Sequencing of mutant genomes

Bioinformatics Practical – BC.7107

Rares Cristea1, Lionel Rohner2, Alain Rohrbasser1, Thibault Schowing1

1University of Fribourg, Fribourg, Switzerland

2University of Bern, Switzerland

**Working title:**

Identification of genomic suppressor mutations by high-throughput sequencing of yeast genomes

**Abstract:**

The discovery of S. cerevisiae has dramatically contributed to great advancement lead to the development of…

The budding Yeast Saccharomyces cerevisiae is a common organism used for genetics manipulation. With a genome with 16 chromosomes (haploid, Mat a or α) or 32 chromosomes (diploid). 99% of the genome is without introns, make this organism handy to manipulate. The short mating time and growth is also short.

Write about aim and main methods….

* Introduction to Quality Control and Remapping (LF)…
* SNP calling and VCF format (LMS,LF)…
* SNP effect and filtering variants
* Visualization with IGV and identification of mutants

Conclude…

* We found xxx potential candidates for temp-sensitive phenotype
* Implication for further studies…

**Key words:** SNP filtering, variant comparison, tom1, riboprotein aggregation, s. cerevisea

**Abbreviations:**

BAM, SAM, VCF, TOM1, RPL, SNP

**Introduction:**

The budding yeast Saccharomyces cerevisiae (S.cerevisiae) is a single-celled lower eukaryote belonging to the kingdom of fungi. Ever since its discovery, S.cerevisae has nourished human advancements in the field of fermented food products, alcoholic beverages (e.g. beer, which is the epony of S.cerevisae) and the production of biofuel.1 In addition to the contribution industrial fermentation, S.cerevisiae has become one of the most popular model organism for eukaryotic biology, due to its simple cellular architecture, cheap maintenance cost, fast growth, non-pathogenic nature \cite{perez-torrado\_opportunistic\_2016} , and homologies to human cells (e.g. ribosomes), which cannot be studied in prokaryotic model organism, such as E.coli \cite{botstein\_yeast\_2011}. In particular, the genetic analysis of S.cerevisiae has gained popularity in the scientific community since it was the first eukaryotic organism whose genome was fully sequenced. The haploid genome of S.cerevisiae consists of 16 linear chromosomes containing 6604 genes encoded within approximately 12 megabase-pairs (Mbp).1 The fact that genome of S.cerevisiae is quite small and almost completely void of intronic DNA, thus making it an ideal microorganism for the identification of mutations and single nucleotide polymorphisms.

Mating of two haploid yeast of opposite mating type (i.e. Mat a or $\alpha $) gives rise to diploid cells that possess 32 chromosomes.

This is a single eukaryotic organism with a division cycle of 90 minutes. Through the process of budding in which smaller daughter cells pinch, or bud, off the mother cell. S.cerevisiae forms colonies on agar plates in the laboratory in a few days with no special incubators required (best grow at 30°).

Tom1 and ribosome proteins in yeast….

The deletion of Tom1 has been associated with aggregation of ribosomal proteins, which results in a temperature-sensitive phenotype of S.cerevisiae that renders them incapable of growing at temperature exceeding 20°C.

The aim of this study is to screen several ∆tom1 S.cerevisiae strains for mutations in genes associated to tom1 (positive and negative regulators) and genes coding for riboproteins…

In this project, we present a number of potential candidates that might help deciphering the temperature-sensitive phenotype of ∆tom1 S.cerevisiae

**Materials and Methods**

**Strains:**

* ∆tom1 S.cerevisiae ??? on sait plus que ca ?

**Genome Sequencing:**

* Illumina … Filtered with Trimmomatic (get rid of shitty reads)
* Write about reference R92…
* Sam to Bam conversion
* Annotation Tools

**Identification of variants:**

* Samtools mpileup or bcftools
* SNP filteration >> snpEff
* Only take variants that are found in less than 2 strains
* Visualization > gensearchNGS and IGV

**References:**

1. Belda, I., Ruiz, J., Santos, A., Van Wyk, N. & Pretorius, I. S. Saccharomyces cerevisiae. *Trends Genet.* S0168952519301829 (2019) doi:10.1016/j.tig.2019.08.009.

