Life Sciences Engineering: Genome to Function

Project 2

Manon Dorster, Alexandra-Elena Preda, Viola Renne, Jiayi Shen

February 29, 2024

Question 1

By deleting the YAF9 gene in S. Cerevisiae, we would be able to compare the viability of the yeast and determine whether YAF9 is an essential or non-essential gene. It is important to remember that RNAi-mediated inactivation does not exist in S. cerevisiae, thus we would use a deletion cassette inserted in the genome via homologous recombination to delete the YAF9 gene or CRISPR/Cas9 to induce a complete deletion through error-prone non-homologous end joining. An advantage of CRISPR/Cas9 is that it rarely has off-target effects. By comparing the viability of the YAF9 knock-out strain with the wild-type strain, we could determine if YAF9 is an essential gene or not. Furthermore, assessing the fitness of the deletion strain would enable us to determine whether the YAF9 gene is essential for UV resistance or whether it is just partially contributing to it.

In order to determine if YAF9 is involved in a DNA repair pathway, we could detect synthetic lethal interactions with genes known to be involved in DNA repair mechanisms. We would combine the YAF9 deletion mutation with mutants in known DNA repair genes and observe the viability of the resulting double mutants after UV-induced double strand DNA breaks. The presence of synthetic lethality would suggest YAF9's involvement in double strand break DNA repair pathways. Viability or lack thereof in double mutants would also provide insights into whether YAF9 acts alone or interacts with other genes for UV resistance.

To visualize the localization of the gene product of YAF9, we could genetically engineer a fusion of the YAF9 protein with a fluorescent protein (for example GFP). In this way, we could determine in which cellular compartment the gene product is expressed.

To better understand in which pathways the YAF9 protein product is involved, we could use proximity-based labeling or yeast 2-hybrid system to search and see which proteins interact with the YAF9 protein.

It would also be interesting to quantify the expression of the YAF9 protein product. We could overexpress the YAF9 gene in S. Cerevisiae with plasmid transfection using yeast artificial chromosomes (YACs) and assess whether increased expression enhances UV resistance. This could help establish a direct correlation between YAF9 protein levels and UV resistance.

Question 2

a) The experiment shown in the graph is a synthetic interaction experiment. The goal of this experiment is to understand if there is an interaction between two genes. This is performed by looking at the phenotype of the wild type, the single mutant (yeasts with only gene A deleted and yeasts with only gene B deleted), and the double mutant (yeasts with both genes deleted).

b) A drug-blocking gene A should have a similar effect as the deletion of gene A. For this reason, we would expect, in cells lacking gene B and treated with a drug blocking gene A, a similar phenotype as double mutants (yellow curve of the given graph). It means that by treating cells lacking gene B with a drug-blocking gene A, the fitness of the cells increases.

c) Yeasts having gene A deleted have a greater fitness than yeasts having gene B deleted. If there were no interaction between the two genes, the fitness of the double mutants would be lower than the fitness of both single mutants. Since the double mutant has a less severe phenotype than expected for combining the two single mutants, it means that there is an interaction between the two genes. In particular, this is a positive genetic interaction. Since the fitness of the double mutant is similar to the fitness of yeasts having only gene A deleted, gene A is masking the loss of fitness phenotype of gene B. For this reason, the two genes are in the same biological pathway, and the product of gene A may be acting before the one of gene B. Since the pathway may involve different cellular compartments, we cannot conclude that the two genes are in the same cellular compartment.

d) The idea is to search for a compound that is synthetic lethal (or sick) with the liver cancer cells, but it is not synthetic lethal (or sick) with cells that normally express gene B. We would consider two cell lines: healthy liver cells (that normally express gene B) and liver cancer cells (that over-express gene B). We would test different compounds to find one that is synthetic lethal with cancer cells, but not with healthy cells. It could also be possible to search for a gene that is synthetic lethal with cancer cells and not with healthy cells. We could then obtain its synthetic lethal profile with a synthetic lethal screen. Lastly, we would find a drug that has a similar synthetic lethal profile to that gene. This second approach would be useful also to understand what is the target of the selected drug. It is a chemical genomics experiment.

