those interaction that appear always when using different 'times'. Hence, although the number of total interactions using this new approach increases, the number of key interactions decreases, since the conditions to be considered 'key interactions' are more restrictive.

- Using ACOGeneInteractions allows to predict more than just two interactions per gene. Therefore, sorting the genes by the number of interactions with other genes represents which genes have a more central role in the network.

Future directions:

It would be interesting that the interactions found by ACOGeneInteraction could be tested in the lab or by other means to verify if the predictions made are consistent.

The main problem found when using the algorithm has been the time it takes when working with big numbers of 'times'. In our case, datasets have 100 genes and when 'times' is bigger than 15 takes more than 1 hour. Using faster algorithms would allow to work with bigger networks getting a wider view how the cells works internally.

Self-evaluation:

Through all the work performed during the final project the objectives have been achieved. Also, deadlines have been met, although some minor modifications were applied when needed.

The modifications applied to the ACO algorithm for predicting gene interactions have given a positive result, allowing us to work with a new approach that is able to predict more interactions, sort those interactions by the number of times they appear (which gives an idea which interactions are more ‘key’ than others), and also, sort the genes by the number of interactions with other genes (central gens in the network).

7. GLOSSARY

ACO:	Ant colony optimization
ACS:	Ant colony system
ACOGI:	Ant colony optimization for gene interactions
ACOTSP:	Ant colony optimization along with traveling salesman problem
AD:	Alzheimer’s disease
AS:	Ant system
CQ:	Clioquinol
DMSO:	Dimethyl sulfoxide
DNA:	Deoxyribonucleic acid
GRN:	Gene regulatory network
PEC:	‘Prova d’avaluació continua’
PIN:	Protein interaction network
RNA:	Ribonucleic acid
SI:	Swarm intelligence
TFM:	‘Treball Final de Master’ (Final project)
TSP:	Traveling salesman problem

8. BIBLIOGRAPHY

1. URL: http://www.techrepublic.com/blog/it-security/swarm-intelligence-are-digital-ants-the-answer-to-malware/, April 2017.

2. Frank Emmert-Streib, Matthias Dehmer, & Benjamin Haibe-Kains, ‘Gene regulatory networks and their applications: understanding biological and medical problems in terms of networks’, Front Cell Dev Biol., 2: 38. 2014.

3. Bianchi, Leonora; Marco Dorigo; Luca Maria Gambardella & Walter J. Gutjahr (2009). "A survey on metaheuristics for stochastic combinatorial optimization". *Natural Computing: an international journal*. 8 (2): 239–287, 2008.

4. Blum, C. & Roli, A., ‘Metaheuristics in combinatorial optimization: Overview and conceptual comparison’, 35 (3). ACM Computing Surveys: 268–308, 2003.

5. Eric Bonabeau, Marco Dorigo & Guy Theraulaz “SWARM INTELLIGENCE From Natural to Artificial Systems” Oxford University Press, New York, 1999.

6. URL: http://www.cleveralgorithms.com/nature-inspired/swarm.html, April 2017

7. Morteza Atabati , Kobra Zarei & Azam Borhani, Ant colony optimization as a descriptor selection in QSPR modeling: Estimation of the λmax of anthraquinones-based dyes’, Journal of Saudi Chemical Society 293(S1), April 2013. [Image Fig1]

8.URL:https://optimization.mccormick.northwestern.edu/index.php/Traveling_salesman_problems, April 2017.

9. Gurkan Bebek, “Identifying Gene Interaction Networks”, Methods Mol Biol., 850: 483–494, 2012.

10. Raza, K. & Kohli, M. “Ant colony optimization for inferring key gene interactions” 2nd International Conference on Computing for Sustainable Global Development, p. 1242-1246, IEEE. (IEEE Xplore), 2015.

11. Jamaludin Sallim, Rosni Abdullah & Ahamad Tajudin Khader "Ant Colony Optimization for optimal path finding in protein interaction network", 2008 First International Conference on Distributed Framework and Applications, Page(s):41 – 44, 2008.

12. URL: https://mla-lima.wikispaces.com/Southern+Blotting+and+DNA+Microarray, May 2017.

13. Brazma A & Vilo J. 'Gene expression data analysis.' FEBS Letters, Aug 25;480(1):17-24, 2000.

14. Li C., Wang J. & Zhou B., 'The metal chelating and chaperoning effects of clioquinol: insights from yeast studies', Journal Alzheimer's Disease. 21(4):1249-62, 2010.

15. Marjolein A. Wijngaarden, Leontine E. H. Bakker, Gerard C. van der Zon, Peter A. C. Hoen, Ko Willems van Dijk, Ingrid M. Jazet, Hanno Pijl & Bruno Guigas, "Regulation of skeletal muscle energy/nutrient-sensing pathways during metabolic adaptation to fasting in healthy humans", 2014.

9. ANNEXES

Annex 1. Table of predicted interactions Dataset1

	N	Gene1	Gene2	Description1	Description2
1	20	PCA1	PET10	Cadmium transporting P-type ATPase; may also have a role in copper and iron homeostasis; stabilized by Cd binding, which prevents ubiquitination; S288C and other lab strains contain a G970R mutation which eliminates Cd transport function	Protein of unknown function that co-purifies with lipid particles; expression pattern suggests a role in respiratory growth; computational analysis of large-scale protein-protein interaction data suggests a role in ATP/ADP exchange
2	20	DFG5	HEM13	Putative mannosidase, essential glycosylphosphatidylinositol (GPI)-anchored membrane protein required for cell wall biogenesis in bud formation, involved in filamentous growth, homologous to Dcw1p	Coproporphyrinogen III oxidase, an oxygen requiring enzyme that catalyzes the sixth step in the heme biosynthetic pathway; localizes to the mitochondrial inner membrane; transcription is repressed by oxygen and heme (via Rox1p and Hap1p)
3	20	EGD1	URA1	Subunit beta1 of the nascent polypeptide-associated complex (NAC) involved in protein targeting, associated with cytoplasmic ribosomes; enhances DNA binding of the Gal4p activator; homolog of human BTF3b	Dihydroorotate dehydrogenase, catalyzes the fourth enzymatic step in the de novo biosynthesis of pyrimidines, converting dihydroorotic acid into orotic acid
4	20	ERG1	FAA4	Squalene epoxidase, catalyzes the epoxidation of squalene to 2,3-oxidosqualene; plays an essential role in the ergosterol-biosynthesis pathway and is the specific target of the antifungal drug terbinafine	Long chain fatty acyl-CoA synthetase, regulates protein modification during growth in the presence of ethanol, functions to incorporate palmitic acid into phospholipids and neutral lipids
5	20	INO1	TRP5	Inositol 1-phosphate synthase, involved in synthesis of inositol phosphates and inositol-containing phospholipids; transcription is coregulated with other phospholipid biosynthetic genes by Ino2p and Ino4p, which bind the UASINO DNA element	Tryptophan synthase involved in tryptophan biosynthesis, regulated by the general control system of amino acid biosynthesis
6	20	1779659_at	GDB1	0	Glycogen debranching enzyme containing glucanotranferase and alpha-1,6-amyloglucosidase activities, required for glycogen degradation; phosphorylated in mitochondria
7	20	1778446_at	MCH2	0	Protein with similarity to mammalian monocarboxylate permeases, which are involved in transport of monocarboxylic acids across the plasma membrane; mutant is not deficient in monocarboxylate transport
8	20	RNR2	SSB2	Ribonucleotide-diphosphate reductase (RNR), small subunit; the RNR complex catalyzes the rate-limiting step in dNTP synthesis and is regulated by DNA replication and DNA damage checkpoint pathways via localization of the small subunits	Cytoplasmic ATPase that is a ribosome-associated molecular chaperone, functions with J-protein partner Zuo1p; may be involved in the folding of newly-synthesized polypeptide chains; member of the HSP70 family; homolog of SSB1
9	20	OLE1	ZPS1	Delta(9) fatty acid desaturase, required for monounsaturated fatty acid synthesis and for normal distribution of mitochondria	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH
10	19	TIS11	VBA3 /// VBA5	mRNA-binding protein expressed during iron starvation; binds to a sequence element in the 3'-untranslated regions of specific mRNAs to mediate their degradation; involved in iron homeostasis	Permease of basic amino acids in the vacuolar membrane /// Putative transporter of the Major Facilitator Superfamily (MFS); proposed role as a basic amino acid permease based on phylogeny
11	19	DAN1	TIR3	Cell wall mannoprotein with similarity to Tir1p, Tir2p, Tir3p, and Tir4p; expressed under anaerobic conditions, completely repressed during aerobic growth	Cell wall mannoprotein of the Srp1p/Tip1p family of serine-alanine-rich proteins; expressed under anaerobic conditions and required for anaerobic growth
12	19	AGX1	PDR5	Alanine:glyoxylate aminotransferase (AGT), catalyzes the synthesis of glycine from glyoxylate, which is one of three pathways for glycine biosynthesis in yeast; has similarity to mammalian and plant alanine:glyoxylate aminotransferases	Plasma membrane ATP-binding cassette (ABC) transporter, multidrug transporter actively regulated by Pdr1p; also involved in steroid transport, cation resistance, and cellular detoxification during exponential growth
13	19	FAA4	FIT2	Long chain fatty acyl-CoA synthetase, regulates protein modification during growth in the presence of ethanol, functions to incorporate palmitic acid into phospholipids and neutral lipids	Mannoprotein that is incorporated into the cell wall via a glycosylphosphatidylinositol (GPI) anchor, involved in the retention of siderophore-iron in the cell wall
14	19	AMS1	HTB1	Vacuolar alpha mannosidase, involved in free oligosaccharide (fOS) degradation; delivered to the vacuole in a novel pathway separate from the secretory pathway	Histone H2B, core histone protein required for chromatin assembly and chromosome function; nearly identical to HTB2; Rad6p-Bre1p-Lge1p mediated ubiquitination regulates transcriptional activation, meiotic DSB formation and H3 methylation
15	19	ILV3	TIS11	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids	mRNA-binding protein expressed during iron starvation; binds to a sequence element in the 3'-untranslated regions of specific mRNAs to mediate their degradation; involved in iron homeostasis

16	19	ARN1	PDR10	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to ferrirubin, ferrirhodin, and related siderophores	ATP-binding cassette (ABC) transporter, multidrug transporter involved in the pleiotropic drug resistance network; regulated by Pdr1p and Pdr3p
17	19	CTT1	ICY1	Cytosolic catalase T, has a role in protection from oxidative damage by hydrogen peroxide	Protein of unknown function, required for viability in rich media of cells lacking mitochondrial DNA; mutants have an invasive growth defect with elongated morphology; induced by amino acid starvation
18	18	1771986_at	MDH2	0	Cytoplasmic malate dehydrogenase, one of three isozymes that catalyze interconversion of malate and oxaloacetate; involved in the glyoxylate cycle and gluconeogenesis during growth on two-carbon compounds; interacts with Pck1p and Fbp1
19	18	AGA1	SAH1	Anchorage subunit of a-agglutinin of a-cells, highly O-glycosylated protein with N-terminal secretion signal and C-terminal signal for addition of GPI anchor to cell wall, linked to adhesion subunit Aga2p via two disulfide bonds	S-adenosyl-L-homocysteine hydrolase, catabolizes S-adenosyl-L-homocysteine which is formed after donation of the activated methyl group of S-adenosyl-L-methionine (AdoMet) to an acceptor
20	18	YKE4	ZRT3	Zinc transporter; localizes to the ER; null mutant is sensitive to calcofluor white, leads to zinc accumulation in cytosol; ortholog of the mouse KE4 and member of the ZIP (ZRT, IRT-like Protein) family	Vacuolar membrane zinc transporter, transports zinc from storage in the vacuole to the cytoplasm when needed; transcription is induced under conditions of zinc deficiency
21	18	1770120_at	CHA1	0	Catabolic L-serine (L-threonine) deaminase, catalyzes the degradation of both L-serine and L-threonine; required to use serine or threonine as the sole nitrogen source, transcriptionally induced by serine and threonine
22	18	CIT2	SCS7	Citrate synthase, catalyzes the condensation of acetyl coenzyme A and oxaloacetate to form citrate, peroxisomal isozyme involved in glyoxylate cycle; expression is controlled by Rtg1p and Rtg2p transcription factors	Sphingolipid alpha-hydroxylase, functions in the alpha-hydroxylation of sphingolipid-associated very long chain fatty acids, has both cytochrome b5-like and hydroxylase/desaturase domains, not essential for growth
23	18	PDR10	SOD2	ATP-binding cassette (ABC) transporter, multidrug transporter involved in the pleiotropic drug resistance network; regulated by Pdr1p and Pdr3p	Mitochondrial superoxide dismutase, protects cells against oxygen toxicity; phosphorylated
24	18	DED81	TKL1	Cytosolic asparaginyl-tRNA synthetase, required for protein synthesis, catalyzes the specific attachment of asparagine to its cognate tRNA	Transketolase, similar to Tkl2p; catalyzes conversion of xylulose-5-phosphate and ribose-5-phosphate to sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate in the pentose phosphate pathway; needed for synthesis of aromatic amino acids
25	18	GLN1	SCS7	Glutamine synthetase (GS), synthesizes glutamine from glutamate and ammonia; with Glt1p, forms the secondary pathway for glutamate biosynthesis from ammonia; expression regulated by nitrogen source and by amino acid limitation	Sphingolipid alpha-hydroxylase, functions in the alpha-hydroxylation of sphingolipid-associated very long chain fatty acids, has both cytochrome b5-like and hydroxylase/desaturase domains, not essential for growth
26	18	BAP2	TKL2	High-affinity leucine permease, functions as a branched-chain amino acid permease involved in the uptake of leucine, isoleucine and valine; contains 12 predicted transmembrane domains	Transketolase, similar to Tkl1p; catalyzes conversion of xylulose-5-phosphate and ribose-5-phosphate to sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate in the pentose phosphate pathway; needed for synthesis of aromatic amino acids
27	17	1773710_at	DED81	0	Cytosolic asparaginyl-tRNA synthetase, required for protein synthesis, catalyzes the specific attachment of asparagine to its cognate tRNA
28	17	1774824_at	1778446_at	0	0
29	17	ILV5	INO1	Acetohydroxyacid reductoisomerase, mitochondrial protein involved in branched-chain amino acid biosynthesis, also required for maintenance of wild-type mitochondrial DNA and found in mitochondrial nucleoids	Inositol 1-phosphate synthase, involved in synthesis of inositol phosphates and inositol-containing phospholipids; transcription is coregulated with other phospholipid biosynthetic genes by Ino2p and Ino4p, which bind the UASINO DNA element
30	17	CIT2	YRO2	Citrate synthase, catalyzes the condensation of acetyl coenzyme A and oxaloacetate to form citrate, peroxisomal isozyme involved in glyoxylate cycle; expression is controlled by Rtg1p and Rtg2p transcription factors	Putative protein of unknown function; the authentic, non-tagged protein is detected in a phosphorylated state in highly purified mitochondria in high-throughput studies; transcriptionally regulated by Haa1p
31	17	CIT2	GLN1	Citrate synthase, catalyzes the condensation of acetyl coenzyme A and oxaloacetate to form citrate, peroxisomal isozyme involved in glyoxylate cycle; expression is controlled by Rtg1p and Rtg2p transcription factors	Glutamine synthetase (GS), synthesizes glutamine from glutamate and ammonia; with Glt1p, forms the secondary pathway for glutamate biosynthesis from ammonia; expression regulated by nitrogen source and by amino acid limitation
32	17	FET3	SAG1	Ferro-O2-oxidoreductase required for high-affinity iron uptake and involved in mediating resistance to copper ion toxicity, belongs to class of integral membrane multicopper oxidases	Alpha-agglutinin of alpha-cells, binds to Aga1p during agglutination, N-terminal half is homologous to the immunoglobulin superfamily and contains binding site for a-agglutinin, C-terminal half is highly glycosylated and contains GPI anchor

33	17	ENB1	URA2	Endosomal ferric enterobactin transporter, expressed under conditions of iron deprivation; member of the major facilitator superfamily; expression is regulated by Rcs1p and affected by chloroquine treatment	Bifunctional carbamoylphosphate synthetase (CPSase)-aspartate transcarbamylase (ATCase), catalyzes the first two enzymatic steps in the de novo biosynthesis of pyrimidines; both activities are subject to feedback inhibition by UTP
34	17	RNR4	WTM1	Ribonucleotide-diphosphate reductase (RNR), small subunit; the RNR complex catalyzes the rate-limiting step in dNTP synthesis and is regulated by DNA replication and DNA damage checkpoint pathways via localization of the small subunits	Transcriptional modulator involved in regulation of meiosis, silencing, and expression of RNR genes; required for nuclear localization of the ribonucleotide reductase small subunit Rnr2p and Rnr4p; contains WD repeats
35	17	FIT3	GRE2	Mannoprotein that is incorporated into the cell wall via a glycosylphosphatidylinositol (GPI) anchor, involved in the retention of siderophore-iron in the cell wall	3-methylbutanal reductase and NADPH-dependent methylglyoxal reductase (D-lactaldehyde dehydrogenase); stress induced (osmotic, ionic, oxidative, heat shock and heavy metals); regulated by the HOG pathway
36	17	MF(ALPHA)1	PDR10	Mating pheromone alpha-factor, made by alpha cells; interacts with mating type a cells to induce cell cycle arrest and other responses leading to mating; also encoded by MF(ALPHA)2, although MF(ALPHA)1 produces most alpha-factor	ATP-binding cassette (ABC) transporter, multidrug transporter involved in the pleiotropic drug resistance network; regulated by Pdr1p and Pdr3p
37	16	PIS1	SOD2	Phosphatidylinositol synthase, required for biosynthesis of phosphatidylinositol, which is a precursor for polyphosphoinositides, sphingolipids, and glycolipid anchors for some of the plasma membrane proteins	Mitochondrial superoxide dismutase, protects cells against oxygen toxicity; phosphorylated
38	16	AHA1	FIT2	Co-chaperone that binds to Hsp82p and activates its ATPase activity; similar to Hch1p; expression is regulated by stresses such as heat shock	Mannoprotein that is incorporated into the cell wall via a glycosylphosphatidylinositol (GPI) anchor, involved in the retention of siderophore-iron in the cell wall
39	16	ARN2	VMR1	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to the siderophore triacetylfusarinine C	Protein of unknown function that may interact with ribosomes, based on co-purification experiments; member of the ATP-binding cassette (ABC) family; potential Cdc28p substrate; detected in purified mitochondria in high-throughput studies
40	16	ILV3	VBA3 /// VBA5	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids	Permease of basic amino acids in the vacuolar membrane /// Putative transporter of the Major Facilitator Superfamily (MFS); proposed role as a basic amino acid permease based on phylogeny
41	16	HXT5	ILV5	Hexose transporter with moderate affinity for glucose, induced in the presence of non-fermentable carbon sources, induced by a decrease in growth rate, contains an extended N-terminal domain relative to other HXTs	Acetohydroxyacid reductoisomerase, mitochondrial protein involved in branched-chain amino acid biosynthesis, also required for maintenance of wild-type mitochondrial DNA and found in mitochondrial nucleoids
42	16	PDR5	YKE4	Plasma membrane ATP-binding cassette (ABC) transporter, multidrug transporter actively regulated by Pdr1p; also involved in steroid transport, cation resistance, and cellular detoxification during exponential growth	Zinc transporter; localizes to the ER; null mutant is sensitive to calcofluor white, leads to zinc accumulation in cytosol; ortholog of the mouse KE4 and member of the ZIP (ZRT, IRT-like Protein) family
43	16	1777315_at	ZRT1	0	High-affinity zinc transporter of the plasma membrane, responsible for the majority of zinc uptake; transcription is induced under low-zinc conditions by the Zap1p transcription factor
44	16	1774824_at	AGX1	0	Alanine:glyoxylate aminotransferase (AGT), catalyzes the synthesis of glycine from glyoxylate, which is one of three pathways for glycine biosynthesis in yeast; has similarity to mammalian and plant alanine:glyoxylate aminotransferases
45	16	ERG1	PDR12	Squalene epoxidase, catalyzes the epoxidation of squalene to 2,3-oxidosqualene; plays an essential role in the ergosterol-biosynthesis pathway and is the specific target of the antifungal drug terbinafine	Plasma membrane ATP-binding cassette (ABC) transporter, weak-acid-inducible multidrug transporter required for weak organic acid resistance; induced by sorbate and benzoate and regulated by War1p; mutants exhibit sorbate hypersensitivity
46	16	1773710_at	HEM13	0	Coproporphyrinogen III oxidase, an oxygen requiring enzyme that catalyzes the sixth step in the heme biosynthetic pathway; localizes to the mitochondrial inner membrane; transcription is repressed by oxygen and heme (via Rox1p and Hap1p)
47	16	GSC2	SAG1	Catalytic subunit of 1,3-beta-glucan synthase, involved in formation of the inner layer of the spore wall; activity positively regulated by Rho1p and negatively by Smk1p; has similarity to an alternate catalytic subunit, Fks1p (Gsc1p)	Alpha-agglutinin of alpha-cells, binds to Aga1p during agglutination, N-terminal half is homologous to the immunoglobulin superfamily and contains binding site for a-agglutinin, C-terminal half is highly glycosylated and contains GPI anchor
48	15	SIT1	ZPS1	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH
49	15	1770120_at	SOD2	0	Mitochondrial superoxide dismutase, protects cells against oxygen toxicity; phosphorylated

