BIOS 247 Mini-Course, Summer 2024

Whole-Genome Sequencing: From Yeast to Fruit Flies

**Exercise 3: Looking or copy number variation in IGV**

In the last exercise, we curated SNPs called by GATK in IGV for five mutants.

* Two of the five had mutations in the gene AFT1, a transcription factor involved in iron-sensing. These mutations may help the cells sequester iron in an environment with alternating H2O2 stress, which can react with iron.
* Another of the five had a heterozygous SNP, indicating it is a diploid. Its diploidy gives it a fitness advantage, so the SNP might not be relevant.
* The last two mutants have no recurrent SNPs, so we still don’t know why they’re adaptive. Today we’ll look for copy number variation to solve this mystery.

Directions:

1. Open the IGV desktop user interface. In the upper left dropdown menu, choose the reference genome: “S. cerevisiae (sacCer3).”
2. To load the five genomes you downloaded from Dropbox, go to File -> Load from File. Upload the five .bam files.
3. Download the tables from the CNV\_candidates subfolder in the data folder in [GitHub](https://github.com/clare-abreu/WGS_Bios247/tree/main/data/). Note that these tables are long, but many rows belong to the same area within a chromosome. Scrolling through these areas is quick.
4. Look up the flagged locations. Do you see any copy number variation that isn’t common to all mutants?

Questions:

1. You should find a section of three genes that have been amplified in two of the mutants compared to the other three. Look up these genes in [SGD](https://www.yeastgenome.org/). Can you explain why this amplification might happen in mutants evolved in environments with high salt stress?
2. Go back to the barcode region in Benchling. A promoter from a certain gene was inserted twice into this region to trigger Cre recombination, which allows the barcode to be inserted. What gene does this promoter correspond to? Look up its location in [SGD](https://www.yeastgenome.org/) and find it in IGV. What does the coverage of this region look like? Can you explain the pattern of SNPs there?