Unit 8

SAJUNG YUN

10 months ago

SGD Overview page

Reply | Quote | Email Author

PAMELA PAN

1 month ago

RE: SGD Overview page

The SGD Overview page is very similar to the gene overview page on NCBI. As stated in the lecture notes, some other information that may be found on this page include the translated protein, gene ontology, phenotypes of the gene, and data from expression studies. The overview page gives a summary of each, and by clicking on their individual tabs, it further expands into that category. For example on the Protein tab, there is more information from experimental data, protein domains and their classification, the amino acid sequence and composition, and identifiers for other databases.

For the same gene as the lecture notes, SEC1/YDR164C, we can see it is located on chromosome IV at position 784215..782041. It translates into a SNARE-binding protein that is 724 amino acids long and weighs 83479.1 Da. The expressed protein is localized to the plasma membrane and is involved in regulating vesicle fusion and docking during exocytosis. Based on experimental results, the protein has a half-life of 10.6 hours.

https://www.yeastgenome.org/locus/S000002571

▲ Hide 1 reply

ALEX GILSON

1 month ago

RE: SGD Overview page

Hello Pamela, thank you for your information about SGD! I have used this database previously in an undergraduate synthetic biology class , where we were transforming S. cerrevisiae in the lab to be more resistant to alcohol or caffiene. It was interesting to see a database house all the information we needed for one specific organism. You're right to say that the page looks similar to the gene overview page at NCBI, but it is much more useful that NCBI should you only be focused on this species.

SAJUNG YUN

10 months ago

Why C. elegans?

Why C. elegans? What makes it a good study?

Reply

ALEX GILSON

1 month ago

RE: Why C. elegans?

COLLAPSE

C. elegans makes a great model organism for a multitude of reasons. Namely, its simple physiological form and transparency mean studying C. elegans with a microscope can allow one to glean much information on physiology, development, and individual cell fate within the organism while it is still alive. On the molecular level, it has a relatively small genome that can be used for comparative genomics and in clinical analysis of specific genes such as examining PARK1/alpha-synuclein and its effects in Parkinsons(https://www.news-medical.net/life-sciences/C-elegans-as-a-Model-Organism.aspx).

Dr. Maho Yokohama. (2020). *C. elegans as a Model Organism.* Retrieved from: https://www.news-medical.net/life-sciences/C-elegans-as-a-Model-Organism.aspx

Reply | Quote | Edit

▲ Hide 2 replies

SAJUNG YUN

4 years ago

UCE Existence

At first impression, why do you think UCEs exist?

Reply | Quote | Email Author

ALEX GILSON

1 month ago

RE: UCE Existence

At first impression when learning about UCE's, I thought it would make sense for them to exist in areas of the genome that had some sort of function that would be lethal should a mutation occur in that region. This could be tested by examining which UCEs existed between two species, and trying to elucidate the function of that region and why it made sense to be conserved in those two species.

An article I found from Taewoo Ryu et al (2012) confused me a bit, as they attempted to examine the existence of these regions between distant species. They found through comparison of 6 very different species that most of the ultraconserved elements associated with dveleopmental genes, and those essential for cell functions. For this reason, it would seem that UCE's are part of the genetic code in regions that cannot be mutated, for the interference of the responsibilities of those regions would result in lethality. The study also showed that UCEs are not only found in "higher modern eukaryotes".

Ryu, T., Seridi, L., & Ravasi, T. (2012). The evolution of ultraconserved elements with different phylogenetic origins. *BMC evolutionary biology*, *12*, 236. https://doi.org/10.1186/1471-2148-12-236

Reply | Quote | Edit

▲ Hide 1 reply

SAJUNG YUN

4 years ago

UCE Mutants

What would you imagine to be the population dynamics of the allele frequency of a UCE mutant?

Reply | Quote | Email Author

BENJAMIN TENNER

1 month ago

RE: UCE Mutants

As UCE mutants are extremely rare I would assume that the allele frequency of a particular mutant in a population is extremely small.

▲ Hide 1 reply

ALEX GILSON

1 month ago

RE: UCE Mutants

Hi Benjamin, I found a study that goes more into detail about the known polymorphisms in UCE's within humans. Interestingly, there are 112 mutations within UCE's in humans that are known to be associated with clinical phenotypes. In total, they found almost 30,000 different mutations, but note that all have a very low minor allele frequency. They exist, but are so few and far between that its hard to get a true read on their actual frequencies.

Habic, A., Mattick, J. S., Calin, G. A., Krese, R., Konc, J., & Kunej, T. (2019). Genetic Variations of Ultraconserved Elements in the Human Genome. *Omics : a journal of integrative biology*, *23*(11), 549–559. https://doi.org/10.1089/omi.2019.0156

Reply | Quote | Edit

Unit 10

SAJUNG YUN

4 years ago

Assembly Problems

What assembly problems are caused by: low read length, structural variation (inversions), sequencing errors, SNPs, repeat regions?