50	15	CHS1	MRH1	Chitin synthase I, requires activation from zymogenic form in order to catalyze the transfer of N-acetylglucosamine (GlcNAc) to chitin; required for repairing the chitin septum during cytokinesis; transcription activated by mating factor	Protein that localizes primarily to the plasma membrane, also found at the nuclear envelope; the authentic, non-tagged protein is detected in mitochondria in a phosphorylated state; has similarity to Hsp30p and Yro2p
51	15	MDH2	MRH1	Cytoplasmic malate dehydrogenase, one of three isozymes that catalyze interconversion of malate and oxaloacetate; involved in the glyoxylate cycle and gluconeogenesis during growth on two-carbon compounds; interacts with Pck1p and Fbp1	Protein that localizes primarily to the plasma membrane, also found at the nuclear envelope; the authentic, non-tagged protein is detected in mitochondria in a phosphorylated state; has similarity to Hsp30p and Yro2p
52	15	CHS1	SUR7	Chitin synthase I, requires activation from zymogenic form in order to catalyze the transfer of N-acetylglucosamine (GlcNAc) to chitin; required for repairing the chitin septum during cytokinesis; transcription activated by mating factor	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants
53	15	AGP1	YRO2	Low-affinity amino acid permease with broad substrate range, involved in uptake of asparagine, glutamine, and other amino acids; expression is regulated by the SPS plasma membrane amino acid sensor system (Ssy1p-Ptr3p-Ssy5p)	Putative protein of unknown function; the authentic, non-tagged protein is detected in a phosphorylated state in highly purified mitochondria in high-throughput studies; transcriptionally regulated by Haa1p
54	15	ARN1	PIS1	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to ferrirubin, ferrirhodin, and related siderophores	Phosphatidylinositol synthase, required for biosynthesis of phosphatidylinositol, which is a precursor for polyphosphoinositides, sphingolipids, and glycolipid anchors for some of the plasma membrane proteins
55	15	ARN1	MF(ALPHA)1	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to ferrirubin, ferrirhodin, and related siderophores	Mating pheromone alpha-factor, made by alpha cells; interacts with mating type a cells to induce cell cycle arrest and other responses leading to mating; also encoded by MF(ALPHA)2, although MF(ALPHA)1 produces most alpha-factor
56	15	MDH2	YRO2	Cytoplasmic malate dehydrogenase, one of three isozymes that catalyze interconversion of malate and oxaloacetate; involved in the glyoxylate cycle and gluconeogenesis during growth on two-carbon compounds; interacts with Pck1p and Fbp1	Putative protein of unknown function; the authentic, non-tagged protein is detected in a phosphorylated state in highly purified mitochondria in high-throughput studies; transcriptionally regulated by Haa1p
57	15	ILV3	SIT1	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p
58	15	AHA1	ERG1	Co-chaperone that binds to Hsp82p and activates its ATPase activity; similar to Hch1p; expression is regulated by stresses such as heat shock	Squalene epoxidase, catalyzes the epoxidation of squalene to 2,3-oxidosqualene; plays an essential role in the ergosterol-biosynthesis pathway and is the specific target of the antifungal drug terbinafine
59	14	CHS1	PCA1	Chitin synthase I, requires activation from zymogenic form in order to catalyze the transfer of N-acetylglucosamine (GlcNAc) to chitin; required for repairing the chitin septum during cytokinesis; transcription activated by mating factor	Cadmium transporting P-type ATPase; may also have a role in copper and iron homeostasis; stabilized by Cd binding, which prevents ubiquitination; S288C and other lab strains contain a G970R mutation which eliminates Cd transport function
60	14	DFG5	GDE1	Putative mannosidase, essential glycosylphosphatidylinositol (GPI)-anchored membrane protein required for cell wall biogenesis in bud formation, involved in filamentous growth, homologous to Dcw1p	Glycerophosphocholine (GroPCho) phosphodiesterase; hydrolyzes GroPCho to choline and glycerolphosphate, for use as a phosphate source and as a precursor for phosphocholine synthesis; may interact with ribosomes
61	14	SIT1	SUR7	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants
62	14	GLN1	YRO2	Glutamine synthetase (GS), synthesizes glutamine from glutamate and ammonia; with Glt1p, forms the secondary pathway for glutamate biosynthesis from ammonia; expression regulated by nitrogen source and by amino acid limitation	Putative protein of unknown function; the authentic, non-tagged protein is detected in a phosphorylated state in highly purified mitochondria in high-throughput studies; transcriptionally regulated by Haa1p
63	14	CHS1	PET10	Chitin synthase I, requires activation from zymogenic form in order to catalyze the transfer of N-acetylglucosamine (GlcNAc) to chitin; required for repairing the chitin septum during cytokinesis; transcription activated by mating factor	Protein of unknown function that co-purifies with lipid particles; expression pattern suggests a role in respiratory growth; computational analysis of large-scale protein-protein interaction data suggests a role in ATP/ADP exchange
64	14	BAP2	SOD2	High-affinity leucine permease, functions as a branched-chain amino acid permease involved in the uptake of leucine, isoleucine and valine; contains 12 predicted transmembrane domains	Mitochondrial superoxide dismutase, protects cells against oxygen toxicity; phosphorylated
65	14	1770120_at	BAP2	0	High-affinity leucine permease, functions as a branched-chain amino acid permease involved in the uptake of

					leucine, isoleucine and valine; contains 12 predicted transmembrane domains
66	14	1773710_at	TKL1	0	Transketolase, similar to Tkl2p; catalyzes conversion of xylulose-5-phosphate and ribose-5-phosphate to sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate in the pentose phosphate pathway; needed for synthesis of aromatic amino acids
67	14	HXT5	INO1	Hexose transporter with moderate affinity for glucose, induced in the presence of non-fermentable carbon sources, induced by a decrease in growth rate, contains an extended N-terminal domain relative to other HXTs	Inositol 1-phosphate synthase, involved in synthesis of inositol phosphates and inositol-containing phospholipids; transcription is coregulated with other phospholipid biosynthetic genes by Ino2p and Ino4p, which bind the UASINO DNA element
68	14	ERG1	FIT2	Squalene epoxidase, catalyzes the epoxidation of squalene to 2,3-oxidosqualene; plays an essential role in the ergosterol-biosynthesis pathway and is the specific target of the antifungal drug terbinafine	Mannoprotein that is incorporated into the cell wall via a glycosylphosphatidylinositol (GPI) anchor, involved in the retention of siderophore-iron in the cell wall
69	14	OLE1	PDR15	Delta(9) fatty acid desaturase, required for monounsaturated fatty acid synthesis and for normal distribution of mitochondria	Plasma membrane ATP binding cassette (ABC) transporter, multidrug transporter and general stress response factor implicated in cellular detoxification; regulated by Pdr1p, Pdr3p and Pdr8p; promoter contains a PDR responsive element
70	13	ATO2	FET3	Putative transmembrane protein involved in export of ammonia, a starvation signal that promotes cell death in aging colonies; phosphorylated in mitochondria; member of the TC 9.B.33 YaaH family; homolog of Ady2p and Y. lipolytica Gpr1p	Ferro-O2-oxidoreductase required for high-affinity iron uptake and involved in mediating resistance to copper ion toxicity, belongs to class of integral membrane multicopper oxidases
71	13	1773710_at	RNR2	0	Ribonucleotide-diphosphate reductase (RNR), small subunit; the RNR complex catalyzes the rate-limiting step in dNTP synthesis and is regulated by DNA replication and DNA damage checkpoint pathways via localization of the small subunits
72	13	ATO2	SOD2	Putative transmembrane protein involved in export of ammonia, a starvation signal that promotes cell death in aging colonies; phosphorylated in mitochondria; member of the TC 9.B.33 YaaH family; homolog of Ady2p and Y. lipolytica Gpr1p	Mitochondrial superoxide dismutase, protects cells against oxygen toxicity; phosphorylated
73	13	1774824_at	FET3	0	Ferro-O2-oxidoreductase required for high-affinity iron uptake and involved in mediating resistance to copper ion toxicity, belongs to class of integral membrane multicopper oxidases
74	13	MRH1	PDR12	Protein that localizes primarily to the plasma membrane, also found at the nuclear envelope; the authentic, non-tagged protein is detected in mitochondria in a phosphorylated state; has similarity to Hsp30p and Yro2p	Plasma membrane ATP-binding cassette (ABC) transporter, weak-acid-inducible multidrug transporter required for weak organic acid resistance; induced by sorbate and benzoate and regulated by War1p; mutants exhibit sorbate hypersensitivity
75	13	PDR10	PIS1	ATP-binding cassette (ABC) transporter, multidrug transporter involved in the pleiotropic drug resistance network; regulated by Pdr1p and Pdr3p	Phosphatidylinositol synthase, required for biosynthesis of phosphatidylinositol, which is a precursor for polyphosphoinositides, sphingolipids, and glycolipid anchors for some of the plasma membrane proteins
76	13	FIT2	GRE2	Mannoprotein that is incorporated into the cell wall via a glycosylphosphatidylinositol (GPI) anchor, involved in the retention of siderophore-iron in the cell wall	3-methylbutanal reductase and NADPH-dependent methylglyoxal reductase (D-lactaldehyde dehydrogenase); stress induced (osmotic, ionic, oxidative, heat shock and heavy metals); regulated by the HOG pathway
77	13	ARN1	MCH2	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to ferrirubin, ferrirhodin, and related siderophores	Protein with similarity to mammalian monocarboxylate permeases, which are involved in transport of monocarboxylic acids across the plasma membrane; mutant is not deficient in monocarboxylate transport
78	13	EGD1	HXK2	Subunit beta1 of the nascent polypeptide-associated complex (NAC) involved in protein targeting, associated with cytoplasmic ribosomes; enhances DNA binding of the Gal4p activator; homolog of human BTF3b	Hexokinase isoenzyme 2 that catalyzes phosphorylation of glucose in the cytosol; predominant hexokinase during growth on glucose; functions in the nucleus to repress expression of HXK1 and GLK1 and to induce expression of its own gene
79	12	ENB1	INO1	Endosomal ferric enterobactin transporter, expressed under conditions of iron deprivation; member of the major facilitator superfamily; expression is regulated by Rcs1p and affected by chloroquine treatment	Inositol 1-phosphate synthase, involved in synthesis of inositol phosphates and inositol-containing phospholipids; transcription is coregulated with other phospholipid biosynthetic genes by Ino2p and Ino4p, which bind the UASINO DNA element
80	12	AHA1	HXT5	Co-chaperone that binds to Hsp82p and activates its ATPase activity; similar to Hch1p; expression is regulated by stresses such as heat shock	Hexose transporter with moderate affinity for glucose, induced in the presence of non-fermentable carbon sources, induced by a decrease in growth rate, contains an extended N-terminal domain relative to other HXTs
81	12	ARN2	HXT5	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates;	Hexose transporter with moderate affinity for glucose, induced in the presence of non-fermentable carbon sources,

				responsible for uptake of iron bound to the siderophore triacetylfusarinine C	induced by a decrease in growth rate, contains an extended N-terminal domain relative to other HXTs
82	12	FET3	PET10	Ferro-O2-oxidoreductase required for high-affinity iron uptake and involved in mediating resistance to copper ion toxicity, belongs to class of integral membrane multicopper oxidases	Protein of unknown function that co-purifies with lipid particles; expression pattern suggests a role in respiratory growth; computational analysis of large-scale protein-protein interaction data suggests a role in ATP/ADP exchange
83	12	MCH2	URA2	Protein with similarity to mammalian monocarboxylate permeases, which are involved in transport of monocarboxylic acids across the plasma membrane; mutant is not deficient in monocarboxylate transport	Bifunctional carbamoylphosphate synthetase (CPSase)-aspartate transcarbamylase (ATCase), catalyzes the first two enzymatic steps in the de novo biosynthesis of pyrimidines; both activities are subject to feedback inhibition by UTP
84	12	AGP1	GLN1	Low-affinity amino acid permease with broad substrate range, involved in uptake of asparagine, glutamine, and other amino acids; expression is regulated by the SPS plasma membrane amino acid sensor system (Ssy1p-Ptr3p-Ssy5p)	Glutamine synthetase (GS), synthesizes glutamine from glutamate and ammonia; with Glt1p, forms the secondary pathway for glutamate biosynthesis from ammonia; expression regulated by nitrogen source and by amino acid limitation
85	12	ILV3	SUR7	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants
86	12	1770120_at	MF(ALPHA)1	0	Mating pheromone alpha-factor, made by alpha cells; interacts with mating type a cells to induce cell cycle arrest and other responses leading to mating; also encoded by MF(ALPHA)2, although MF(ALPHA)1 produces most alpha-factor
87	12	FIT3	PDR5	Mannoprotein that is incorporated into the cell wall via a glycosylphosphatidylinositol (GPI) anchor, involved in the retention of siderophore-iron in the cell wall	Plasma membrane ATP-binding cassette (ABC) transporter, multidrug transporter actively regulated by Pdr1p; also involved in steroid transport, cation resistance, and cellular detoxification during exponential growth
88	12	FET3	GSC2	Ferro-O2-oxidoreductase required for high-affinity iron uptake and involved in mediating resistance to copper ion toxicity, belongs to class of integral membrane multicopper oxidases	Catalytic subunit of 1,3-beta-glucan synthase, involved in formation of the inner layer of the spore wall; activity positively regulated by Rho1p and negatively by Smk1p; has similarity to an alternate catalytic subunit, Fks1p (Gsc1p)
89	12	AHA1	OLE1	Co-chaperone that binds to Hsp82p and activates its ATPase activity; similar to Hch1p; expression is regulated by stresses such as heat shock	Delta(9) fatty acid desaturase, required for monounsaturated fatty acid synthesis and for normal distribution of mitochondria
90	12	VBA3 /// VBA5	VMR1	Permease of basic amino acids in the vacuolar membrane /// Putative transporter of the Major Facilitator Superfamily (MFS); proposed role as a basic amino acid permease based on phylogeny	Protein of unknown function that may interact with ribosomes, based on co-purification experiments; member of the ATP-binding cassette (ABC) family; potential Cdc28p substrate; detected in purified mitochondria in high-throughput studies
91	11	ENB1	TKL2	Endosomal ferric enterobactin transporter, expressed under conditions of iron deprivation; member of the major facilitator superfamily; expression is regulated by Rcs1p and affected by chloroquine treatment	Transketolase, similar to Tkl1p; catalyzes conversion of xylulose-5-phosphate and ribose-5-phosphate to sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate in the pentose phosphate pathway; needed for synthesis of aromatic amino acids
92	11	FAA4	PDR12	Long chain fatty acyl-CoA synthetase, regulates protein modification during growth in the presence of ethanol, functions to incorporate palmitic acid into phospholipids and neutral lipids	Plasma membrane ATP-binding cassette (ABC) transporter, weak-acid-inducible multidrug transporter required for weak organic acid resistance; induced by sorbate and benzoate and regulated by War1p; mutants exhibit sorbate hypersensitivity
93	11	SIT1	VBA3 /// VBA5	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p	Permease of basic amino acids in the vacuolar membrane /// Putative transporter of the Major Facilitator Superfamily (MFS); proposed role as a basic amino acid permease based on phylogeny
94	11	CTS1	DED81	Endochitinase, required for cell separation after mitosis; transcriptional activation during the G1 phase of the cell cycle is mediated by transcription factor Ace2p	Cytosolic asparaginyl-tRNA synthetase, required for protein synthesis, catalyzes the specific attachment of asparagine to its cognate tRNA
95	11	AHA1	GLN1	Co-chaperone that binds to Hsp82p and activates its ATPase activity; similar to Hch1p; expression is regulated by stresses such as heat shock	Glutamine synthetase (GS), synthesizes glutamine from glutamate and ammonia; with Glt1p, forms the secondary pathway for glutamate biosynthesis from ammonia; expression regulated by nitrogen source and by amino acid limitation
96	11	DFG5	PDR5	Putative mannosidase, essential glycosylphosphatidylinositol (GPI)-anchored membrane protein required for cell wall biogenesis in bud formation, involved in filamentous growth, homologous to Dcw1p	Plasma membrane ATP-binding cassette (ABC) transporter, multidrug transporter actively regulated by Pdr1p; also involved in steroid transport, cation resistance, and cellular detoxification during exponential growth
97	11	CHS1	GDE1	Chitin synthase I, requires activation from zymogenic form in order to catalyze the transfer of N-acetylglucosamine (GlcNAc) to chitin; required for	Glycerophosphocholine (GroPCho) phosphodiesterase; hydrolyzes GroPCho to choline and glycerolphosphate, for use as a phosphate source and as a precursor for phosphocholine synthesis; may interact with ribosomes