Question 3

a) The image represents a chemical genomics screen. The goal is to create functional annotations for chemical compound libraries, specifically focusing on the NCI oncology library in this instance. The principle of this type of analysis is that gene deletion alleles that show chemical genetic interactions with a compound should also be synthetically lethal or sick with a mutation in the compound target gene. For this experiment, the first step is a synthetic lethality screen, to obtain the genetic interaction profiles of genes. The second step is to find the chemical-genetic interaction profiles of different compounds. At this point, a comparison of chemical genetic profile to the compendium of genetic interaction profile is done. Then a genetic interaction network is constructed, where nodes are genes, and edges connect genes that share similar genetic interaction profiles.

Gene sharing similar patterns are proximal to each other. To divide genes in 17 distinct biological processes, a clustering is performed (each cluster has a different colour). The cluster is performed using functional enrichment annotations. Then it is possible to map compounds onto the genetic interaction network based on the most similar genetic interaction profile to the compound 1.

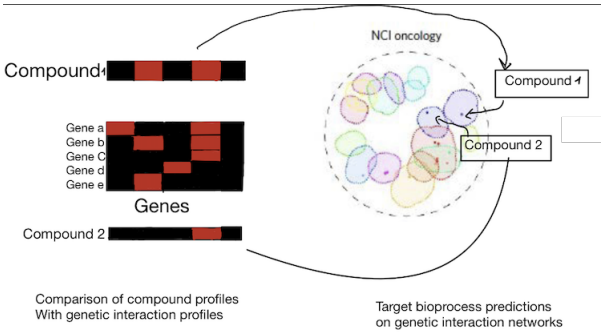

Figure 1: Chemical genomics screening - Compare compound chemical-genetic profile - Adapted from [7].

b) Colored nodes represent chemical compounds. Each compound was placed on the map at the position of the gene with the most similar genetic interaction profile to it. In this particular case, dot 1 is positioned in the blue area that represents the biological process of protein degradation. Dot 2 is positioned in the purple area that is associated with the biological process of respiration, oxidative phosphorylation and mitochondria targeting.

c) For the dot 1, which is a compound in the biological process of protein degradation, we found bortezomib (NSC681239). Bortezomib, a proteasome inhibitor, can induce apoptosis by accumulating proteins in cells. The proteasome plays a crucial role in protein degradation, recognizing and selectively unfolding ubiquitinated proteins for controlled cellular protein levels [5]. We also checked the Supplementary Information of the paper seen in class [4], and in the Supplementary Dataset 19, we found the confirmation that Bortezomib (Cond-011318, NSC681239) is involved in the bioprocess of protein degradation.

For the dot 2, which is a compound in the biological process of respiration, oxidative phosphorylation and mitochondrial targeting, we found the compound Nilotinib (NSC747599). Nilotinib is a tyrosine kinase inhibitor compound used in the treatment of certain types of cancer. It acts by blocking the protein kinase called Bcr-Abl kinase which is produced by leukaemia cells that have the Philadelphia chromosome, and causes them to multiply uncontrollably[1]. We found this compound by checking the Supplementary Information of the paper seen in class [4], and in the Supplementary Dataset 19 we found this compound (Cond-011342, NSC747599).

Question 4

We need to target S. rosetta C-type lectin gene since it is necessary for multicellular rosette development. To create a conditional mutant allele of C-type lectin gene, one viable approach involves employing the Tet-Off system. This system enables temporal

control over gene expression. The Tetracycline-controlled transactivator (tTA) is responsible for modulating gene expression. In the absence of the drug doxycycline (Dox), tTA dimers selectively bind to tetO sequences, thereby initiating the transcription of the target gene. Conversely, in the presence of Dox, tTA undergoes a conformational change, rendering it incapable of binding to tetO sequences, and inhibiting the expression of the gene. To introduce the system, we would use two different constructs. The first construct would contain a promoter with Tet operator sequences and the C-type lectin gene. To remove the endogenous gene, we could produce two double strand breaks upstream and downstream the gene, and then use homologous recombination to insert the C-type lectin gene with the Tet operator sequences. The second construct would contain Tet-Off transactivator gene (tTA). Since we need to insert two plasmids, we wold use sequential transformation. For each transformation, it would be useful to use a selectable marker to identify cells that have successfully incorporated the transgene. This approach allows us to precisely control the expression of the C-type lectin gene in response to doxycycline.