Reply

ALEX GILSON

1 month ago

RE: Assembly Problems

Tandem repeat regions present sequence assembly obstacles in most systems, due to the nature of their sequences. Should a region be full of tandem repeats, with shorter read length, differentiating which parts are from the same read/a different part of the same sequence becomes challenging. I found an article from Ole K Torresen et al (2019) that discusses this issue. The authors's main concern for these repeat regions was how they can affect multiple levels of analysis for the regions they reside in, and when not accounted for correctly, can lead to multiple low confidence reads/sequences that yield low confidence proteins in an area that may simply have a few repeats within it.

Tørresen, O. K., Star, B., Mier, P., Andrade-Navarro, M. A., Bateman, A., Jarnot, P., Gruca, A., Grynberg, M., Kajava, A. V., Promponas, V. J., Anisimova, M., Jakobsen, K. S., & Linke, D. (2019). Tandem repeats lead to sequence assembly errors and impose multi-level challenges for genome and protein databases. *Nucleic acids research*, *47*(21), 10994–11006. https://doi.org/10.1093/nar/gkz841

Reply Quote

Unit 11

SAJUNG YUN

4 years ago

Epigenetics

What exactly is epigenetics? How is epigenetics different from genetics?

Reply

ALEX GILSON

1 month ago

RE: Epigenetics

Epigenetics is the study of biological mechanisms of inheritance, involving the control of gene expression that does is not directly associated with the underlying DNA sequences of the genome, and these actions that are also reversible (Kovalchuk and Kovalchuk, 2012). Comparatively, geneitcs is a more broad category that is the study of the mechanism of inheritance in general and particularly interested in genes (Kovalchuk and Kovalchuk, 2012). The difference between the two areas of study lie in the fact that epigenetics is concerned with things such as histone modifications and chemical associations with DNA, rather than looking at its sequence.

Igor Kobalchuk & Olga Kovalchuk. (2012). "Epigenetics in Health and Disease". *Pearson Education, Inc.* Upper Saddle Rive, NJ 07458. ISBN 978-0-13-259708-1.

Reply Quote Edit

▲ Hide 2 replies

Unit 12

SAJUNG YUN

4 years ago

Paired-end in RNA-seq

COLLAPSE

For RNA-seq, paired-end sequencing is strongly preferred. Give the presence of introns, why might that be the case?

Reply | Quote | Email Author

ALEX GILSON

26 days ago

RE: Paired-end in RNA-seq

COLLAPSE

Paired-end reads are preferred over single-end for RNA-seq data due in part to the fact that paired-end reads can compensate for introns by examining the gaps/length between the ends seen by their alignment to a reference genome. According to an article from Freedman et al (2020), short paired-end sequencing provides more accurate accounting for expression estimation/ differential expression, as well as being a more cost effective solution compared to using longer single-end sequencing.

Freedman, A.H., Gaspar, J.M. & Sackton, T.B. (2020). Short paired-end reads trump long single-end reads for expression analysis. *BMC Bioinformatics* **21,** 149. https://doi.org/10.1186/s12859-020-3484-z

Reply | Quote | Edit

Unit 13

SAJUNG YUN

2 years ago

Single-cell RNA-seq

What are the (possible) useful applications for single-cell RNA-seq

Reply | Quote | Email Author

ALEX GILSON

20 days ago

RE: Single-cell RNA-seq

COLLAPSE

ScRNA-seq is useful in multiple different scenarios, and I was interested to see novel methods that have been published recently using this method. I found an article from Trung Nghia Vu et al. (2019) that used scRNA-seq to examine somatic cells for individual cell genetic mutations, which is so specific that I wasn't sure it warranted its own use. The authors of the paper used scRNA-seq to determine whether this was even a possibility at such a level (T. N. Vu et al, 2019). They found that this tool can be used to find mutations at the individual somatic cell level and discover differences in cell homogeneity, which could be applied in some fashion to examining patients for cancer prognosis and other diagnosable diseases using this method (T. N. Vu et al, 2019).

Vu, T. N., Nguyen, H. N., Calza, S., Kalari, K. R., Wang, L., & Pawitan, Y. (2019). Cell-level somatic mutation detection from single-cell RNA sequencing. *Bioinformatics (Oxford, England)*, *35*(22), 4679–4687. https://doi.org/10.1093/bioinformatics/btz288

Reply | Quote | Edit

▲ Hide 1 reply

SARA CARSANARO

18 days ago

RE: Single-cell RNA-seq

Hey Alex!
I like that you found evidence for the utility of single-cell RNA-seq in cancer for identifying somatic variants and understanding tumor heterogeneity. Clonal expansion is the signature of natural selection in cancer, and cancer treatment can act as a form of artificial selection by selecting for cells with variants that convey resistance. Some interesting examples of this include: EGFR T790M for resistance to Gefitinib, ALK L1196 for resistance to Crizotiniib, and BRCA reversion variants for resistance to PARP inhibitors. Single-cell RNA-seq is one method of screening for these resistance variants, which can be helpful when deciding on a treatment path for a patient and can provide valuable insight for next steps when a patient stops responding to treatment.