				repairing the chitin septum during cytokinesis; transcription activated by mating factor	
98	11	ATO2	URA2	Putative transmembrane protein involved in export of ammonia, a starvation signal that promotes cell death in aging colonies; phosphorylated in mitochondria; member of the TC 9.B.33 YaaH family; homolog of Ady2p and Y. lipolytica Gpr1p	Bifunctional carbamoylphosphate synthetase (CPSase)-aspartate transcarbamylase (ATCase), catalyzes the first two enzymatic steps in the de novo biosynthesis of pyrimidines; both activities are subject to feedback inhibition by UTP
99	11	ENB1	SIT1	Endosomal ferric enterobactin transporter, expressed under conditions of iron deprivation; member of the major facilitator superfamily; expression is regulated by Rcs1p and affected by chloroquine treatment	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p
100	11	GDE1	GSC2	Glycerophosphocholine (GroPCho) phosphodiesterase; hydrolyzes GroPCho to choline and glycerolphosphate, for use as a phosphate source and as a precursor for phosphocholine synthesis; may interact with ribosomes	Catalytic subunit of 1,3-beta-glucan synthase, involved in formation of the inner layer of the spore wall; activity positively regulated by Rho1p and negatively by Smk1p; has similarity to an alternate catalytic subunit, Fks1p (Gsc1p)
101	11	HTB1	URA1	Histone H2B, core histone protein required for chromatin assembly and chromosome function; nearly identical to HTB2; Rad6p-Bre1p-Lge1p mediated ubiquitination regulates transcriptional activation, meiotic DSB formation and H3 methylation	Dihydroorotate dehydrogenase, catalyzes the fourth enzymatic step in the de novo biosynthesis of pyrimidines, converting dihydroorotic acid into orotic acid
102	10	SAH1	SUR7	S-adenosyl-L-homocysteine hydrolase, catabolizes S-adenosyl-L-homocysteine which is formed after donation of the activated methyl group of S-adenosyl-L-methionine (AdoMet) to an acceptor	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants
103	10	ENB1	HXT5	Endosomal ferric enterobactin transporter, expressed under conditions of iron deprivation; member of the major facilitator superfamily; expression is regulated by Rcs1p and affected by chloroquine treatment	Hexose transporter with moderate affinity for glucose, induced in the presence of non-fermentable carbon sources, induced by a decrease in growth rate, contains an extended N-terminal domain relative to other HXTs
104	10	HEM13	PDR5	Coproporphyrinogen III oxidase, an oxygen requiring enzyme that catalyzes the sixth step in the heme biosynthetic pathway; localizes to the mitochondrial inner membrane; transcription is repressed by oxygen and heme (via Rox1p and Hap1p)	Plasma membrane ATP-binding cassette (ABC) transporter, multidrug transporter actively regulated by Pdr1p; also involved in steroid transport, cation resistance, and cellular detoxification during exponential growth
105	10	GRE2	PDR12	3-methylbutanal reductase and NADPH-dependent methylglyoxal reductase (D-lactaldehyde dehydrogenase); stress induced (osmotic, ionic, oxidative, heat shock and heavy metals); regulated by the HOG pathway	Plasma membrane ATP-binding cassette (ABC) transporter, weak-acid-inducible multidrug transporter required for weak organic acid resistance; induced by sorbate and benzoate and regulated by War1p; mutants exhibit sorbate hypersensitivity
106	10	GDE1	PDR12	Glycerophosphocholine (GroPCho) phosphodiesterase; hydrolyzes GroPCho to choline and glycerolphosphate, for use as a phosphate source and as a precursor for phosphocholine synthesis; may interact with ribosomes	Plasma membrane ATP-binding cassette (ABC) transporter, weak-acid-inducible multidrug transporter required for weak organic acid resistance; induced by sorbate and benzoate and regulated by War1p; mutants exhibit sorbate hypersensitivity
107	10	CHA1	SUR7	Catabolic L-serine (L-threonine) deaminase, catalyzes the degradation of both L-serine and L-threonine; required to use serine or threonine as the sole nitrogen source, transcriptionally induced by serine and threonine	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants
108	10	FIT3	ZRT3	Mannoprotein that is incorporated into the cell wall via a glycosylphosphatidylinositol (GPI) anchor, involved in the retention of siderophore-iron in the cell wall	Vacuolar membrane zinc transporter, transports zinc from storage in the vacuole to the cytoplasm when needed; transcription is induced under conditions of zinc deficiency
109	10	FET3	PCA1	Ferro-O2-oxidoreductase required for high-affinity iron uptake and involved in mediating resistance to copper ion toxicity, belongs to class of integral membrane multicopper oxidases	Cadmium transporting P-type ATPase; may also have a role in copper and iron homeostasis; stabilized by Cd binding, which prevents ubiquitination; S288C and other lab strains contain a G970R mutation which eliminates Cd transport function
110	10	CHA1	RNR4	Catabolic L-serine (L-threonine) deaminase, catalyzes the degradation of both L-serine and L-threonine; required to use serine or threonine as the sole nitrogen source, transcriptionally induced by serine and threonine	Ribonucleotide-diphosphate reductase (RNR), small subunit; the RNR complex catalyzes the rate-limiting step in dNTP synthesis and is regulated by DNA replication and DNA damage checkpoint pathways via localization of the small subunits
111	10	SUR7	ZPS1	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH
112	9	DFG5	YKE4	Putative mannosidase, essential glycosylphosphatidylinositol (GPI)-anchored membrane protein required for cell wall biogenesis in bud formation, involved in filamentous growth, homologous to Dcw1p	Zinc transporter; localizes to the ER; null mutant is sensitive to calcofluor white, leads to zinc accumulation in cytosol; ortholog of the mouse KE4 and member of the ZIP (ZRT, IRT-like Protein) family
113	9	ATO2	PDR10	Putative transmembrane protein involved in export of ammonia, a starvation signal that promotes cell death in aging colonies; phosphorylated in mitochondria;	ATP-binding cassette (ABC) transporter, multidrug transporter involved in the pleiotropic drug resistance network; regulated by Pdr1p and Pdr3p

				member of the TC 9.B.33 YaaH family; homolog of Ady2p and Y. lipolytica Gpr1p	
114	9	GDE1	MRH1	Glycerophosphocholine (GroPCho) phosphodiesterase; hydrolyzes GroPCho to choline and glycerolphosphate, for use as a phosphate source and as a precursor for phosphocholine synthesis; may interact with ribosomes	Protein that localizes primarily to the plasma membrane, also found at the nuclear envelope; the authentic, non-tagged protein is detected in mitochondria in a phosphorylated state; has similarity to Hsp30p and Yro2p
115	9	CHA1	MF(ALPHA)1	Catabolic L-serine (L-threonine) deaminase, catalyzes the degradation of both L-serine and L-threonine; required to use serine or threonine as the sole nitrogen source, transcriptionally induced by serine and threonine	Mating pheromone alpha-factor, made by alpha cells; interacts with mating type a cells to induce cell cycle arrest and other responses leading to mating; also encoded by MF(ALPHA)2, although MF(ALPHA)1 produces most alpha-factor
116	9	MF(ALPHA)1	RNR4	Mating pheromone alpha-factor, made by alpha cells; interacts with mating type a cells to induce cell cycle arrest and other responses leading to mating; also encoded by MF(ALPHA)2, although MF(ALPHA)1 produces most alpha-factor	Ribonucleotide-diphosphate reductase (RNR), small subunit; the RNR complex catalyzes the rate-limiting step in dNTP synthesis and is regulated by DNA replication and DNA damage checkpoint pathways via localization of the small subunits
117	9	ILV3	TRP5	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids	Tryptophan synthase involved in tryptophan biosynthesis, regulated by the general control system of amino acid biosynthesis
118	9	SUR7	TKL1	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants	Transketolase, similar to Tkl2p; catalyzes conversion of xylulose-5-phosphate and ribose-5-phosphate to sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate in the pentose phosphate pathway; needed for synthesis of aromatic amino acids
119	9	1774824_at	MCH2	0	Protein with similarity to mammalian monocarboxylate permeases, which are involved in transport of monocarboxylic acids across the plasma membrane; mutant is not deficient in monocarboxylate transport
120	9	SSB2	TKL1	Cytoplasmic ATPase that is a ribosome-associated molecular chaperone, functions with J-protein partner Zuo1p; may be involved in the folding of newly-synthesized polypeptide chains; member of the HSP70 family; homolog of SSB1	Transketolase, similar to Tkl2p; catalyzes conversion of xylulose-5-phosphate and ribose-5-phosphate to sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate in the pentose phosphate pathway; needed for synthesis of aromatic amino acids
121	9	FET3	GDE1	Ferro-O2-oxidoreductase required for high-affinity iron uptake and involved in mediating resistance to copper ion toxicity, belongs to class of integral membrane multicopper oxidases	Glycerophosphocholine (GroPCho) phosphodiesterase; hydrolyzes GroPCho to choline and glycerolphosphate, for use as a phosphate source and as a precursor for phosphocholine synthesis; may interact with ribosomes
122	9	TKL1	TRP5	Transketolase, similar to Tkl2p; catalyzes conversion of xylulose-5-phosphate and ribose-5-phosphate to sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate in the pentose phosphate pathway; needed for synthesis of aromatic amino acids	Tryptophan synthase involved in tryptophan biosynthesis, regulated by the general control system of amino acid biosynthesis
123	9	ENB1	MCH2	Endosomal ferric enterobactin transporter, expressed under conditions of iron deprivation; member of the major facilitator superfamily; expression is regulated by Rcs1p and affected by chloroquine treatment	Protein with similarity to mammalian monocarboxylate permeases, which are involved in transport of monocarboxylic acids across the plasma membrane; mutant is not deficient in monocarboxylate transport
124	9	1778446_at	AMS1	0	Vacuolar alpha mannosidase, involved in free oligosaccharide (fOS) degradation; delivered to the vacuole in a novel pathway separate from the secretory pathway
125	9	BAP2	ENB1	High-affinity leucine permease, functions as a branched-chain amino acid permease involved in the uptake of leucine, isoleucine and valine; contains 12 predicted transmembrane domains	Endosomal ferric enterobactin transporter, expressed under conditions of iron deprivation; member of the major facilitator superfamily; expression is regulated by Rcs1p and affected by chloroquine treatment
126	8	1774824_at	VMR1	0	Protein of unknown function that may interact with ribosomes, based on co-purification experiments; member of the ATP-binding cassette (ABC) family; potential Cdc28p substrate; detected in purified mitochondria in high-throughput studies
127	8	ILV5	TRP5	Acetohydroxyacid reductoisomerase, mitochondrial protein involved in branched-chain amino acid biosynthesis, also required for maintenance of wild-type mitochondrial DNA and found in mitochondrial nucleoids	Tryptophan synthase involved in tryptophan biosynthesis, regulated by the general control system of amino acid biosynthesis
128	8	FIT3	YKE4	Mannoprotein that is incorporated into the cell wall via a glycosylphosphatidylinositol (GPI) anchor, involved in the retention of siderophore-iron in the cell wall	Zinc transporter; localizes to the ER; null mutant is sensitive to calcofluor white, leads to zinc accumulation in cytosol; ortholog of the mouse KE4 and member of the ZIP (ZRT, IRT-like Protein) family
129	8	1770120_at	PDR10	0	ATP-binding cassette (ABC) transporter, multidrug transporter involved in the pleiotropic drug resistance network; regulated by Pdr1p and Pdr3p
130	8	SCS7	YRO2	Sphingolipid alpha-hydroxylase, functions in the alpha-hydroxylation of sphingolipid-associated very long chain fatty acids, has both cytochrome b5-like and	Putative protein of unknown function; the authentic, non-tagged protein is detected in a phosphorylated state in highly purified mitochondria in high-throughput studies; transcriptionally regulated by Haa1p

				hydroxylase/desaturase domains, not essential for growth	
131	8	HXT5	URA2	Hexose transporter with moderate affinity for glucose, induced in the presence of non-fermentable carbon sources, induced by a decrease in growth rate, contains an extended N-terminal domain relative to other HXTs	Bifunctional carbamoylphosphate synthetase (CPSase)-aspartate transcarbamylase (ATCase), catalyzes the first two enzymatic steps in the de novo biosynthesis of pyrimidines; both activities are subject to feedback inhibition by UTP
132	8	CHA1	ENB1	Catabolic L-serine (L-threonine) deaminase, catalyzes the degradation of both L-serine and L-threonine; required to use serine or threonine as the sole nitrogen source, transcriptionally induced by serine and threonine	Endosomal ferric enterobactin transporter, expressed under conditions of iron deprivation; member of the major facilitator superfamily; expression is regulated by Rcs1p and affected by chloroquine treatment
133	8	GDE1	SAG1	Glycerophosphocholine (GroPCho) phosphodiesterase; hydrolyzes GroPCho to choline and glycerolphosphate, for use as a phosphate source and as a precursor for phosphocholine synthesis; may interact with ribosomes	Alpha-agglutinin of alpha-cells, binds to Aga1p during agglutination, N-terminal half is homologous to the immunoglobulin superfamily and contains binding site for a-agglutinin, C-terminal half is highly glycosylated and contains GPI anchor
134	8	CHS1	FET3	Chitin synthase I, requires activation from zymogenic form in order to catalyze the transfer of N-acetylglucosamine (GlcNAc) to chitin; required for repairing the chitin septum during cytokinesis; transcription activated by mating factor	Ferro-O2-oxidoreductase required for high-affinity iron uptake and involved in mediating resistance to copper ion toxicity, belongs to class of integral membrane multicopper oxidases
135	7	DED81	WTM1	Cytosolic asparaginyl-tRNA synthetase, required for protein synthesis, catalyzes the specific attachment of asparagine to its cognate tRNA	Transcriptional modulator involved in regulation of meiosis, silencing, and expression of RNR genes; required for nuclear localization of the ribonucleotide reductase small subunit Rnr2p and Rnr4p; contains WD repeats
136	7	SIT1	TRP5	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p	Tryptophan synthase involved in tryptophan biosynthesis, regulated by the general control system of amino acid biosynthesis
137	7	AGP1	CIT2	Low-affinity amino acid permease with broad substrate range, involved in uptake of asparagine, glutamine, and other amino acids; expression is regulated by the SPS plasma membrane amino acid sensor system (Ssy1p-Ptr3p-Ssy5p)	Citrate synthase, catalyzes the condensation of acetyl coenzyme A and oxaloacetate to form citrate, peroxisomal isozyme involved in glyoxylate cycle; expression is controlled by Rtg1p and Rtg2p transcription factors
138	7	AHA1	FAA4	Co-chaperone that binds to Hsp82p and activates its ATPase activity; similar to Hch1p; expression is regulated by stresses such as heat shock	Long chain fatty acyl-CoA synthetase, regulates protein modification during growth in the presence of ethanol, functions to incorporate palmitic acid into phospholipids and neutral lipids
139	7	FIT3	PDR12	Mannoprotein that is incorporated into the cell wall via a glycosylphosphatidylinositol (GPI) anchor, involved in the retention of siderophore-iron in the cell wall	Plasma membrane ATP-binding cassette (ABC) transporter, weak-acid-inducible multidrug transporter required for weak organic acid resistance; induced by sorbate and benzoate and regulated by War1p; mutants exhibit sorbate hypersensitivity
140	7	HXT5	TRP5	Hexose transporter with moderate affinity for glucose, induced in the presence of non-fermentable carbon sources, induced by a decrease in growth rate, contains an extended N-terminal domain relative to other HXTs	Tryptophan synthase involved in tryptophan biosynthesis, regulated by the general control system of amino acid biosynthesis
141	7	BAP2	CHA1	High-affinity leucine permease, functions as a branched-chain amino acid permease involved in the uptake of leucine, isoleucine and valine; contains 12 predicted transmembrane domains	Catabolic L-serine (L-threonine) deaminase, catalyzes the degradation of both L-serine and L-threonine; required to use serine or threonine as the sole nitrogen source, transcriptionally induced by serine and threonine
142	7	1774824_at	ARN2	0	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to the siderophore triacetylfusarinine C
143	7	GDE1	YKE4	Glycerophosphocholine (GroPCho) phosphodiesterase; hydrolyzes GroPCho to choline and glycerolphosphate, for use as a phosphate source and as a precursor for phosphocholine synthesis; may interact with ribosomes	Zinc transporter; localizes to the ER; null mutant is sensitive to calcofluor white, leads to zinc accumulation in cytosol; ortholog of the mouse KE4 and member of the ZIP (ZRT, IRT-like Protein) family
144	7	MRH1	OLE1	Protein that localizes primarily to the plasma membrane, also found at the nuclear envelope; the authentic, non-tagged protein is detected in mitochondria in a phosphorylated state; has similarity to Hsp30p and Yro2p	Delta(9) fatty acid desaturase, required for monounsaturated fatty acid synthesis and for normal distribution of mitochondria
145	7	1773710_at	SUR7	0	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants
146	7	CHA1	PDR10	Catabolic L-serine (L-threonine) deaminase, catalyzes the degradation of both L-serine and L-threonine; required to use serine or threonine as the sole nitrogen source, transcriptionally induced by serine and threonine	ATP-binding cassette (ABC) transporter, multidrug transporter involved in the pleiotropic drug resistance network; regulated by Pdr1p and Pdr3p