2. Botstein, D., Chervitz, S. A. & Cherry, J. M. Yeast as a Model Organism. 4 (2011).

The budding Yeast Saccharomyces cerevisiae is a common organism used for genetics manipulation. This organism is well conserved among the eukaryote and can be used correlate with human pathways. With a genome with 16 chromosomes (haploid, Mat a or α) or 32 chromosomes (diploid). 99% of the genome is without introns, make this organism handy to manipulate. 12 million bases pair and contains tween 5 800 à 6 572 gènes1,2. The homology with human is estimate to 23%, which is a good candidate for preliminary studies regarding human pathways. The short mating time and growth is also short. Thus, the identification of potential mutant is grandly enhanced. This is a single eukaryotic organism with a division cycle of 90 minutes. Through the process of budding in which smaller daughter cells pinch, or bud, off the mother cell. Due to the microscopic size (~5 microM, between bacteria and human cell size) and simple growth environment, yeasts are inexpensive and easy to grow in silico. Saccharomyces cerevisiae is also no-pathogen, and forms colonies on agar plates in the laboratory in a few days with no special incubators required (best grow at 30°).

Gene TOM ou autre polymorphisme trouvé à décrire dans le fichier vcf.

Arabidopsis Thaliana

Les deux références.

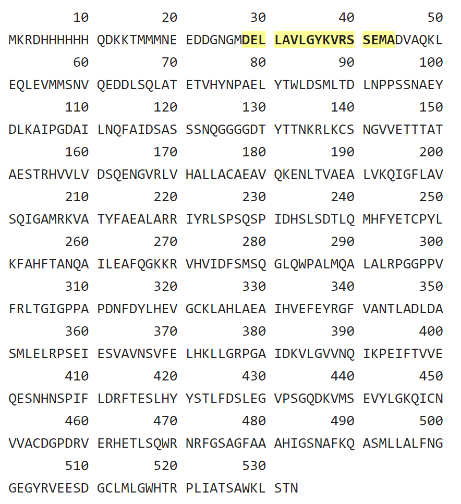
1. www.yeastgenome.org
2. Herskowitz I. Life cycle of the budding yeast Saccharomyces cerevisiae. *Microbiol Rev*. 1988;52(4):536–553.

ATH intro

The plant Arabidopsis thaliana is a genetic model worldwide used in plant biology since 1995, when it has been promoted as model for molecular genetics. The genome is entirely sequenced in 200. ATH is a diploid organism of 114,5 to 125 million base pairs within 5 chromosomes (haploid). The germination to mature seed is done about 6 weeks and easy to cultivate in restricted space and produce a lot of seed. A wide range of mutants are already available and it growth from year to another through multinational research community of academic, government and industry laboratories. The importance of ATH is crucial and invaluable resources to fight the loss of crops dur to plants diseases.

**GAI gene2**

The GAI gene regulates the sensitivity of the plant to gibberellin*, apical growth hormone of the plants*, tetracyclic diterpenoid growth factors that are essential regulators of stem elongation and other plant developmental processes3. The main mutation involved a deletion of a 17 amino acid segment at the terminal region of this DNA-binding transcription factor activity and causes a dwarf phenotype. The GAI protein as normally a length of 533 AA and is normally located in the nucleus. The deleted segment is shown in yellow.



The DELLA protein GAI comes from the GRAS family protein 3 and located on the chromosome 1, locus:2006747 AT1G14920. Call DELLA because of the 5 AA deleted.

Transcriptional regulator that acts as a repressor of the gibberellin (GA) signalling pathway. Transcription coactivator of the zinc finger transcription factors GAF1/IDD2 and ENY/IDD1 in regulation of gibberellin homeostasis and signalling. No effect of the BOI proteins on its stability. Probably acts by participating in large multiprotein complexes that repress transcription of GA-inducible genes. Positively regulates XERICO expression. In contrast to RGA, it is less sensitive to GA. Its activity is probably regulated by other phytohormones such as auxin and ethylene

1. https://www.ncbi.nlm.nih.gov/genome/4
2. Peng J, Carol P, Richards DE, et al. The Arabidopsis GAI gene defines a signaling pathway that negatively regulates gibberellin responses. *Genes Dev*. 1997;11(23):3194–3205. doi:10.1101/gad.11.23.3194
3. Gibberellins: perception, transduction and responses. *Hooley R Plant Mol Biol. 1994 Dec; 26(5):1529-55.*

Lactobacillus helveticus intro

The lactobacillus helveticus D75 has been isolated from health intestinal tract of children. It’s a circular DNA with 2,05 million base pairs with differences between strains and has been sequenced in 2012. This is then ancient biotechnology application. This is a Gramm+ bacteria is worldwide employed in the production of a wide range of fermented milk, meat, and plant products. used in the dairy industry as a starter or an adjunct culture for extra-hard, semi-hard and Swiss type cheese manufacture as well as for some types of fermented milk processing. By releasing intracellular enzymes, autolysis of L. helveticus has many impacts on cheese ripening and its extent is known to be strain dependent.

References

From my ass’s hairs and bubblebut my imaginary friend whom loves cheese.