A second method would be targeting the product of the C-type lectin gene. This can be done using on of the two conditional degron-based technologies, the temperature-sensitive (ts) degron and the auxin-inducible degron (AID) systems, which allow us to selectively deplete Proteins of Interest (POIs) under specific conditions. We would use the AID system. Auxin is a family of plant hormones that regulate gene expression across various stages of growth and development in plants. We would use two constructs. The first plasmid contains DNA sequence encoding the protein of interest (C-type lectin gene in this case) fused with the AID sequence. We can introduce the AID-tagged C-type lectin gene construct into the genome, using homologous recombination that can also remove the endogenous gene (as described in the previous point). The second plasmid encodes the TIR1 protein, which is a component of the SCF (Skp1-Cullin-F-box) E3 ubiquitin ligase complex. TIR1 recognizes the AID sequence and facilitates the ubiquitination and degradation of the AID-tagged protein in the presence of auxin. The degron (AID) fused with the C-type lectin protein, is recognized and poly-ubiquitylated by the SCF TIR1 E3 ubiquitin ligase in the presence of auxin. Since we need to insert two plasmids, we wold use sequential transformation. For each transformation, it would be useful to use a selectable marker to identify cells that have successfully incorporated the transgene. New researchs show that it is possible also to use a All-in-One Vector [6].

Question 5

Step 2: Identifying potential gene pairs, is crucial, and the proposed approaches (genomic analysis and pathway analysis) are appropriate. Determining frequently co-mutated or co-expressed gene pairs is essential for identifying potential therapeutic targets. The use of bioinformatics is a standard and cost-effective approach, and it is imperative in this case. Furthermore, pathway analysis is important to understand how these genes might be involved in the disease mechanism, in order to target specific molecules within the pathway. It is however unclear why the JAK-STAT and PI3K-AKT pathways are mentioned. The BCR-ABL pathway is key in CML, which is not very clearly pin-pointed in the chatGPT answer, and there are perhaps other pathways than JAK-STAT and PI3K-AKT to take into consideration as well.

Step 3: It is important to choose an appropriate cell line. The ones suggested by ChatGPT seem correct since they are cell lines obtained from CML in blast crisis (BC) phase, which is the most advanced phase in the CML disease progression [2]. There are several other cell lines which can be used e.g. NALM-1,KCL-22 [2]. Gene silencing and overexpression are also standard in experimental design and they are important, the use of RNAi and CRISP/Cas9 is indeed needed for overexpression and silencing. Silencing or overexpressing the identified gene pairs in a pairwise individual manner is an accurate approach in our case. It is however important to keep in mind the potential off-target effects when using RNAi, which is not mentioned in step 3 of the ChatGPT response.

Step 4: In first point of this step, it would be helpful to know which information each assay provides. MTT assay is used to measure cell proliferation and viability, and could thus aid in discriminating which targeted gene pairs do not work to kill CML cells. However, this technique would probably not be the most precise to detect cell apoptosis. CellTiter-Glo assay measure the concentration of ATP in cells, indicating metabolic activity. As in the MTT assay, CellTiter-Glo would be an adequate technique to detect ineffective drug combinations but would not indicate which cells are undergoing apoptosis. Hence, flow cytometry is the best choice among the proposed assays to detect apoptotic cells. It is important to mention that we would use the Annexin V flow cytometry assay for detection of apoptotic cells, which should be included in the experimental design protocol. The second step proposed by ChatGPT is correct. After identifying gene pairs that are synthetic lethal, we should assess the viability, proliferation, and apoptosis of cells targeted with single drugs and combination of two drugs. It is important to consider also non-cancer cells to evaluate possible side effects. In addition, it is important to keep in mind that cancer cells may develop drug resistance due to their highly mutational nature.

Alternative Strategy: Another possible approach to identify gene pairs is synthetic lethal screen using CRISPR double-knockout (CDKO) system 2. The aim is to identify two drugs that combined can be used for cancer therapy. For the pairwise sgRNA expression system, we could use a dual-promoter system. The two distinct promoters are human and mouse U6, to limit the homologous region. The first step would be designing sgRNAs that target the genes of interest. In particular, we can design a sgRNA library focused on genes relevant to potential CML drug therapies. In the study seen in the lecture [3], 207 genes plus controls were tested. We would obtain a dual sgRNA library, by ligating together the mU6-driven and hU6-driven single-sgRNA libraries. Then we would lentiviral vectors to deliver them. We are able to perform a pooled screen, by using as barcode the sgRNAs. The frequency of the dual sgRNAs, can be optained by deep sequencing. We would look for gene pairs with specific growth phenotypes and genetic interactions. The gene pair with a synergistic action could then be validated in vitro, using Cas9-expressing K562 cells. We would then evaluate the experiment's findings for their potential relevance to developing effective combination drug therapies for CML.