147	7	1778446_at	ARN1	0	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to ferrirubin, ferrirhodin, and related siderophores
148	6	CHA1	SIT1	Catabolic L-serine (L-threonine) deaminase, catalyzes the degradation of both L-serine and L-threonine; required to use serine or threonine as the sole nitrogen source, transcriptionally induced by serine and threonine	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p
149	6	FET3	OLE1	Ferro-O2-oxidoreductase required for high-affinity iron uptake and involved in mediating resistance to copper ion toxicity, belongs to class of integral membrane multicopper oxidases	Delta(9) fatty acid desaturase, required for monounsaturated fatty acid synthesis and for normal distribution of mitochondria
150	6	AHA1	URA2	Co-chaperone that binds to Hsp82p and activates its ATPase activity; similar to Hch1p; expression is regulated by stresses such as heat shock	Bifunctional carbamoylphosphate synthetase (CPSase)-aspartate transcarbamylase (ATCase), catalyzes the first two enzymatic steps in the de novo biosynthesis of pyrimidines; both activities are subject to feedback inhibition by UTP
151	6	SIT1	TIS11	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p	mRNA-binding protein expressed during iron starvation; binds to a sequence element in the 3'-untranslated regions of specific mRNAs to mediate their degradation; involved in iron homeostasis
152	6	DED81	TRP5	Cytosolic asparaginyl-tRNA synthetase, required for protein synthesis, catalyzes the specific attachment of asparagine to its cognate tRNA	Tryptophan synthase involved in tryptophan biosynthesis, regulated by the general control system of amino acid biosynthesis
153	6	GDE1	SUR7	Glycerophosphocholine (GroPCho) phosphodiesterase; hydrolyzes GroPCho to choline and glycerolphosphate, for use as a phosphate source and as a precursor for phosphocholine synthesis; may interact with ribosomes	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants
154	6	AHA1	YRO2	Co-chaperone that binds to Hsp82p and activates its ATPase activity; similar to Hch1p; expression is regulated by stresses such as heat shock	Putative protein of unknown function; the authentic, non-tagged protein is detected in a phosphorylated state in highly purified mitochondria in high-throughput studies; transcriptionally regulated by Haa1p
155	6	URA1	YPS1	Dihydroorotate dehydrogenase, catalyzes the fourth enzymatic step in the de novo biosynthesis of pyrimidines, converting dihydroorotic acid into orotic acid	Aspartic protease, attached to the plasma membrane via a glycosylphosphatidylinositol (GPI) anchor
156	5	DED81	HEM13	Cytosolic asparaginyl-tRNA synthetase, required for protein synthesis, catalyzes the specific attachment of asparagine to its cognate tRNA	Coproporphyrinogen III oxidase, an oxygen requiring enzyme that catalyzes the sixth step in the heme biosynthetic pathway; localizes to the mitochondrial inner membrane; transcription is repressed by oxygen and heme (via Rox1p and Hap1p)
157	5	1773710_at	DFG5	0	Putative mannosidase, essential glycosylphosphatidylinositol (GPI)-anchored membrane protein required for cell wall biogenesis in bud formation, involved in filamentous growth, homologous to Dcw1p
158	5	1773710_at	WTM1	0	Transcriptional modulator involved in regulation of meiosis, silencing, and expression of RNR genes; required for nuclear localization of the ribonucleotide reductase small subunit Rnr2p and Rnr4p; contains WD repeats
159	5	ARN2	URA2	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to the siderophore triacetylfusarinine C	Bifunctional carbamoylphosphate synthetase (CPSase)-aspartate transcarbamylase (ATCase), catalyzes the first two enzymatic steps in the de novo biosynthesis of pyrimidines; both activities are subject to feedback inhibition by UTP
160	5	CHS1	GSC2	Chitin synthase I, requires activation from zymogenic form in order to catalyze the transfer of N-acetylglucosamine (GlcNAc) to chitin; required for repairing the chitin septum during cytokinesis; transcription activated by mating factor	Catalytic subunit of 1,3-beta-glucan synthase, involved in formation of the inner layer of the spore wall; activity positively regulated by Rho1p and negatively by Smk1p; has similarity to an alternate catalytic subunit, Fks1p (Gsc1p)
161	5	1773710_at	CHS1	0	Chitin synthase I, requires activation from zymogenic form in order to catalyze the transfer of N-acetylglucosamine (GlcNAc) to chitin; required for repairing the chitin septum during cytokinesis; transcription activated by mating factor
162	5	DED81	RNR2	Cytosolic asparaginyl-tRNA synthetase, required for protein synthesis, catalyzes the specific attachment of asparagine to its cognate tRNA	Ribonucleotide-diphosphate reductase (RNR), small subunit; the RNR complex catalyzes the rate-limiting step in dNTP synthesis and is regulated by DNA replication and DNA damage checkpoint pathways via localization of the small subunits
163	5	AHA1	ZPS1	Co-chaperone that binds to Hsp82p and activates its ATPase activity; similar to Hch1p; expression is regulated by stresses such as heat shock	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH
164	5	PDR12	YRO2	Plasma membrane ATP-binding cassette (ABC) transporter, weak-acid-inducible multidrug transporter required for weak organic acid resistance; induced by	Putative protein of unknown function; the authentic, non-tagged protein is detected in a phosphorylated state in

				sorbate and benzoate and regulated by War1p; mutants exhibit sorbate hypersensitivity	highly purified mitochondria in high-throughput studies; transcriptionally regulated by Haa1p
165	5	RNR2	TKL1	Ribonucleotide-diphosphate reductase (RNR), small subunit; the RNR complex catalyzes the rate-limiting step in dNTP synthesis and is regulated by DNA replication and DNA damage checkpoint pathways via localization of the small subunits	Transketolase, similar to Tkl2p; catalyzes conversion of xylulose-5-phosphate and ribose-5-phosphate to sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate in the pentose phosphate pathway; needed for synthesis of aromatic amino acids
166	5	1773710_at	TRP5	0	Tryptophan synthase involved in tryptophan biosynthesis, regulated by the general control system of amino acid biosynthesis
167	5	OLE1	PDR12	Delta(9) fatty acid desaturase, required for monounsaturated fatty acid synthesis and for normal distribution of mitochondria	Plasma membrane ATP-binding cassette (ABC) transporter, weak-acid-inducible multidrug transporter required for weak organic acid resistance; induced by sorbate and benzoate and regulated by War1p; mutants exhibit sorbate hypersensitivity
168	5	OLE1	SUR7	Delta(9) fatty acid desaturase, required for monounsaturated fatty acid synthesis and for normal distribution of mitochondria	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants
169	5	AHA1	PDR12	Co-chaperone that binds to Hsp82p and activates its ATPase activity; similar to Hch1p; expression is regulated by stresses such as heat shock	Plasma membrane ATP-binding cassette (ABC) transporter, weak-acid-inducible multidrug transporter required for weak organic acid resistance; induced by sorbate and benzoate and regulated by War1p; mutants exhibit sorbate hypersensitivity
170	5	SCS7	URA2	Sphingolipid alpha-hydroxylase, functions in the alpha-hydroxylation of sphingolipid-associated very long chain fatty acids, has both cytochrome b5-like and hydroxylase/desaturase domains, not essential for growth	Bifunctional carbamoylphosphate synthetase (CPSase)-aspartate transcarbamylase (ATCase), catalyzes the first two enzymatic steps in the de novo biosynthesis of pyrimidines; both activities are subject to feedback inhibition by UTP
171	5	EGD1	MF(ALPHA)1	Subunit beta1 of the nascent polypeptide-associated complex (NAC) involved in protein targeting, associated with cytoplasmic ribosomes; enhances DNA binding of the Gal4p activator; homolog of human BTF3b	Mating pheromone alpha-factor, made by alpha cells; interacts with mating type a cells to induce cell cycle arrest and other responses leading to mating; also encoded by MF(ALPHA)2, although MF(ALPHA)1 produces most alpha-factor
172	5	DAN1	DED81	Cell wall mannoprotein with similarity to Tir1p, Tir2p, Tir3p, and Tir4p; expressed under anaerobic conditions, completely repressed during aerobic growth	Cytosolic asparaginyl-tRNA synthetase, required for protein synthesis, catalyzes the specific attachment of asparagine to its cognate tRNA
173	5	MDH2	OLE1	Cytoplasmic malate dehydrogenase, one of three isozymes that catalyze interconversion of malate and oxaloacetate; involved in the glyoxylate cycle and gluconeogenesis during growth on two-carbon compounds; interacts with Pck1p and Fbp1	Delta(9) fatty acid desaturase, required for monounsaturated fatty acid synthesis and for normal distribution of mitochondria
174	4	FIT3	GDE1	Mannoprotein that is incorporated into the cell wall via a glycosylphosphatidylinositol (GPI) anchor, involved in the retention of siderophore-iron in the cell wall	Glycerophosphocholine (GroPCho) phosphodiesterase; hydrolyzes GroPCho to choline and glycerolphosphate, for use as a phosphate source and as a precursor for phosphocholine synthesis; may interact with ribosomes
175	4	ARN2	ENB1	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to the siderophore triacetylfusarinine C	Endosomal ferric enterobactin transporter, expressed under conditions of iron deprivation; member of the major facilitator superfamily; expression is regulated by Rcs1p and affected by chloroquine treatment
176	4	AMS1	MCH2	Vacuolar alpha mannosidase, involved in free oligosaccharide (fOS) degradation; delivered to the vacuole in a novel pathway separate from the secretory pathway	Protein with similarity to mammalian monocarboxylate permeases, which are involved in transport of monocarboxylic acids across the plasma membrane; mutant is not deficient in monocarboxylate transport
177	4	FET3	VBA3 /// VBA5	Ferro-O2-oxidoreductase required for high-affinity iron uptake and involved in mediating resistance to copper ion toxicity, belongs to class of integral membrane multicopper oxidases	Permease of basic amino acids in the vacuolar membrane /// Putative transporter of the Major Facilitator Superfamily (MFS); proposed role as a basic amino acid permease based on phylogeny
178	4	FET3	MRH1	Ferro-O2-oxidoreductase required for high-affinity iron uptake and involved in mediating resistance to copper ion toxicity, belongs to class of integral membrane multicopper oxidases	Protein that localizes primarily to the plasma membrane, also found at the nuclear envelope; the authentic, non-tagged protein is detected in mitochondria in a phosphorylated state; has similarity to Hsp30p and Yro2p
179	4	AMS1	ARN1	Vacuolar alpha mannosidase, involved in free oligosaccharide (fOS) degradation; delivered to the vacuole in a novel pathway separate from the secretory pathway	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to ferrirubin, ferrirhodin, and related siderophores
180	4	1773710_at	GDE1	0	Glycerophosphocholine (GroPCho) phosphodiesterase; hydrolyzes GroPCho to choline and glycerolphosphate, for use as a phosphate source and as a precursor for phosphocholine synthesis; may interact with ribosomes
181	4	GDE1	HEM13	Glycerophosphocholine (GroPCho) phosphodiesterase; hydrolyzes GroPCho to choline and glycerolphosphate,	Coproporphyrinogen III oxidase, an oxygen requiring enzyme that catalyzes the sixth step in the heme biosynthetic pathway; localizes to the mitochondrial inner

				for use as a phosphate source and as a precursor for phosphocholine synthesis; may interact with ribosomes	membrane; transcription is repressed by oxygen and heme (via Rox1p and Hap1p)
182	4	ARN1	HTB1	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to ferrirubin, ferrirhodin, and related siderophores	Histone H2B, core histone protein required for chromatin assembly and chromosome function; nearly identical to HTB2; Rad6p-Bre1p-Lge1p mediated ubiquitination regulates transcriptional activation, meiotic DSB formation and H3 methylation
183	4	TRP5	VBA3 /// VBA5	Tryptophan synthase involved in tryptophan biosynthesis, regulated by the general control system of amino acid biosynthesis	Permease of basic amino acids in the vacuolar membrane /// Putative transporter of the Major Facilitator Superfamily (MFS); proposed role as a basic amino acid permease based on phylogeny
184	4	1778446_at	ICY1	0	Protein of unknown function, required for viability in rich media of cells lacking mitochondrial DNA; mutants have an invasive growth defect with elongated morphology; induced by amino acid starvation
185	4	HEM13	VBA3 /// VBA5	Coproporphyrinogen III oxidase, an oxygen requiring enzyme that catalyzes the sixth step in the heme biosynthetic pathway; localizes to the mitochondrial inner membrane; transcription is repressed by oxygen and heme (via Rox1p and Hap1p)	Permease of basic amino acids in the vacuolar membrane /// Putative transporter of the Major Facilitator Superfamily (MFS); proposed role as a basic amino acid permease based on phylogeny
186	4	HXT5	VMR1	Hexose transporter with moderate affinity for glucose, induced in the presence of non-fermentable carbon sources, induced by a decrease in growth rate, contains an extended N-terminal domain relative to other HXTs	Protein of unknown function that may interact with ribosomes, based on co-purification experiments; member of the ATP-binding cassette (ABC) family; potential Cdc28p substrate; detected in purified mitochondria in high-throughput studies
187	4	1774824_at	VBA3 /// VBA5	0	Permease of basic amino acids in the vacuolar membrane /// Putative transporter of the Major Facilitator Superfamily (MFS); proposed role as a basic amino acid permease based on phylogeny
188	4	GDE1	ZPS1	Glycerophosphocholine (GroPCho) phosphodiesterase; hydrolyzes GroPCho to choline and glycerolphosphate, for use as a phosphate source and as a precursor for phosphocholine synthesis; may interact with ribosomes	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH
189	4	MRH1	SAG1	Protein that localizes primarily to the plasma membrane, also found at the nuclear envelope; the authentic, non-tagged protein is detected in mitochondria in a phosphorylated state; has similarity to Hsp30p and Yro2p	Alpha-agglutinin of alpha-cells, binds to Aga1p during agglutination, N-terminal half is homologous to the immunoglobulin superfamily and contains binding site for a-agglutinin, C-terminal half is highly glycosylated and contains GPI anchor
190	4	FET3	SUR7	Ferro-O2-oxidoreductase required for high-affinity iron uptake and involved in mediating resistance to copper ion toxicity, belongs to class of integral membrane multicopper oxidases	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants
191	3	GSC2	MRH1	Catalytic subunit of 1,3-beta-glucan synthase, involved in formation of the inner layer of the spore wall; activity positively regulated by Rho1p and negatively by Smk1p; has similarity to an alternate catalytic subunit, Fks1p (Gsc1p)	Protein that localizes primarily to the plasma membrane, also found at the nuclear envelope; the authentic, non-tagged protein is detected in mitochondria in a phosphorylated state; has similarity to Hsp30p and Yro2p
192	3	TIS11	VMR1	mRNA-binding protein expressed during iron starvation; binds to a sequence element in the 3'-untranslated regions of specific mRNAs to mediate their degradation; involved in iron homeostasis	Protein of unknown function that may interact with ribosomes, based on co-purification experiments; member of the ATP-binding cassette (ABC) family; potential Cdc28p substrate; detected in purified mitochondria in high-throughput studies
193	3	HEM13	ILV3	Coproporphyrinogen III oxidase, an oxygen requiring enzyme that catalyzes the sixth step in the heme biosynthetic pathway; localizes to the mitochondrial inner membrane; transcription is repressed by oxygen and heme (via Rox1p and Hap1p)	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids
194	3	INO1	SIT1	Inositol 1-phosphate synthase, involved in synthesis of inositol phosphates and inositol-containing phospholipids; transcription is coregulated with other phospholipid biosynthetic genes by Ino2p and Ino4p, which bind the UASINO DNA element	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p
195	3	PCA1	WTM1	Cadmium transporting P-type ATPase; may also have a role in copper and iron homeostasis; stabilized by Cd binding, which prevents ubiquitination; S288C and other lab strains contain a G970R mutation which eliminates Cd transport function	Transcriptional modulator involved in regulation of meiosis, silencing, and expression of RNR genes; required for nuclear localization of the ribonucleotide reductase small subunit Rnr2p and Rnr4p; contains WD repeats
196	3	GDE1	SIT1	Glycerophosphocholine (GroPCho) phosphodiesterase; hydrolyzes GroPCho to choline and glycerolphosphate, for use as a phosphate source and as a precursor for phosphocholine synthesis; may interact with ribosomes	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p
197	3	ILV3	ZPS1	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH

198	3	GSC2	PET10	Catalytic subunit of 1,3-beta-glucan synthase, involved in formation of the inner layer of the spore wall; activity positively regulated by Rho1p and negatively by Smk1p; has similarity to an alternate catalytic subunit, Fks1p (Gsc1p)	Protein of unknown function that co-purifies with lipid particles; expression pattern suggests a role in respiratory growth; computational analysis of large-scale protein-protein interaction data suggests a role in ATP/ADP exchange
199	3	PET10	SAG1	Protein of unknown function that co-purifies with lipid particles; expression pattern suggests a role in respiratory growth; computational analysis of large-scale protein-protein interaction data suggests a role in ATP/ADP exchange	Alpha-agglutinin of alpha-cells, binds to Aga1p during agglutination, N-terminal half is homologous to the immunoglobulin superfamily and contains binding site for a-agglutinin, C-terminal half is highly glycosylated and contains GPI anchor
200	3	FET3	ZPS1	Ferro-O2-oxidoreductase required for high-affinity iron uptake and involved in mediating resistance to copper ion toxicity, belongs to class of integral membrane multicopper oxidases	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH
201	3	ERG1	OLE1	Squalene epoxidase, catalyzes the epoxidation of squalene to 2,3-oxidosqualene; plays an essential role in the ergosterol-biosynthesis pathway and is the specific target of the antifungal drug terbinafine	Delta(9) fatty acid desaturase, required for monounsaturated fatty acid synthesis and for normal distribution of mitochondria
202	3	MF(ALPHA)1	WTM1	Mating pheromone alpha-factor, made by alpha cells; interacts with mating type a cells to induce cell cycle arrest and other responses leading to mating; also encoded by MF(ALPHA)2, although MF(ALPHA)1 produces most alpha-factor	Transcriptional modulator involved in regulation of meiosis, silencing, and expression of RNR genes; required for nuclear localization of the ribonucleotide reductase small subunit Rnr2p and Rnr4p; contains WD repeats
203	3	OLE1	URA2	Delta(9) fatty acid desaturase, required for monounsaturated fatty acid synthesis and for normal distribution of mitochondria	Bifunctional carbamoylphosphate synthetase (CPSase)-aspartate transcarbamylase (ATCase), catalyzes the first two enzymatic steps in the de novo biosynthesis of pyrimidines; both activities are subject to feedback inhibition by UTP
204	3	AGX1	DFG5	Alanine:glyoxylate aminotransferase (AGT), catalyzes the synthesis of glycine from glyoxylate, which is one of three pathways for glycine biosynthesis in yeast; has similarity to mammalian and plant alanine:glyoxylate aminotransferases	Putative mannosidase, essential glycosylphosphatidylinositol (GPI)-anchored membrane protein required for cell wall biogenesis in bud formation, involved in filamentous growth, homologous to Dcw1p
205	3	MRH1	YRO2	Protein that localizes primarily to the plasma membrane, also found at the nuclear envelope; the authentic, non-tagged protein is detected in mitochondria in a phosphorylated state; has similarity to Hsp30p and Yro2p	Putative protein of unknown function; the authentic, non-tagged protein is detected in a phosphorylated state in highly purified mitochondria in high-throughput studies; transcriptionally regulated by Haa1p
206	3	DED81	ILV3	Cytosolic asparaginyl-tRNA synthetase, required for protein synthesis, catalyzes the specific attachment of asparagine to its cognate tRNA	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids
207	3	AGX1	HEM13	Alanine:glyoxylate aminotransferase (AGT), catalyzes the synthesis of glycine from glyoxylate, which is one of three pathways for glycine biosynthesis in yeast; has similarity to mammalian and plant alanine:glyoxylate aminotransferases	Coproporphyrinogen III oxidase, an oxygen requiring enzyme that catalyzes the sixth step in the heme biosynthetic pathway; localizes to the mitochondrial inner membrane; transcription is repressed by oxygen and heme (via Rox1p and Hap1p)
208	3	1774824_at	ARN1	0	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to ferrirubin, ferrirhodin, and related siderophores
209	3	DFG5	FIT3	Putative mannosidase, essential glycosylphosphatidylinositol (GPI)-anchored membrane protein required for cell wall biogenesis in bud formation, involved in filamentous growth, homologous to Dcw1p	Mannoprotein that is incorporated into the cell wall via a glycosylphosphatidylinositol (GPI) anchor, involved in the retention of siderophore-iron in the cell wall
210	3	OLE1	SIT1	Delta(9) fatty acid desaturase, required for monounsaturated fatty acid synthesis and for normal distribution of mitochondria	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p
211	2	OLE1	SCS7	Delta(9) fatty acid desaturase, required for monounsaturated fatty acid synthesis and for normal distribution of mitochondria	Sphingolipid alpha-hydroxylase, functions in the alpha-hydroxylation of sphingolipid-associated very long chain fatty acids, has both cytochrome b5-like and hydroxylase/desaturase domains, not essential for growth
212	2	RNR4	SUR7	Ribonucleotide-diphosphate reductase (RNR), small subunit; the RNR complex catalyzes the rate-limiting step in dNTP synthesis and is regulated by DNA replication and DNA damage checkpoint pathways via localization of the small subunits	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants
213	2	HXT5	SIT1	Hexose transporter with moderate affinity for glucose, induced in the presence of non-fermentable carbon sources, induced by a decrease in growth rate, contains an extended N-terminal domain relative to other HXTs	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p
214	2	SUR7	TRP5	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with	Tryptophan synthase involved in tryptophan biosynthesis, regulated by the general control system of amino acid biosynthesis

				Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants	
215	2	AHA1	SIT1	Co-chaperone that binds to Hsp82p and activates its ATPase activity; similar to Hch1p; expression is regulated by stresses such as heat shock	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p
216	2	GDE1	VBA3 /// VBA5	Glycerophosphocholine (GroPCho) phosphodiesterase; hydrolyzes GroPCho to choline and glycerolphosphate, for use as a phosphate source and as a precursor for phosphocholine synthesis; may interact with ribosomes	Permease of basic amino acids in the vacuolar membrane /// Putative transporter of the Major Facilitator Superfamily (MFS); proposed role as a basic amino acid permease based on phylogeny
217	2	1774824_at	SIT1	0	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p
218	2	DED81	SUR7	Cytosolic asparaginyl-tRNA synthetase, required for protein synthesis, catalyzes the specific attachment of asparagine to its cognate tRNA	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants
219	2	OLE1	YRO2	Delta(9) fatty acid desaturase, required for monounsaturated fatty acid synthesis and for normal distribution of mitochondria	Putative protein of unknown function; the authentic, non-tagged protein is detected in a phosphorylated state in highly purified mitochondria in high-throughput studies; transcriptionally regulated by Haa1p
220	2	AGA1	ILV3	Anchorage subunit of a-agglutinin of a-cells, highly O-glycosylated protein with N-terminal secretion signal and C-terminal signal for addition of GPI anchor to cell wall, linked to adhesion subunit Aga2p via two disulfide bonds	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids
221	2	TIS11	ZPS1	mRNA-binding protein expressed during iron starvation; binds to a sequence element in the 3'-untranslated regions of specific mRNAs to mediate their degradation; involved in iron homeostasis	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH
222	2	GDE1	OLE1	Glycerophosphocholine (GroPCho) phosphodiesterase; hydrolyzes GroPCho to choline and glycerolphosphate, for use as a phosphate source and as a precursor for phosphocholine synthesis; may interact with ribosomes	Delta(9) fatty acid desaturase, required for monounsaturated fatty acid synthesis and for normal distribution of mitochondria
223	2	CHS1	SIT1	Chitin synthase I, requires activation from zymogenic form in order to catalyze the transfer of N-acetylglucosamine (GlcNAc) to chitin; required for repairing the chitin septum during cytokinesis; transcription activated by mating factor	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p
224	2	GLN1	URA2	Glutamine synthetase (GS), synthesizes glutamine from glutamate and ammonia; with Glt1p, forms the secondary pathway for glutamate biosynthesis from ammonia; expression regulated by nitrogen source and by amino acid limitation	Bifunctional carbamoylphosphate synthetase (CPSase)-aspartate transcarbamylase (ATCase), catalyzes the first two enzymatic steps in the de novo biosynthesis of pyrimidines; both activities are subject to feedback inhibition by UTP
225	2	GSC2	PDR12	Catalytic subunit of 1,3-beta-glucan synthase, involved in formation of the inner layer of the spore wall; activity positively regulated by Rho1p and negatively by Smk1p; has similarity to an alternate catalytic subunit, Fks1p (Gsc1p)	Plasma membrane ATP-binding cassette (ABC) transporter, weak-acid-inducible multidrug transporter required for weak organic acid resistance; induced by sorbate and benzoate and regulated by War1p; mutants exhibit sorbate hypersensitivity
226	2	1773710_at	ILV3	0	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids
227	2	ENB1	TRP5	Endosomal ferric enterobactin transporter, expressed under conditions of iron deprivation; member of the major facilitator superfamily; expression is regulated by Rcs1p and affected by chloroquine treatment	Tryptophan synthase involved in tryptophan biosynthesis, regulated by the general control system of amino acid biosynthesis
228	2	PDR10	RNR4	ATP-binding cassette (ABC) transporter, multidrug transporter involved in the pleiotropic drug resistance network; regulated by Pdr1p and Pdr3p	Ribonucleotide-diphosphate reductase (RNR), small subunit; the RNR complex catalyzes the rate-limiting step in dNTP synthesis and is regulated by DNA replication and DNA damage checkpoint pathways via localization of the small subunits
229	2	ILV3	VMR1	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids	Protein of unknown function that may interact with ribosomes, based on co-purification experiments; member of the ATP-binding cassette (ABC) family; potential Cdc28p substrate; detected in purified mitochondria in high-throughput studies
230	2	ATO2	SCS7	Putative transmembrane protein involved in export of ammonia, a starvation signal that promotes cell death in aging colonies; phosphorylated in mitochondria; member of the TC 9.B.33 YaaH family; homolog of Ady2p and Y. lipolytica Gpr1p	Sphingolipid alpha-hydroxylase, functions in the alpha-hydroxylation of sphingolipid-associated very long chain fatty acids, has both cytochrome b5-like and hydroxylase/desaturase domains, not essential for growth
231	2	INO1	TKL1	Inositol 1-phosphate synthase, involved in synthesis of inositol phosphates and inositol-containing	Transketolase, similar to Tkl2p; catalyzes conversion of xylulose-5-phosphate and ribose-5-phosphate to

				phospholipids; transcription is coregulated with other phospholipid biosynthetic genes by Ino2p and Ino4p, which bind the UASINO DNA element	sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate in the pentose phosphate pathway; needed for synthesis of aromatic amino acids
232	2	CHA1	WTM1	Catabolic L-serine (L-threonine) deaminase, catalyzes the degradation of both L-serine and L-threonine; required to use serine or threonine as the sole nitrogen source, transcriptionally induced by serine and threonine	Transcriptional modulator involved in regulation of meiosis, silencing, and expression of RNR genes; required for nuclear localization of the ribonucleotide reductase small subunit Rnr2p and Rnr4p; contains WD repeats
233	2	DFG5	PDR12	Putative mannosidase, essential glycosylphosphatidylinositol (GPI)-anchored membrane protein required for cell wall biogenesis in bud formation, involved in filamentous growth, homologous to Dcw1p	Plasma membrane ATP-binding cassette (ABC) transporter, weak-acid-inducible multidrug transporter required for weak organic acid resistance; induced by sorbate and benzoate and regulated by War1p; mutants exhibit sorbate hypersensitivity
234	2	AGA1	SUR7	Anchorage subunit of a-agglutinin of a-cells, highly O-glycosylated protein with N-terminal secretion signal and C-terminal signal for addition of GPI anchor to cell wall, linked to adhesion subunit Aga2p via two disulfide bonds	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants
235	2	ERG1	ZPS1	Squalene epoxidase, catalyzes the epoxidation of squalene to 2,3-oxidosqualene; plays an essential role in the ergosterol-biosynthesis pathway and is the specific target of the antifungal drug terbinafine	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH
236	2	ARN2	VBA3 /// VBA5	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to the siderophore triacetylfusarinine C	Permease of basic amino acids in the vacuolar membrane /// Putative transporter of the Major Facilitator Superfamily (MFS); proposed role as a basic amino acid permease based on phylogeny
237	2	VMR1	ZPS1	Protein of unknown function that may interact with ribosomes, based on co-purification experiments; member of the ATP-binding cassette (ABC) family; potential Cdc28p substrate; detected in purified mitochondria in high-throughput studies	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH
238	2	ARN1	SOD2	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to ferrirubin, ferrirhodin, and related siderophores	Mitochondrial superoxide dismutase, protects cells against oxygen toxicity; phosphorylated
239	2	PCA1	RNR4	Cadmium transporting P-type ATPase; may also have a role in copper and iron homeostasis; stabilized by Cd binding, which prevents ubiquitination; S288C and other lab strains contain a G970R mutation which eliminates Cd transport function	Ribonucleotide-diphosphate reductase (RNR), small subunit; the RNR complex catalyzes the rate-limiting step in dNTP synthesis and is regulated by DNA replication and DNA damage checkpoint pathways via localization of the small subunits
240	2	CTS1	WTM1	Endochitinase, required for cell separation after mitosis; transcriptional activation during the G1 phase of the cell cycle is mediated by transcription factor Ace2p	Transcriptional modulator involved in regulation of meiosis, silencing, and expression of RNR genes; required for nuclear localization of the ribonucleotide reductase small subunit Rnr2p and Rnr4p; contains WD repeats
241	2	AGA1	ZPS1	Anchorage subunit of a-agglutinin of a-cells, highly O-glycosylated protein with N-terminal secretion signal and C-terminal signal for addition of GPI anchor to cell wall, linked to adhesion subunit Aga2p via two disulfide bonds	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH
242	1	ENB1	VMR1	Endosomal ferric enterobactin transporter, expressed under conditions of iron deprivation; member of the major facilitator superfamily; expression is regulated by Rcs1p and affected by chloroquine treatment	Protein of unknown function that may interact with ribosomes, based on co-purification experiments; member of the ATP-binding cassette (ABC) family; potential Cdc28p substrate; detected in purified mitochondria in high-throughput studies
243	1	ARN2	SIT1	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to the siderophore triacetylfusarinine C	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p
244	1	DFG5	GSC2	Putative mannosidase, essential glycosylphosphatidylinositol (GPI)-anchored membrane protein required for cell wall biogenesis in bud formation, involved in filamentous growth, homologous to Dcw1p	Catalytic subunit of 1,3-beta-glucan synthase, involved in formation of the inner layer of the spore wall; activity positively regulated by Rho1p and negatively by Smk1p; has similarity to an alternate catalytic subunit, Fks1p (Gsc1p)
245	1	FIT3	HEM13	Mannoprotein that is incorporated into the cell wall via a glycosylphosphatidylinositol (GPI) anchor, involved in the retention of siderophore-iron in the cell wall	Coproporphyrinogen III oxidase, an oxygen requiring enzyme that catalyzes the sixth step in the heme biosynthetic pathway; localizes to the mitochondrial inner membrane; transcription is repressed by oxygen and heme (via Rox1p and Hap1p)
246	1	ILV3	OLE1	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids	Delta(9) fatty acid desaturase, required for monounsaturated fatty acid synthesis and for normal distribution of mitochondria
247	1	PDR12	SAG1	Plasma membrane ATP-binding cassette (ABC) transporter, weak-acid-inducible multidrug transporter required for weak organic acid resistance; induced by	Alpha-agglutinin of alpha-cells, binds to Aga1p during agglutination, N-terminal half is homologous to the immunoglobulin superfamily and contains binding site for a-

				sorbate and benzoate and regulated by War1p; mutants exhibit sorbate hypersensitivity	agglutinin, C-terminal half is highly glycosylated and contains GPI anchor
248	1	GSC2	HEM13	Catalytic subunit of 1,3-beta-glucan synthase, involved in formation of the inner layer of the spore wall; activity positively regulated by Rho1p and negatively by Smk1p; has similarity to an alternate catalytic subunit, Fks1p (Gsc1p)	Coproporphyrinogen III oxidase, an oxygen requiring enzyme that catalyzes the sixth step in the heme biosynthetic pathway; localizes to the mitochondrial inner membrane; transcription is repressed by oxygen and heme (via Rox1p and Hap1p)
249	1	HXT5	ILV3	Hexose transporter with moderate affinity for glucose, induced in the presence of non-fermentable carbon sources, induced by a decrease in growth rate, contains an extended N-terminal domain relative to other HXTs	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids
250	1	1778446_at	VMR1	0	Protein of unknown function that may interact with ribosomes, based on co-purification experiments; member of the ATP-binding cassette (ABC) family; potential Cdc28p substrate; detected in purified mitochondria in high-throughput studies
251	1	PDR12	ZPS1	Plasma membrane ATP-binding cassette (ABC) transporter, weak-acid-inducible multidrug transporter required for weak organic acid resistance; induced by sorbate and benzoate and regulated by War1p; mutants exhibit sorbate hypersensitivity	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH
252	1	SIT1	URA2	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p	Bifunctional carbamoylphosphate synthetase (CPSase)-aspartate transcarbamylase (ATCase), catalyzes the first two enzymatic steps in the de novo biosynthesis of pyrimidines; both activities are subject to feedback inhibition by UTP
253	1	AGA1	CHS1	Anchorage subunit of a-agglutinin of a-cells, highly O-glycosylated protein with N-terminal secretion signal and C-terminal signal for addition of GPI anchor to cell wall, linked to adhesion subunit Aga2p via two disulfide bonds	Chitin synthase I, requires activation from zymogenic form in order to catalyze the transfer of N-acetylglucosamine (GlcNAc) to chitin; required for repairing the chitin septum during cytokinesis; transcription activated by mating factor
254	1	GDE1	ILV3	Glycerophosphocholine (GroPCho) phosphodiesterase; hydrolyzes GroPCho to choline and glycerolphosphate, for use as a phosphate source and as a precursor for phosphocholine synthesis; may interact with ribosomes	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids
255	1	DFG5	ILV3	Putative mannosidase, essential glycosylphosphatidylinositol (GPI)-anchored membrane protein required for cell wall biogenesis in bud formation, involved in filamentous growth, homologous to Dcw1p	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids
256	1	CIT2	OLE1	Citrate synthase, catalyzes the condensation of acetyl coenzyme A and oxaloacetate to form citrate, peroxisomal isozyme involved in glyoxylate cycle; expression is controlled by Rtg1p and Rtg2p transcription factors	Delta(9) fatty acid desaturase, required for monounsaturated fatty acid synthesis and for normal distribution of mitochondria
257	1	GSC2	PCA1	Catalytic subunit of 1,3-beta-glucan synthase, involved in formation of the inner layer of the spore wall; activity positively regulated by Rho1p and negatively by Smk1p; has similarity to an alternate catalytic subunit, Fks1p (Gsc1p)	Cadmium transporting P-type ATPase; may also have a role in copper and iron homeostasis; stabilized by Cd binding, which prevents ubiquitination; S288C and other lab strains contain a G970R mutation which eliminates Cd transport function
258	1	TRP5	WTM1	Tryptophan synthase involved in tryptophan biosynthesis, regulated by the general control system of amino acid biosynthesis	Transcriptional modulator involved in regulation of meiosis, silencing, and expression of RNR genes; required for nuclear localization of the ribonucleotide reductase small subunit Rnr2p and Rnr4p; contains WD repeats
259	1	TIS11	TKL1	mRNA-binding protein expressed during iron starvation; binds to a sequence element in the 3'-untranslated regions of specific mRNAs to mediate their degradation; involved in iron homeostasis	Transketolase, similar to Tkl2p; catalyzes conversion of xylulose-5-phosphate and ribose-5-phosphate to sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate in the pentose phosphate pathway; needed for synthesis of aromatic amino acids
260	1	SIT1	VMR1	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p	Protein of unknown function that may interact with ribosomes, based on co-purification experiments; member of the ATP-binding cassette (ABC) family; potential Cdc28p substrate; detected in purified mitochondria in high-throughput studies
261	1	ILV3	RNR4	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids	Ribonucleotide-diphosphate reductase (RNR), small subunit; the RNR complex catalyzes the rate-limiting step in dNTP synthesis and is regulated by DNA replication and DNA damage checkpoint pathways via localization of the small subunits
262	1	HXT5	ZPS1	Hexose transporter with moderate affinity for glucose, induced in the presence of non-fermentable carbon sources, induced by a decrease in growth rate, contains an extended N-terminal domain relative to other HXTs	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH