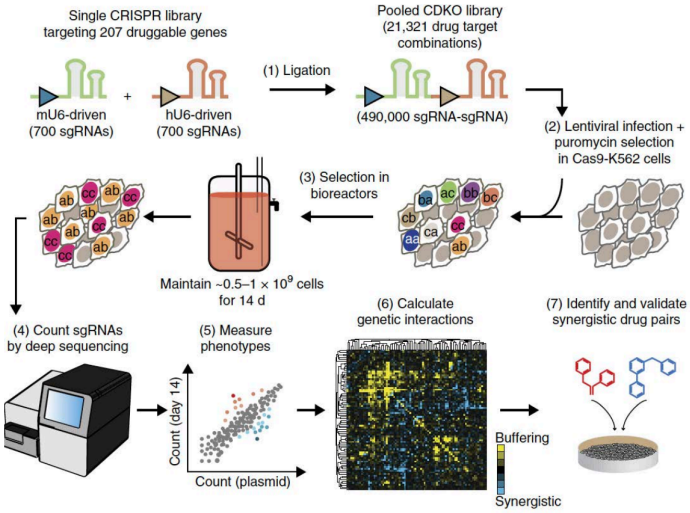

Figure 2: Pipeline for an alternative strategy

References

[1] Jean-Yves Blay and Margaret von Mehren. Nilotinib: A novel, selective tyrosine kinase inhibitor. *Seminars in Oncology*, 38(01):S3–S9, April 2011.

[2] Holyoake TL Clarke CJ. Preclinical approaches in chronic myeloid leukemia: from cells to systems. *Experimental Hematology*, 47:13–24, March 2017.

[3] Kyuho Han, Edwin E. Jeng, Gaelen T. Hess, David W. Morgens, Amy Li, and Michael C. Bassik. Synergistic drug combinations for cancer identified in a crispr screen for pairwise genetic interactions. *Nature Biotechnology*, 35(5):463–74, May 2017.

[4] Jeff S Piotrowski, Sheena C Li, Raamesh Deshpande, Scott W Simpkins, Justin Nelson, Yoko Yashiroda, Jacqueline M Barber, Hamid Safizadeh, Erin Wilson, Hiroki Okada, Abraham A Gebre, Karen Kubo, Nikko P Torres, Marissa A LeBlanc, Kerry Andrusiak, Reika Okamoto, Mami Yoshimura, Eva DeRango-Adem, Jolanda van Leeuwen, Katsuhiko Shirahige, Anastasia Baryshnikova, Grant W Brown, Hiroyuki Hirano, Michael Costanzo, Brenda Andrews, Yoshikazu Ohya, Hiroyuki Osada, Minoru Yoshida, Chad L Myers, and Charles Boone. Functional annotation of chemical libraries across diverse biological processes. *Nature Chemical Biology*, 13(9):982–993, July 2017.

[5] Cancer Research UK. Bortezomib (velcade). `https://www.cancerresearchuk.org/about-cancer/treatment/drugs/bortezomib#:~:text=Bortezomib%20is%20a%20type%20of%20targeted%20cancer%20drug%20called%20a,build%20up%20inside%20the%20cell.` Accessed November 12, 2023.

[6] Aisha Yesbolatova, Yuichiro Saito, and Masato T. Kanemaki. Constructing auxin-inducible degron mutants using an all-in-one vector. *Pharmaceuticals*, 13(5):103, May 2020.

[7] Fu-lai Zhou, Sheena C. Li, Yue Zhu, Wan-jing Guo, Li-jun Shao, Justin Nelson, Scott Simpkins, and et al. Integrating yeast chemical genomics and mammalian cell pathway analysis. *Acta Pharmacologica Sinica*, 40(9):1245–1255, September 2019.