263	1	PCA1	SAG1	Cadmium transporting P-type ATPase; may also have a role in copper and iron homeostasis; stabilized by Cd binding, which prevents ubiquitination; S288C and other lab strains contain a G970R mutation which eliminates Cd transport function	Alpha-agglutinin of alpha-cells, binds to Aga1p during agglutination, N-terminal half is homologous to the immunoglobulin superfamily and contains binding site for a-agglutinin, C-terminal half is highly glycosylated and contains GPI anchor
264	1	GSC2	SUR7	Catalytic subunit of 1,3-beta-glucan synthase, involved in formation of the inner layer of the spore wall; activity positively regulated by Rho1p and negatively by Smk1p; has similarity to an alternate catalytic subunit, Fks1p (Gsc1p)	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants
265	1	AHA1	CIT2	Co-chaperone that binds to Hsp82p and activates its ATPase activity; similar to Hch1p; expression is regulated by stresses such as heat shock	Citrate synthase, catalyzes the condensation of acetyl coenzyme A and oxaloacetate to form citrate, peroxisomal isozyme involved in glyoxylate cycle; expression is controlled by Rtg1p and Rtg2p transcription factors
266	1	CHS1	SAH1	Chitin synthase I, requires activation from zymogenic form in order to catalyze the transfer of N-acetylglucosamine (GlcNAc) to chitin; required for repairing the chitin septum during cytokinesis; transcription activated by mating factor	S-adenosyl-L-homocysteine hydrolase, catabolizes S-adenosyl-L-homocysteine which is formed after donation of the activated methyl group of S-adenosyl-L-methionine (AdoMet) to an acceptor
267	1	ENB1	ZPS1	Endosomal ferric enterobactin transporter, expressed under conditions of iron deprivation; member of the major facilitator superfamily; expression is regulated by Rcs1p and affected by chloroquine treatment	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH
268	1	HEM13	YKE4	Coproporphyrinogen III oxidase, an oxygen requiring enzyme that catalyzes the sixth step in the heme biosynthetic pathway; localizes to the mitochondrial inner membrane; transcription is repressed by oxygen and heme (via Rox1p and Hap1p)	Zinc transporter; localizes to the ER; null mutant is sensitive to calcofluor white, leads to zinc accumulation in cytosol; ortholog of the mouse KE4 and member of the ZIP (ZRT, IRT-like Protein) family
269	1	AHA1	SCS7	Co-chaperone that binds to Hsp82p and activates its ATPase activity; similar to Hch1p; expression is regulated by stresses such as heat shock	Sphingolipid alpha-hydroxylase, functions in the alpha-hydroxylation of sphingolipid-associated very long chain fatty acids, has both cytochrome b5-like and hydroxylase/desaturase domains, not essential for growth
270	1	FET3	TIS11	Ferro-O2-oxidoreductase required for high-affinity iron uptake and involved in mediating resistance to copper ion toxicity, belongs to class of integral membrane multicopper oxidases	mRNA-binding protein expressed during iron starvation; binds to a sequence element in the 3'-untranslated regions of specific mRNAs to mediate their degradation; involved in iron homeostasis
271	1	RNR4	SIT1	Ribonucleotide-diphosphate reductase (RNR), small subunit; the RNR complex catalyzes the rate-limiting step in dNTP synthesis and is regulated by DNA replication and DNA damage checkpoint pathways via localization of the small subunits	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p
272	1	PDR15	ZPS1	Plasma membrane ATP binding cassette (ABC) transporter, multidrug transporter and general stress response factor implicated in cellular detoxification; regulated by Pdr1p, Pdr3p and Pdr8p; promoter contains a PDR responsive element	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH
273	1	TRP5	ZPS1	Tryptophan synthase involved in tryptophan biosynthesis, regulated by the general control system of amino acid biosynthesis	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH
274	1	ARN2	MCH2	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to the siderophore triacetylfusarinine C	Protein with similarity to mammalian monocarboxylate permeases, which are involved in transport of monocarboxylic acids across the plasma membrane; mutant is not deficient in monocarboxylate transport
275	1	FET3	SIT1	Ferro-O2-oxidoreductase required for high-affinity iron uptake and involved in mediating resistance to copper ion toxicity, belongs to class of integral membrane multicopper oxidases	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p
276	1	HEM13	VMR1	Coproporphyrinogen III oxidase, an oxygen requiring enzyme that catalyzes the sixth step in the heme biosynthetic pathway; localizes to the mitochondrial inner membrane; transcription is repressed by oxygen and heme (via Rox1p and Hap1p)	Protein of unknown function that may interact with ribosomes, based on co-purification experiments; member of the ATP-binding cassette (ABC) family; potential Cdc28p substrate; detected in purified mitochondria in high-throughput studies
277	1	SUR7	WTM1	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants	Transcriptional modulator involved in regulation of meiosis, silencing, and expression of RNR genes; required for nuclear localization of the ribonucleotide reductase small subunit Rnr2p and Rnr4p; contains WD repeats
278	1	1774824_at	GSC2	0	Catalytic subunit of 1,3-beta-glucan synthase, involved in formation of the inner layer of the spore wall; activity positively regulated by Rho1p and negatively by Smk1p; has similarity to an alternate catalytic subunit, Fks1p (Gsc1p)
279	1	AHA1	ARN2	Co-chaperone that binds to Hsp82p and activates its ATPase activity; similar to Hch1p; expression is regulated by stresses such as heat shock	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to the siderophore triacetylfusarinine C

280	1	MRH1	ZPS1	Protein that localizes primarily to the plasma membrane, also found at the nuclear envelope; the authentic, non-tagged protein is detected in mitochondria in a phosphorylated state; has similarity to Hsp30p and Yro2p	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH
281	1	AGA1	SIT1	Anchorage subunit of a-agglutinin of a-cells, highly O-glycosylated protein with N-terminal secretion signal and C-terminal signal for addition of GPI anchor to cell wall, linked to adhesion subunit Aga2p via two disulfide bonds	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p
282	1	SAG1	SUR7	Alpha-agglutinin of alpha-cells, binds to Aga1p during agglutination, N-terminal half is homologous to the immunoglobulin superfamily and contains binding site for a-agglutinin, C-terminal half is highly glycosylated and contains GPI anchor	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants
283	1	CHS1	SAG1	Chitin synthase I, requires activation from zymogenic form in order to catalyze the transfer of N-acetylglucosamine (GlcNAc) to chitin; required for repairing the chitin septum during cytokinesis; transcription activated by mating factor	Alpha-agglutinin of alpha-cells, binds to Aga1p during agglutination, N-terminal half is homologous to the immunoglobulin superfamily and contains binding site for a-agglutinin, C-terminal half is highly glycosylated and contains GPI anchor
284	1	HEM13	TRP5	Coproporphyrinogen III oxidase, an oxygen requiring enzyme that catalyzes the sixth step in the heme biosynthetic pathway; localizes to the mitochondrial inner membrane; transcription is repressed by oxygen and heme (via Rox1p and Hap1p)	Tryptophan synthase involved in tryptophan biosynthesis, regulated by the general control system of amino acid biosynthesis
285	1	CHS1	ILV3	Chitin synthase I, requires activation from zymogenic form in order to catalyze the transfer of N-acetylglucosamine (GlcNAc) to chitin; required for repairing the chitin septum during cytokinesis; transcription activated by mating factor	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids
286	1	TKL1	WTM1	Transketolase, similar to Tkl2p; catalyzes conversion of xylulose-5-phosphate and ribose-5-phosphate to sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate in the pentose phosphate pathway; needed for synthesis of aromatic amino acids	Transcriptional modulator involved in regulation of meiosis, silencing, and expression of RNR genes; required for nuclear localization of the ribonucleotide reductase small subunit Rnr2p and Rnr4p; contains WD repeats
287	1	ARN2	TRP5	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to the siderophore triacetylfusarinine C	Tryptophan synthase involved in tryptophan biosynthesis, regulated by the general control system of amino acid biosynthesis
288	1	AGP1	SCS7	Low-affinity amino acid permease with broad substrate range, involved in uptake of asparagine, glutamine, and other amino acids; expression is regulated by the SPS plasma membrane amino acid sensor system (Ssy1p-Ptr3p-Ssy5p)	Sphingolipid alpha-hydroxylase, functions in the alpha-hydroxylation of sphingolipid-associated very long chain fatty acids, has both cytochrome b5-like and hydroxylase/desaturase domains, not essential for growth
289	1	FET3	SCS7	Ferro-O2-oxidoreductase required for high-affinity iron uptake and involved in mediating resistance to copper ion toxicity, belongs to class of integral membrane multicopper oxidases	Sphingolipid alpha-hydroxylase, functions in the alpha-hydroxylation of sphingolipid-associated very long chain fatty acids, has both cytochrome b5-like and hydroxylase/desaturase domains, not essential for growth
290	1	AGA1	ERG1	Anchorage subunit of a-agglutinin of a-cells, highly O-glycosylated protein with N-terminal secretion signal and C-terminal signal for addition of GPI anchor to cell wall, linked to adhesion subunit Aga2p via two disulfide bonds	Squalene epoxidase, catalyzes the epoxidation of squalene to 2,3-oxidosqualene; plays an essential role in the ergosterol-biosynthesis pathway and is the specific target of the antifungal drug terbinafine
291	1	ARN2	INO1	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to the siderophore triacetylfusarinine C	Inositol 1-phosphate synthase, involved in synthesis of inositol phosphates and inositol-containing phospholipids; transcription is coregulated with other phospholipid biosynthetic genes by Ino2p and Ino4p, which bind the UASINO DNA element

Annex 2. Table of central genes in the network Dataset1

	Gene	N interactions	Description
1	ZPS1	17	Putative GPI-anchored protein; transcription is induced under low-zinc conditions, as mediated by the Zap1p transcription factor, and at alkaline pH
2	SUR7	14	Putative integral membrane protein; component of eisosomes; associated with endocytosis, along with Pil1p and Lsp1p; sporulation and plasma membrane sphingolipid content are altered in mutants
3	SIT1	14	Ferrioxamine B transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; transcription is induced during iron deprivation and diauxic shift; potentially phosphorylated by Cdc28p

4	TRP5	12	Tryptophan synthase involved in tryptophan biosynthesis, regulated by the general control system of amino acid biosynthesis
5	URA2	10	Bifunctional carbamoylphosphate synthetase (CPSase)-aspartate transcarbamylase (ATCase), catalyzes the first two enzymatic steps in the de novo biosynthesis of pyrimidines; both activities are subject to feedback inhibition by UTP
6	WTM1	10	Transcriptional modulator involved in regulation of meiosis, silencing, and expression of RNR genes; required for nuclear localization of the ribonucleotide reductase small subunit Rnr2p and Rnr4p; contains WD repeats
7	VMR1	10	Protein of unknown function that may interact with ribosomes, based on co-purification experiments; member of the ATP-binding cassette (ABC) family; potential Cdc28p substrate; detected in purified mitochondria in high-throughput studies
8	PDR12	10	Plasma membrane ATP-binding cassette (ABC) transporter, weak-acid-inducible multidrug transporter required for weak organic acid resistance; induced by sorbate and benzoate and regulated by War1p; mutants exhibit sorbate hypersensitivity
9	VBA3 /// VBA5	9	Permease of basic amino acids in the vacuolar membrane /// Putative transporter of the Major Facilitator Superfamily (MFS); proposed role as a basic amino acid permease based on phylogeny
10	YRO2	9	Putative protein of unknown function; the authentic, non-tagged protein is detected in a phosphorylated state in highly purified mitochondria in high-throughput studies; transcriptionally regulated by Haa1p
11	SAG1	8	Alpha-agglutinin of alpha-cells, binds to Aga1p during agglutination, N-terminal half is homologous to the immunoglobulin superfamily and contains binding site for a-agglutinin, C-terminal half is highly glycosylated and contains GPI anchor
12	OLE1	8	Delta(9) fatty acid desaturase, required for monounsaturated fatty acid synthesis and for normal distribution of mitochondria
13	ILV3	8	Dihydroxyacid dehydratase, catalyzes third step in the common pathway leading to biosynthesis of branched-chain amino acids
14	HEM13	7	Coproporphyrinogen III oxidase, an oxygen requiring enzyme that catalyzes the sixth step in the heme biosynthetic pathway; localizes to the mitochondrial inner membrane; transcription is repressed by oxygen and heme (via Rox1p and Hap1p)
15	SCS7	7	Sphingolipid alpha-hydroxylase, functions in the alpha-hydroxylation of sphingolipid-associated very long chain fatty acids, has both cytochrome b5-like and hydroxylase/desaturase domains, not essential for growth
16	TKL1	7	Transketolase, similar to Tkl2p; catalyzes conversion of xylulose-5-phosphate and ribose-5-phosphate to sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate in the pentose phosphate pathway; needed for synthesis of aromatic amino acids
17	MCH2	6	Protein with similarity to mammalian monocarboxylate permeases, which are involved in transport of monocarboxylic acids across the plasma membrane; mutant is not deficient in monocarboxylate transport
18	SOD2	6	Mitochondrial superoxide dismutase, protects cells against oxygen toxicity; phosphorylated
19	PDR10	5	ATP-binding cassette (ABC) transporter, multidrug transporter involved in the pleiotropic drug resistance network; regulated by Pdr1p and Pdr3p
20	YKE4	5	Zinc transporter; localizes to the ER; null mutant is sensitive to calcofluor white, leads to zinc accumulation in cytosol; ortholog of the mouse KE4 and member of the ZIP (ZRT, IRT-like Protein) family
21	MRH1	5	Protein that localizes primarily to the plasma membrane, also found at the nuclear envelope; the authentic, non-tagged protein is detected in mitochondria in a phosphorylated state; has similarity to Hsp30p and Yro2p
22	GDE1	5	Glycerophosphocholine (GroPCho) phosphodiesterase; hydrolyzes GroPCho to choline and glycerolphosphate, for use as a phosphate source and as a precursor for phosphocholine synthesis; may interact with ribosomes
23	GSC2	5	Catalytic subunit of 1,3-beta-glucan synthase, involved in formation of the inner layer of the spore wall; activity positively regulated by Rho1p and negatively by Smk1p; has similarity to an alternate catalytic subunit, Fks1p (Gsc1p)
24	RNR4	5	Ribonucleotide-diphosphate reductase (RNR), small subunit; the RNR complex catalyzes the rate-limiting step in dNTP synthesis and is regulated by DNA replication and DNA damage checkpoint pathways via localization of the small subunits
25	PET10	4	Protein of unknown function that co-purifies with lipid particles; expression pattern suggests a role in respiratory growth; computational analysis of large-scale protein-protein interaction data suggests a role in ATP/ADP exchange
26	PDR5	4	Plasma membrane ATP-binding cassette (ABC) transporter, multidrug transporter actively regulated by Pdr1p; also involved in steroid transport, cation resistance, and cellular detoxification during exponential growth
27	INO1	4	Inositol 1-phosphate synthase, involved in synthesis of inositol phosphates and inositol-containing phospholipids; transcription is coregulated with other phospholipid biosynthetic genes by Ino2p and Ino4p, which bind the UASINO DNA element
28	MF(ALPHA)1	4	Mating pheromone alpha-factor, made by alpha cells; interacts with mating type a cells to induce cell cycle arrest and other responses leading to mating; also encoded by MF(ALPHA)2, although MF(ALPHA)1 produces most alpha-factor
29	FIT2	3	Mannoprotein that is incorporated into the cell wall via a glycosylphosphatidylinositol (GPI) anchor, involved in the retention of siderophore-iron in the cell wall
30	TIS11	3	mRNA-binding protein expressed during iron starvation; binds to a sequence element in the 3'-untranslated regions of specific mRNAs to mediate their degradation; involved in iron homeostasis

31	DED81	3	Cytosolic asparaginyl-tRNA synthetase, required for protein synthesis, catalyzes the specific attachment of asparagine to its cognate tRNA
32	GLN1	3	Glutamine synthetase (GS), synthesizes glutamine from glutamate and ammonia; with Glt1p, forms the secondary pathway for glutamate biosynthesis from ammonia; expression regulated by nitrogen source and by amino acid limitation
33	PCA1	3	Cadmium transporting P-type ATPase; may also have a role in copper and iron homeostasis; stabilized by Cd binding, which prevents ubiquitination; S288C and other lab strains contain a G970R mutation which eliminates Cd transport function
34	FET3	3	Ferro-O2-oxidoreductase required for high-affinity iron uptake and involved in mediating resistance to copper ion toxicity, belongs to class of integral membrane multicopper oxidases
35	HXT5	3	Hexose transporter with moderate affinity for glucose, induced in the presence of non-fermentable carbon sources, induced by a decrease in growth rate, contains an extended N-terminal domain relative to other HXTs
36	ENB1	3	Endosomal ferric enterobactin transporter, expressed under conditions of iron deprivation; member of the major facilitator superfamily; expression is regulated by Rcs1p and affected by chloroquine treatment
37	ARN1	3	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to ferrirubin, ferrirhodin, and related siderophores
38	URA1	2	Dihydroorotate dehydrogenase, catalyzes the fourth enzymatic step in the de novo biosynthesis of pyrimidines, converting dihydroorotic acid into orotic acid
39	FAA4	2	Long chain fatty acyl-CoA synthetase, regulates protein modification during growth in the presence of ethanol, functions to incorporate palmitic acid into phospholipids and neutral lipids
40	HTB1	2	Histone H2B, core histone protein required for chromatin assembly and chromosome function; nearly identical to HTB2; Rad6p-Bre1p-Lge1p mediated ubiquitination regulates transcriptional activation, meiotic DSB formation and H3 methylation
41	ICY1	2	Protein of unknown function, required for viability in rich media of cells lacking mitochondrial DNA; mutants have an invasive growth defect with elongated morphology; induced by amino acid starvation
42	SAH1	2	S-adenosyl-L-homocysteine hydrolase, catabolizes S-adenosyl-L-homocysteine which is formed after donation of the activated methyl group of S-adenosyl-L-methionine (AdoMet) to an acceptor
43	ZRT3	2	Vacuolar membrane zinc transporter, transports zinc from storage in the vacuole to the cytoplasm when needed; transcription is induced under conditions of zinc deficiency
44	CHA1	2	Catabolic L-serine (L-threonine) deaminase, catalyzes the degradation of both L-serine and L-threonine; required to use serine or threonine as the sole nitrogen source, transcriptionally induced by serine and threonine
45	TKL2	2	Transketolase, similar to Tkl1p; catalyzes conversion of xylulose-5-phosphate and ribose-5-phosphate to sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate in the pentose phosphate pathway; needed for synthesis of aromatic amino acids
46	GRE2	2	3-methylbutanal reductase and NADPH-dependent methylglyoxal reductase (D-lactaldehyde dehydrogenase); stress induced (osmotic, ionic, oxidative, heat shock and heavy metals); regulated by the HOG pathway
47	PIS1	2	Phosphatidylinositol synthase, required for biosynthesis of phosphatidylinositol, which is a precursor for polyphosphoinositides, sphingolipids, and glycolipid anchors for some of the plasma membrane proteins
48	ERG1	2	Squalene epoxidase, catalyzes the epoxidation of squalene to 2,3-oxidosqualene; plays an essential role in the ergosterol-biosynthesis pathway and is the specific target of the antifungal drug terbinafine
49	RNR2	2	Ribonucleotide-diphosphate reductase (RNR), small subunit; the RNR complex catalyzes the rate-limiting step in dNTP synthesis and is regulated by DNA replication and DNA damage checkpoint pathways via localization of the small subunits
50	CIT2	2	Citrate synthase, catalyzes the condensation of acetyl coenzyme A and oxaloacetate to form citrate, peroxisomal isozyme involved in glyoxylate cycle; expression is controlled by Rtg1p and Rtg2p transcription factors
51	ARN2	2	Transporter, member of the ARN family of transporters that specifically recognize siderophore-iron chelates; responsible for uptake of iron bound to the siderophore triacetylfusarinine C
52	DFG5	2	Putative mannosidase, essential glycosylphosphatidylinositol (GPI)-anchored membrane protein required for cell wall biogenesis in bud formation, involved in filamentous growth, homologous to Dcw1p
53	CHS1	2	Chitin synthase I, requires activation from zymogenic form in order to catalyze the transfer of N-acetylglucosamine (GlcNAc) to chitin; required for repairing the chitin septum during cytokinesis; transcription activated by mating factor
54	GDB1	1	Glycogen debranching enzyme containing glucanotranferase and alpha-1,6-amyloglucosidase activities, required for glycogen degradation; phosphorylated in mitochondria
55	SSB2	1	Cytoplasmic ATPase that is a ribosome-associated molecular chaperone, functions with J-protein partner Zuo1p; may be involved in the folding of newly-synthesized polypeptide chains; member of the HSP70 family; homolog of SSB1
56	TIR3	1	Cell wall mannoprotein of the Srp1p/Tip1p family of serine-alanine-rich proteins; expressed under anaerobic conditions and required for anaerobic growth
57	MDH2	1	Cytoplasmic malate dehydrogenase, one of three isozymes that catalyze interconversion of malate and oxaloacetate; involved in the glyoxylate cycle and gluconeogenesis during growth on two-carbon compounds; interacts with Pck1p and Fbp1
58	1778446_at	1	0

59	ILV5	1	Acetohydroxyacid reductoisomerase, mitochondrial protein involved in branched-chain amino acid biosynthesis, also required for maintenance of wild-type mitochondrial DNA and found in mitochondrial nucleoids
60	ZRT1	1	High-affinity zinc transporter of the plasma membrane, responsible for the majority of zinc uptake; transcription is induced under low-zinc conditions by the Zap1p transcription factor
61	AGX1	1	Alanine:glyoxylate aminotransferase (AGT), catalyzes the synthesis of glycine from glyoxylate, which is one of three pathways for glycine biosynthesis in yeast; has similarity to mammalian and plant alanine:glyoxylate aminotransferases
62	BAP2	1	High-affinity leucine permease, functions as a branched-chain amino acid permease involved in the uptake of leucine, isoleucine and valine; contains 12 predicted transmembrane domains
63	PDR15	1	Plasma membrane ATP binding cassette (ABC) transporter, multidrug transporter and general stress response factor implicated in cellular detoxification; regulated by Pdr1p, Pdr3p and Pdr8p; promoter contains a PDR responsive element
64	HXK2	1	Hexokinase isoenzyme 2 that catalyzes phosphorylation of glucose in the cytosol; predominant hexokinase during growth on glucose; functions in the nucleus to repress expression of HXK1 and GLK1 and to induce expression of its own gene
65	AMS1	1	Vacuolar alpha mannosidase, involved in free oligosaccharide (fOS) degradation; delivered to the vacuole in a novel pathway separate from the secretory pathway
66	YPS1	1	Aspartic protease, attached to the plasma membrane via a glycosylphosphatidylinositol (GPI) anchor
67	FIT3	1	Mannoprotein that is incorporated into the cell wall via a glycosylphosphatidylinositol (GPI) anchor, involved in the retention of siderophore-iron in the cell wall

Annex 3. Table of predicted interactions Dataset2

	N times	Gene1	Gene2	Description1	Description2
1	20	SHISA2	SLITRK4	Homo sapiens shisa homolog 2 (Xenopus laevis) (SHISA2), mRNA.	Homo sapiens SLIT and NTRK-like family, member 4 (SLITRK4), mRNA.
2	20	AURKA	FRZB	Homo sapiens aurora kinase A (AURKA), transcript variant 5, mRNA.	Homo sapiens frizzled-related protein (FRZB), mRNA.
3	20	IL32	LOC100008589	Homo sapiens interleukin 32 (IL32), transcript variant 7, mRNA.	Homo sapiens 28S ribosomal RNA (LOC100008589), non-coding RNA.
4	20	THBS4	IL8	Homo sapiens thrombospondin 4 (THBS4), mRNA.	Homo sapiens interleukin 8 (IL8), mRNA.
5	20	SPINK1	ALG1L	Homo sapiens serine peptidase inhibitor, Kazal type 1 (SPINK1), mRNA.	Homo sapiens asparagine-linked glycosylation 1-like (ALG1L), mRNA.
6	20	CCL2	GSTM1	Homo sapiens chemokine (C-C motif) ligand 2 (CCL2), mRNA.	Homo sapiens glutathione S-transferase M1 (GSTM1), transcript variant 1, mRNA.
7	20	SLC25A25	SLITRK4	Homo sapiens solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 25 (SLC25A25), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA.	Homo sapiens SLIT and NTRK-like family, member 4 (SLITRK4), mRNA.
8	20	ATP2A1	LOC649143	Homo sapiens ATPase, Ca++ transporting, cardiac muscle, fast twitch 1 (ATP2A1), transcript variant a, mRNA.	PREDICTED: Homo sapiens similar to HLA class II histocompatibility antigen, DRB1-9 beta chain precursor (MHC class I antigen DRB1*9) (DR-9) (DR9), transcript variant 2 (LOC649143), mRNA.
9	20	MT1E	AURKA	Homo sapiens metallothionein 1E (MT1E), mRNA.	['Homo sapiens aurora kinase A (AURKA), transcript variant 5, mRNA.' 'Homo sapiens aurora kinase A (AURKA), transcript variant 3, mRNA.']
10	20	PTER	ZP3	Homo sapiens phosphotriesterase related (PTER), transcript variant 2, mRNA.	Homo sapiens zona pellucida glycoprotein 3 (sperm receptor) (ZP3), mRNA.
11	20	GSTT1	FAM174A	Homo sapiens glutathione S-transferase theta 1 (GSTT1), mRNA.	Homo sapiens family with sequence similarity 174, member A (FAM174A), mRNA.
12	20	OLFM1	CD79B	Homo sapiens olfactomedin 1 (OLFM1), transcript variant 2, mRNA.	Homo sapiens CD79b molecule, immunoglobulin-associated beta (CD79B), transcript variant 3, mRNA.
13	20	TMEM70	LOC100134364	Homo sapiens transmembrane protein 70 (TMEM70), transcript variant 1, mRNA.	PREDICTED: Homo sapiens hypothetical protein LOC100134364 (LOC100134364), mRNA.
14	20	CLYBL	HLA-A29.1	Homo sapiens citrate lyase beta like (CLYBL), mRNA.	Homo sapiens major histocompatibility complex class I HLA-A29.1 (HLA-A29.1), mRNA. XM_001132736 XM_001132742 XM_001132747 XM_001132750 XM_001132753 XM_001132757 XM_001132760 XM_001132769 XM_001132772 XM_001132776 XM_001132779 XM_001132781 XM_001132783 XM_001132785 XM_001132794
15	19	ACTN3	COL3A1	Homo sapiens actinin, alpha 3 (ACTN3), mRNA.	Homo sapiens collagen, type III, alpha 1 (COL3A1), mRNA.
16	19	POMC	SLC2A3	Homo sapiens proopiomelanocortin (POMC), transcript variant 1, mRNA.	Homo sapiens solute carrier family 2 (facilitated glucose transporter), member 3 (SLC2A3), mRNA.
17	19	POMC	SPINK1	Homo sapiens proopiomelanocortin (POMC), transcript variant 1, mRNA.	Homo sapiens serine peptidase inhibitor, Kazal type 1 (SPINK1), mRNA.
18	19	MFF	LOC100133565	Homo sapiens mitochondrial fission factor (MFF), nuclear gene encoding mitochondrial protein, mRNA.	PREDICTED: Homo sapiens similar to hCG23738 (LOC100133565), mRNA.
19	19	POMC	HLA-DRB5	Homo sapiens proopiomelanocortin (POMC), transcript variant 1, mRNA.	Homo sapiens major histocompatibility complex, class II, DR beta 5 (HLA-DRB5), mRNA.
20	19	CHURC1	ZP3	Homo sapiens churchill domain containing 1 (CHURC1), mRNA.	Homo sapiens zona pellucida glycoprotein 3 (sperm receptor) (ZP3), mRNA.

21	19	ITGB1BP1	HMGCS2	Homo sapiens integrin beta 1 binding protein 1 (ITGB1BP1), transcript variant 1, mRNA.	Homo sapiens 3-hydroxy-3-methylglutaryl-Coenzyme A synthase 2 (mitochondrial) (HMGCS2), nuclear gene encoding mitochondrial protein, mRNA.
22	19	TNNI1	HBG1	Homo sapiens troponin I type 1 (skeletal, slow) (TNNI1), mRNA.	Homo sapiens hemoglobin, gamma A (HBG1), mRNA.
23	19	HLA-DRB5	HBG1	Homo sapiens major histocompatibility complex, class II, DR beta 5 (HLA-DRB5), mRNA.	Homo sapiens hemoglobin, gamma A (HBG1), mRNA.
24	19	AKR1C4	SLC25A25	Homo sapiens aldo-keto reductase family 1, member C4 (chlordecone reductase; 3-alpha hydroxysteroid dehydrogenase, type I; dihydrodiol dehydrogenase 4) (AKR1C4), mRNA.	Homo sapiens solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 25 (SLC25A25), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA.
25	19	BEGAIN	FOSB	Homo sapiens brain-enriched guanylate kinase-associated homolog (rat) (BEGAIN), mRNA.	Homo sapiens FBJ murine osteosarcoma viral oncogene homolog B (FOSB), mRNA.
26	19	IGFN1	AGTPBP1	Homo sapiens immunoglobulin-like and fibronectin type III domain containing 1 (IGFN1), mRNA.	Homo sapiens ATP/GTP binding protein 1 (AGTPBP1), mRNA.
27	19	QPRT	SDHALP1	Homo sapiens quinolinate phosphoribosyltransferase (QPRT), mRNA.	Homo sapiens succinate dehydrogenase complex, subunit A, flavoprotein pseudogene 1 (SDHALP1) on chromosome 3.
28	19	POMC	TNNI1	Homo sapiens proopiomelanocortin (POMC), transcript variant 1, mRNA.	Homo sapiens troponin I type 1 (skeletal, slow) (TNNI1), mRNA.
29	18	PRODH	CHURC1	Homo sapiens proline dehydrogenase (oxidase) 1 (PRODH), nuclear gene encoding mitochondrial protein, mRNA.	Homo sapiens churchill domain containing 1 (CHURC1), mRNA.
30	18	NDUFAF1	AKR1C2	Homo sapiens NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, assembly factor 1 (NDUFAF1), mRNA.	Homo sapiens aldo-keto reductase family 1, member C2 (dihydrodiol dehydrogenase 2; bile acid binding protein; 3-alpha hydroxysteroid dehydrogenase, type III) (AKR1C2), transcript variant 1, mRNA. XM_943424 XM_943425 XM_943427
31	18	MYH1	TINP1	Homo sapiens myosin, heavy polypeptide 1, skeletal muscle, adult (MYH1), mRNA.	Homo sapiens TGF beta-inducible nuclear protein 1 (TINP1), mRNA.
32	18	TNNI1	SPINK1	Homo sapiens troponin I type 1 (skeletal, slow) (TNNI1), mRNA.	Homo sapiens serine peptidase inhibitor, Kazal type 1 (SPINK1), mRNA.
33	18	QPRT	MYBPH	Homo sapiens quinolinate phosphoribosyltransferase (QPRT), mRNA.	Homo sapiens myosin binding protein H (MYBPH), mRNA.
34	18	MYH1	RPL14	Homo sapiens myosin, heavy polypeptide 1, skeletal muscle, adult (MYH1), mRNA.	Homo sapiens ribosomal protein L14 (RPL14), transcript variant 1, mRNA.
35	18	SLC2A3	HBG1	Homo sapiens solute carrier family 2 (facilitated glucose transporter), member 3 (SLC2A3), mRNA.	Homo sapiens hemoglobin, gamma A (HBG1), mRNA.
36	18	ACTC1	AZIN1	Homo sapiens actin, alpha, cardiac muscle 1 (ACTC1), mRNA.	Homo sapiens antizyme inhibitor 1 (AZIN1), transcript variant 2, mRNA.
37	18	EGR1	ATP2A1	Homo sapiens early growth response 1 (EGR1), mRNA.	Homo sapiens ATPase, Ca++ transporting, cardiac muscle, fast twitch 1 (ATP2A1), transcript variant a, mRNA.
38	18	HLA-DRB5	SLC2A3	Homo sapiens major histocompatibility complex, class II, DR beta 5 (HLA-DRB5), mRNA.	Homo sapiens solute carrier family 2 (facilitated glucose transporter), member 3 (SLC2A3), mRNA.
39	18	TNNI1	SLC2A3	Homo sapiens troponin I type 1 (skeletal, slow) (TNNI1), mRNA.	Homo sapiens solute carrier family 2 (facilitated glucose transporter), member 3 (SLC2A3), mRNA.
40	18	IGFN1	CALML6	Homo sapiens immunoglobulin-like and fibronectin type III domain containing 1 (IGFN1), mRNA.	Homo sapiens calmodulin-like 6 (CALML6), mRNA.
41	18	AZIN1	LOC100132394	Homo sapiens antizyme inhibitor 1 (AZIN1), transcript variant 2, mRNA.	PREDICTED: Homo sapiens hypothetical protein LOC100132394 (LOC100132394), mRNA.
42	18	GSDMC	PTER	Homo sapiens gasdermin C (GSDMC), mRNA.	Homo sapiens phosphotriesterase related (PTER), transcript variant 2, mRNA.
43	18	TNNI1	HLA-DRB5	Homo sapiens troponin I type 1 (skeletal, slow) (TNNI1), mRNA.	Homo sapiens major histocompatibility complex, class II, DR beta 5 (HLA-DRB5), mRNA.
44	17	SHISA2	SLC25A25	Homo sapiens shisa homolog 2 (Xenopus laevis) (SHISA2), mRNA.	Homo sapiens solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 25 (SLC25A25), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA.
45	17	AZIN1	LOC647307	Homo sapiens antizyme inhibitor 1 (AZIN1), transcript variant 2, mRNA.	PREDICTED: Homo sapiens misc_RNA (LOC647307), miscRNA.
46	17	IGFN1	TIMM10	Homo sapiens immunoglobulin-like and fibronectin type III domain containing 1 (IGFN1), mRNA.	Homo sapiens translocase of inner mitochondrial membrane 10 homolog (yeast) (TIMM10), nuclear gene encoding mitochondrial protein, mRNA.
47	17	SLC2A3	SPINK1	Homo sapiens solute carrier family 2 (facilitated glucose transporter), member 3 (SLC2A3), mRNA.	Homo sapiens serine peptidase inhibitor, Kazal type 1 (SPINK1), mRNA.
48	17	TTC25	AGTPBP1	Homo sapiens tetratricopeptide repeat domain 25 (TTC25), mRNA.	Homo sapiens ATP/GTP binding protein 1 (AGTPBP1), mRNA.
49	17	HLA-DRB1	LOC100008589	Homo sapiens major histocompatibility complex, class II, DR beta 1 (HLA-DRB1), mRNA.	Homo sapiens 28S ribosomal RNA (LOC100008589), non-coding RNA.
50	17	POMC	HBG1	Homo sapiens proopiomelanocortin (POMC), transcript variant 1, mRNA.	Homo sapiens hemoglobin, gamma A (HBG1), mRNA.
51	17	GSDMC	ZP3	Homo sapiens gasdermin C (GSDMC), mRNA.	Homo sapiens zona pellucida glycoprotein 3 (sperm receptor) (ZP3), mRNA.
52	17	PTER	CHURC1	Homo sapiens phosphotriesterase related (PTER), transcript variant 2, mRNA.	Homo sapiens churchill domain containing 1 (CHURC1), mRNA.
53	17	C17ORF97	LOC647307	Homo sapiens chromosome 17 open reading frame 97 (C17orf97), mRNA.	PREDICTED: Homo sapiens misc_RNA (LOC647307), miscRNA.
54	17	TTC25	GSDMC	Homo sapiens tetratricopeptide repeat domain 25 (TTC25), mRNA.	Homo sapiens gasdermin C (GSDMC), mRNA.
55	17	AKR1C4	SHISA2	Homo sapiens aldo-keto reductase family 1, member C4 (chlordecone reductase; 3-alpha hydroxysteroid dehydrogenase, type I; dihydrodiol dehydrogenase 4) (AKR1C4), mRNA.	Homo sapiens shisa homolog 2 (Xenopus laevis) (SHISA2), mRNA.

56	17	AKR1C4	SLITRK4	Homo sapiens aldo-keto reductase family 1, member C4 (chlordecone reductase; 3-alpha hydroxysteroid dehydrogenase, type I; dihydrodiol dehydrogenase 4) (AKR1C4), mRNA.	Homo sapiens SLIT and NTRK-like family, member 4 (SLITRK4), mRNA.
57	17	NDUFAF1	LOC100133565	Homo sapiens NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, assembly factor 1 (NDUFAF1), mRNA.	PREDICTED: Homo sapiens similar to hCG23738 (LOC100133565), mRNA.
58	17	TTC25	PTER	Homo sapiens tetratricopeptide repeat domain 25 (TTC25), mRNA.	Homo sapiens phosphotriesterase related (PTER), transcript variant 2, mRNA.
59	16	TIMM10	AGTPBP1	Homo sapiens translocase of inner mitochondrial membrane 10 homolog (yeast) (TIMM10), nuclear gene encoding mitochondrial protein, mRNA.	Homo sapiens ATP/GTP binding protein 1 (AGTPBP1), mRNA.
60	16	MAPK8IP1	LOC100131138	Homo sapiens mitogen-activated protein kinase 8 interacting protein 1 (MAPK8IP1), mRNA.	PREDICTED: Homo sapiens similar to hCG2040918 (LOC100131138), mRNA.
61	16	PRODH	LOC100008589	Homo sapiens proline dehydrogenase (oxidase) 1 (PRODH), nuclear gene encoding mitochondrial protein, mRNA.	Homo sapiens 28S ribosomal RNA (LOC100008589), non-coding RNA.
62	16	MYBPH	SDHALP1	Homo sapiens myosin binding protein H (MYBPH), mRNA.	Homo sapiens succinate dehydrogenase complex, subunit A, flavoprotein pseudogene 1 (SDHALP1) on chromosome 3.
63	16	HLA-DRB1	ZP3	Homo sapiens major histocompatibility complex, class II, DR beta 1 (HLA-DRB1), mRNA.	Homo sapiens zona pellucida glycoprotein 3 (sperm receptor) (ZP3), mRNA.
64	16	ACTN3	GSTM1	Homo sapiens actinin, alpha 3 (ACTN3), mRNA.	Homo sapiens glutathione S-transferase M1 (GSTM1), transcript variant 1, mRNA.
65	16	GSTM1	MAPK8IP1	Homo sapiens glutathione S-transferase M1 (GSTM1), transcript variant 1, mRNA.	Homo sapiens mitogen-activated protein kinase 8 interacting protein 1 (MAPK8IP1), mRNA.
66	16	COL3A1	MIR1974	Homo sapiens collagen, type III, alpha 1 (COL3A1), mRNA.	Homo sapiens microRNA 1974 (MIR1974), microRNA.
67	15	SLC2A3	ALG1L	Homo sapiens solute carrier family 2 (facilitated glucose transporter), member 3 (SLC2A3), mRNA.	Homo sapiens asparagine-linked glycosylation 1-like (ALG1L), mRNA.
68	15	TNNI1	ALG1L	Homo sapiens troponin I type 1 (skeletal, slow) (TNNI1), mRNA.	Homo sapiens asparagine-linked glycosylation 1-like (ALG1L), mRNA.
69	15	ACTN3	MAPK8IP1	Homo sapiens actinin, alpha 3 (ACTN3), mRNA.	Homo sapiens mitogen-activated protein kinase 8 interacting protein 1 (MAPK8IP1), mRNA.
70	15	ACTC1	GSTM1	Homo sapiens actin, alpha, cardiac muscle 1 (ACTC1), mRNA.	Homo sapiens glutathione S-transferase M1 (GSTM1), transcript variant 1, mRNA.
71	14	HLA-A29.1	AKR1C2	Homo sapiens major histocompatibility complex class I HLA-A29.1 (HLA-A29.1), mRNA. XM_001132736 XM_001132742 XM_001132747 XM_001132750 XM_001132753 XM_001132757 XM_001132760 XM_001132769 XM_001132772 XM_001132776 XM_001132779 XM_001132781 XM_001132783 XM_001132785 XM_001132794	Homo sapiens aldo-keto reductase family 1, member C2 (dihydrodiol dehydrogenase 2; bile acid binding protein; 3-alpha hydroxysteroid dehydrogenase, type III) (AKR1C2), transcript variant 1, mRNA. XM_943424 XM_943425 XM_943427
72	14	POMC	ALG1L	Homo sapiens proopiomelanocortin (POMC), transcript variant 1, mRNA.	Homo sapiens asparagine-linked glycosylation 1-like (ALG1L), mRNA.
73	14	AKR1C2	LOC100133565	Homo sapiens aldo-keto reductase family 1, member C2 (dihydrodiol dehydrogenase 2; bile acid binding protein; 3-alpha hydroxysteroid dehydrogenase, type III) (AKR1C2), transcript variant 1, mRNA. XM_943424 XM_943425 XM_943427	PREDICTED: Homo sapiens similar to hCG23738 (LOC100133565), mRNA.
74	14	GSDMC	CHURC1	Homo sapiens gasdermin C (GSDMC), mRNA.	Homo sapiens churchill domain containing 1 (CHURC1), mRNA.
75	13	FAM174A	AKR1C2	Homo sapiens family with sequence similarity 174, member A (FAM174A), mRNA.	Homo sapiens aldo-keto reductase family 1, member C2 (dihydrodiol dehydrogenase 2; bile acid binding protein; 3-alpha hydroxysteroid dehydrogenase, type III) (AKR1C2), transcript variant 1, mRNA. XM_943424 XM_943425 XM_943427
76	13	HS3ST2	COL3A1	Homo sapiens heparan sulfate (glucosamine) 3-O-sulfotransferase 2 (HS3ST2), mRNA.	Homo sapiens collagen, type III, alpha 1 (COL3A1), mRNA.
77	13	MAPK8IP1	COL3A1	Homo sapiens mitogen-activated protein kinase 8 interacting protein 1 (MAPK8IP1), mRNA.	Homo sapiens collagen, type III, alpha 1 (COL3A1), mRNA.
78	12	TTC25	AZIN1	Homo sapiens tetratricopeptide repeat domain 25 (TTC25), mRNA.	Homo sapiens antizyme inhibitor 1 (AZIN1), transcript variant 2, mRNA.
79	12	IGFN1	TTC25	Homo sapiens immunoglobulin-like and fibronectin type III domain containing 1 (IGFN1), mRNA.	Homo sapiens tetratricopeptide repeat domain 25 (TTC25), mRNA.
80	12	HS3ST2	CD79B	Homo sapiens heparan sulfate (glucosamine) 3-O-sulfotransferase 2 (HS3ST2), mRNA.	Homo sapiens CD79b molecule, immunoglobulin-associated beta (CD79B), transcript variant 3, mRNA.
81	11	ACTN3	HS3ST2	Homo sapiens actinin, alpha 3 (ACTN3), mRNA.	Homo sapiens heparan sulfate (glucosamine) 3-O-sulfotransferase 2 (HS3ST2), mRNA.
82	11	AKR1C4	AKR1C2	Homo sapiens aldo-keto reductase family 1, member C4 (chlordecone reductase; 3-alpha hydroxysteroid dehydrogenase, type I; dihydrodiol dehydrogenase 4) (AKR1C4), mRNA.	Homo sapiens aldo-keto reductase family 1, member C2 (dihydrodiol dehydrogenase 2; bile acid binding protein; 3-alpha hydroxysteroid dehydrogenase, type III) (AKR1C2), transcript variant 1, mRNA. XM_943424 XM_943425 XM_943427
83	11	HLA-DRB5	ALG1L	Homo sapiens major histocompatibility complex, class II, DR beta 5 (HLA-DRB5), mRNA.	Homo sapiens asparagine-linked glycosylation 1-like (ALG1L), mRNA.
84	11	LOC100008589	PTER	Homo sapiens 28S ribosomal RNA (LOC100008589), non-coding RNA.	Homo sapiens phosphotriesterase related (PTER), transcript variant 2, mRNA.
85	11	GSDMC	LOC647307	Homo sapiens gasdermin C (GSDMC), mRNA.	PREDICTED: Homo sapiens misc_RNA (LOC647307), miscRNA.
86	10	ACTN3	MIR1974	Homo sapiens actinin, alpha 3 (ACTN3), mRNA.	Homo sapiens microRNA 1974 (MIR1974), microRNA.
87	10	GSTM1	COL3A1	Homo sapiens glutathione S-transferase M1 (GSTM1), transcript variant 1, mRNA.	Homo sapiens collagen, type III, alpha 1 (COL3A1), mRNA.
88	10	HS3ST2	OLFM1	Homo sapiens heparan sulfate (glucosamine) 3-O-sulfotransferase 2 (HS3ST2), mRNA.	Homo sapiens olfactomedin 1 (OLFM1), transcript variant 2, mRNA.

89	10	QPRT	GSTM1	Homo sapiens quinolinate phosphoribosyltransferase (QPRT), mRNA.	Homo sapiens glutathione S-transferase M1 (GSTM1), transcript variant 1, mRNA.
90	10	PRODH	PTER	Homo sapiens proline dehydrogenase (oxidase) 1 (PRODH), nuclear gene encoding mitochondrial protein, mRNA.	Homo sapiens phosphotriesterase related (PTER), transcript variant 2, mRNA.
91	9	ACTC1	TTC25	Homo sapiens actin, alpha, cardiac muscle 1 (ACTC1), mRNA.	Homo sapiens tetratricopeptide repeat domain 25 (TTC25), mRNA.
92	9	CCL2	DKFZP564J102	Homo sapiens chemokine (C-C motif) ligand 2 (CCL2), mRNA.	Homo sapiens DKFZP564J102 protein (DKFZP564J102), transcript variant 1, mRNA.
93	9	QPRT	CCL2	Homo sapiens quinolinate phosphoribosyltransferase (QPRT), mRNA.	Homo sapiens chemokine (C-C motif) ligand 2 (CCL2), mRNA.
94	8	AKR1C4	EGR1	Homo sapiens aldo-keto reductase family 1, member C4 (chlordecone reductase; 3-alpha hydroxysteroid dehydrogenase, type I; dihydrodiol dehydrogenase 4) (AKR1C4), mRNA.	Homo sapiens early growth response 1 (EGR1), mRNA.
95	8	ACTN3	CRYM	Homo sapiens actinin, alpha 3 (ACTN3), mRNA.	Homo sapiens crystallin, mu (CRYM), transcript variant 1, mRNA.
96	8	HMGCS2	LOC100008589	Homo sapiens 3-hydroxy-3-methylglutaryl-Coenzyme A synthase 2 (mitochondrial) (HMGCS2), nuclear gene encoding mitochondrial protein, mRNA.	Homo sapiens 28S ribosomal RNA (LOC100008589), non-coding RNA.
97	7	HLA-DRB5	SPINK1	Homo sapiens major histocompatibility complex, class II, DR beta 5 (HLA-DRB5), mRNA.	Homo sapiens serine peptidase inhibitor, Kazal type 1 (SPINK1), mRNA.
98	7	AKR1C4	FOSB	Homo sapiens aldo-keto reductase family 1, member C4 (chlordecone reductase; 3-alpha hydroxysteroid dehydrogenase, type I; dihydrodiol dehydrogenase 4) (AKR1C4), mRNA.	Homo sapiens FBJ murine osteosarcoma viral oncogene homolog B (FOSB), mRNA.
99	6	FRZB	MFF	Homo sapiens frizzled-related protein (FRZB), mRNA.	Homo sapiens mitochondrial fission factor (MFF), nuclear gene encoding mitochondrial protein, mRNA.
100	5	AKR1C4	NDUFAF1	Homo sapiens aldo-keto reductase family 1, member C4 (chlordecone reductase; 3-alpha hydroxysteroid dehydrogenase, type I; dihydrodiol dehydrogenase 4) (AKR1C4), mRNA.	Homo sapiens NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, assembly factor 1 (NDUFAF1), mRNA.
101	5	HBG1	ALG1L	Homo sapiens hemoglobin, gamma A (HBG1), mRNA.	Homo sapiens asparagine-linked glycosylation 1-like (ALG1L), mRNA.
102	5	TINP1	IL8	Homo sapiens TGF beta-inducible nuclear protein 1 (TINP1), mRNA.	Homo sapiens interleukin 8 (IL8), mRNA.
103	5	MYF6	IL8	Homo sapiens myogenic factor 6 (herculin) (MYF6), mRNA.	Homo sapiens interleukin 8 (IL8), mRNA.
104	4	THBS4	MYF6	Homo sapiens thrombospondin 4 (THBS4), mRNA.	Homo sapiens myogenic factor 6 (herculin) (MYF6), mRNA.
105	4	IL32	HMGCS2	Homo sapiens interleukin 32 (IL32), transcript variant 7, mRNA.	Homo sapiens 3-hydroxy-3-methylglutaryl-Coenzyme A synthase 2 (mitochondrial) (HMGCS2), nuclear gene encoding mitochondrial protein, mRNA.
106	4	AKR1C4	ATP2A1	Homo sapiens aldo-keto reductase family 1, member C4 (chlordecone reductase; 3-alpha hydroxysteroid dehydrogenase, type I; dihydrodiol dehydrogenase 4) (AKR1C4), mRNA.	Homo sapiens ATPase, Ca++ transporting, cardiac muscle, fast twitch 1 (ATP2A1), transcript variant a, mRNA.
107	4	ATP2A1	IL32	Homo sapiens ATPase, Ca++ transporting, cardiac muscle, fast twitch 1 (ATP2A1), transcript variant a, mRNA.	Homo sapiens interleukin 32 (IL32), transcript variant 7, mRNA.
108	3	SPINK1	HBG1	Homo sapiens serine peptidase inhibitor, Kazal type 1 (SPINK1), mRNA.	Homo sapiens hemoglobin, gamma A (HBG1), mRNA.
109	3	THBS4	TINP1	Homo sapiens thrombospondin 4 (THBS4), mRNA.	Homo sapiens TGF beta-inducible nuclear protein 1 (TINP1), mRNA.
110	2	ITGB1BP1	IL32	Homo sapiens integrin beta 1 binding protein 1 (ITGB1BP1), transcript variant 1, mRNA.	Homo sapiens interleukin 32 (IL32), transcript variant 7, mRNA.
111	2	HS.137971	MT1E	AV737943 CB Homo sapiens cDNA clone CBDAEG06 5, mRNA sequence	Homo sapiens metallothionein 1E (MT1E), mRNA.
112	2	SHISA2	AKR1C2	Homo sapiens shisa homolog 2 (Xenopus laevis) (SHISA2), mRNA.	Homo sapiens aldo-keto reductase family 1, member C2 (dihydrodiol dehydrogenase 2; bile acid binding protein; 3-alpha hydroxysteroid dehydrogenase, type III) (AKR1C2), transcript variant 1, mRNA. XM_943424 XM_943425 XM_943427
113	2	SLC25A25	AKR1C2	Homo sapiens solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 25 (SLC25A25), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA.	Homo sapiens aldo-keto reductase family 1, member C2 (dihydrodiol dehydrogenase 2; bile acid binding protein; 3-alpha hydroxysteroid dehydrogenase, type III) (AKR1C2), transcript variant 1, mRNA. XM_943424 XM_943425 XM_943427
114	1	NDUFAF1	SHISA2	Homo sapiens NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, assembly factor 1 (NDUFAF1), mRNA.	Homo sapiens shisa homolog 2 (Xenopus laevis) (SHISA2), mRNA.
115	1	EGR1	SLC25A25	Homo sapiens early growth response 1 (EGR1), mRNA.	Homo sapiens solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 25 (SLC25A25), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA.
116	1	BEGAIN	SHISA2	Homo sapiens brain-enriched guanylate kinase-associated homolog (rat) (BEGAIN), mRNA.	Homo sapiens shisa homolog 2 (Xenopus laevis) (SHISA2), mRNA.
117	1	EGR1	SLITRK4	Homo sapiens early growth response 1 (EGR1), mRNA.	Homo sapiens SLIT and NTRK-like family, member 4 (SLITRK4), mRNA.

Annex 4. Table of more central genes in the network Dataset2

	Gene	N interactions	Description
1	ALG1L	6	Homo sapiens asparagine-linked glycosylation 1-like (ALG1L), mRNA.
2	AKR1C2	6	Homo sapiens aldo-keto reductase family 1, member C2 (dihydrodiol dehydrogenase 2; bile acid binding protein; 3-alpha hydroxysteroid dehydrogenase, type III) (AKR1C2), transcript variant 1, mRNA. XM_943424 XM_943425 XM_943427
3	HBG1	5	Homo sapiens hemoglobin, gamma A (HBG1), mRNA.
4	SLITRK4	4	Homo sapiens SLIT and NTRK-like family, member 4 (SLITRK4), mRNA.
5	LOC100008589	4	Homo sapiens 28S ribosomal RNA (LOC100008589), non-coding RNA.
6	GSTM1	4	Homo sapiens glutathione S-transferase M1 (GSTM1), transcript variant 1, mRNA.
7	ZP3	4	Homo sapiens zona pellucida glycoprotein 3 (sperm receptor) (ZP3), mRNA.
8	COL3A1	4	Homo sapiens collagen, type III, alpha 1 (COL3A1), mRNA.
9	SPINK1	4	Homo sapiens serine peptidase inhibitor, Kazal type 1 (SPINK1), mRNA.
10	PTER	4	Homo sapiens phosphotriesterase related (PTER), transcript variant 2, mRNA.
11	IL8	3	Homo sapiens interleukin 8 (IL8), mRNA.
12	SLC2A3	3	Homo sapiens solute carrier family 2 (facilitated glucose transporter), member 3 (SLC2A3), mRNA.
13	LOC100133565	3	PREDICTED: Homo sapiens similar to hCG23738 (LOC100133565), mRNA.
14	SLC25A25	3	Homo sapiens solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 25 (SLC25A25), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA.
15	AGTPBP1	3	Homo sapiens ATP/GTP binding protein 1 (AGTPBP1), mRNA.
16	CHURC1	3	Homo sapiens churchill domain containing 1 (CHURC1), mRNA.
17	LOC647307	3	PREDICTED: Homo sapiens misc_RNA (LOC647307), miscRNA.
18	SHISA2	3	Homo sapiens shisa homolog 2 (Xenopus laevis) (SHISA2), mRNA.
19	CD79B	2	Homo sapiens CD79b molecule, immunoglobulin-associated beta (CD79B), transcript variant 3, mRNA.
20	HLA-DRB5	2	Homo sapiens major histocompatibility complex, class II, DR beta 5 (HLA-DRB5), mRNA.
21	HMGCS2	2	Homo sapiens 3-hydroxy-3-methylglutaryl-Coenzyme A synthase 2 (mitochondrial) (HMGCS2), nuclear gene encoding mitochondrial protein, mRNA.
22	FOSB	2	Homo sapiens FBJ murine osteosarcoma viral oncogene homolog B (FOSB), mRNA.
23	SDHALP1	2	Homo sapiens succinate dehydrogenase complex, subunit A, flavoprotein pseudogene 1 (SDHALP1) on chromosome 3.
24	TINP1	2	Homo sapiens TGF beta-inducible nuclear protein 1 (TINP1), mRNA.
25	AZIN1	2	Homo sapiens antizyme inhibitor 1 (AZIN1), transcript variant 2, mRNA.
26	ATP2A1	2	Homo sapiens ATPase, Ca++ transporting, cardiac muscle, fast twitch 1 (ATP2A1), transcript variant a, mRNA.
27	MAPK8IP1	2	Homo sapiens mitogen-activated protein kinase 8 interacting protein 1 (MAPK8IP1), mRNA.
28	MIR1974	2	Homo sapiens microRNA 1974 (MIR1974), microRNA.
29	TTC25	2	Homo sapiens tetratricopeptide repeat domain 25 (TTC25), mRNA.
30	IL32	2	Homo sapiens interleukin 32 (IL32), transcript variant 7, mRNA.
31	FRZB	1	Homo sapiens frizzled-related protein (FRZB), mRNA.
32	LOC649143	1	PREDICTED: Homo sapiens similar to HLA class II histocompatibility antigen, DRB1-9 beta chain precursor (MHC class I antigen DRB1*9) (DR-9) (DR9), transcript variant 2 (LOC649143), mRNA.
33	AURKA	1	Homo sapiens aurora kinase A (AURKA), transcript variant 5, mRNA.
34	FAM174A	1	Homo sapiens family with sequence similarity 174, member A (FAM174A), mRNA.
35	LOC100134364	1	PREDICTED: Homo sapiens hypothetical protein LOC100134364 (LOC100134364), mRNA.
36	HLA-A29.1	1	Homo sapiens major histocompatibility complex class I HLA-A29.1 (HLA-A29.1), mRNA. XM_001132736 XM_001132742 XM_001132747 XM_001132750 XM_001132753 XM_001132757 XM_001132760 XM_001132769 XM_001132772 XM_001132776 XM_001132779 XM_001132781 XM_001132783 XM_001132785 XM_001132794
37	TNNI1	1	Homo sapiens troponin I type 1 (skeletal, slow) (TNNI1), mRNA.
38	MYBPH	1	Homo sapiens myosin binding protein H (MYBPH), mRNA.

39	RPL14	1	Homo sapiens ribosomal protein L14 (RPL14), transcript variant 1, mRNA.
40	CALML6	1	Homo sapiens calmodulin-like 6 (CALML6), mRNA.
41	LOC100132394	1	PREDICTED: Homo sapiens hypothetical protein LOC100132394 (LOC100132394), mRNA.
42	TIMM10	1	Homo sapiens translocase of inner mitochondrial membrane 10 homolog (yeast) (TIMM10), nuclear gene encoding mitochondrial protein, mRNA.
43	GSDMC	1	Homo sapiens gasdermin C (GSDMC), mRNA.
44	LOC100131138	1	PREDICTED: Homo sapiens similar to hCG2040918 (LOC100131138), mRNA.
45	HS3ST2	1	Homo sapiens heparan sulfate (glucosamine) 3-O-sulfotransferase 2 (HS3ST2), mRNA.
46	OLFM1	1	Homo sapiens olfactomedin 1 (OLFM1), transcript variant 2, mRNA.
47	DKFZP564J102	1	Homo sapiens DKFZP564J102 protein (DKFZP564J102), transcript variant 1, mRNA.
48	CCL2	1	Homo sapiens chemokine (C-C motif) ligand 2 (CCL2), mRNA.
49	EGR1	1	Homo sapiens early growth response 1 (EGR1), mRNA.
50	CRYM	1	Homo sapiens crystallin, mu (CRYM), transcript variant 1, mRNA.
51	MFF	1	Homo sapiens mitochondrial fission factor (MFF), nuclear gene encoding mitochondrial protein, mRNA.
52	NDUFAF1	1	Homo sapiens NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, assembly factor 1 (NDUFAF1), mRNA.
53	MYF6	1	Homo sapiens myogenic factor 6 (herculin) (MYF6), mRNA.
54	MT1E	1	Homo sapiens metallothionein 1E (MT1E), mRNA.