

Genómica Computacional

1. ¿Qué pudo haber pasado con la Cepa 2? (10 pts)

Lo que pudo haber pasado con la Cepa 2 de la bacteria *Escherichia coli* es que todas las cepas son distintas y tienen genes distintos. En el experimento se está tratando de ver si la bacteria utilizada tiene un resultado de lactosa positivo o lactosa negativo al utilizar un inductor (IPTG). Esta bacteria es capaz de metabolizar lactosa únicamente cuando no se encuentren azúcares más fáciles para usar como fuente de energía. Es decir si se encuentra glucosa u otra azúcar los genes necesarios para la metabolización de lactosa están apagados. Los genes que son necesarios para que se de son lactosa permeasa y B-galactosidasa. Es probable que en la cepa 2 para el primer experimento agarraron una bacteria con ambos genes expresados y en el experimento 2 utilizaron una bacteria que o no tenía ninguno de los genes expresados o solo uno. Otra opción que paso con la Cepa 2 es que ocurrió una mutación en la bacteria. Dado a que las bacterias se pueden pasar genes es probable que la bacteria o bacterias utilizadas tengan una mutación. Esta mutación puede haber sido la causante de que los genes utilizados para la metabolización de lactasa no existieran y por lo tanto no se pudo metabolizar la lactasa.

2. ¿Qué estructura genética está asociada con este fenotipo? Descríbala (10 pts)

La estructura genética que está asociada con este fenotipo es el operon *lac*. Este operón incluye los genes lactosa permeasa (*lacY*) y la B-galactosidasa (*lacZ*). Este operón está regulado por factores distintos: el gen activo *lacI* y el sitio activador. El gen activo *lacI* es el encargado de codificar un polipéptido que es un represor transcripcional. En una célula normal hay aproximadamente 10 proteínas represoras *lac*. Estas proteínas son las que se unen a sitios en el promotor del operón *lac*. Cuando estas se unen exitosamente estas no dejan que se expresen los genes. El sitio activador se une con la proteína CAP. Esta proteína se codifica por un gen que no se encuentra del operón *lac* y está regulada por el cofactor cAMP. El cAMP se acumula en la célula cuando los azúcares o nutrientes están en niveles bajos dentro de la célula. Por lo tanto si no hay niveles bajos de azúcares o nutrientes en la célula no hay cAMP. Si no hay cAMP la proteína CAP no se une con el sitio y no activa el operón *lac*. Cuando hay niveles bajos de azúcar como glucosa hay cAMP y la proteína CAP se puede unir al sitio activo y activar el operón *lac*. Con el operón *lac* activado se activa la RNA polimerasa y esto hace que se pueda sintetizar la lactosa permeasa y la B-galactosidasa. Tomando lo mencionado anteriormente, es posible que los niveles de azúcar de la célula sean bajos y los niveles de el cofactor cAMP sean altos y aun así el operón *lac* esté inactivo por ausencia de lactosa.

3. Plantee un método de análisis para confirmar su hipótesis del inciso 1. (suponga que ya cuenta con la extracción de ADN) (20 pts)

El método de análisis para confirmar mi hipótesis del inciso 1 tomando en cuenta que ya cuento con la extracción de ADN es hacer un PCR normal, en un laboratorio, o un PCR *in silico*. Este PCR lo haría para poder ver si las bacterias que utilice para el estudio tienen una mutación en el *lacI* esto me ayuda a ver si él represión nunca se junta con la lactosa. Esta mutación en el *lacI* hace que se no se pueda expresar el operón *lac* aunque esté presente la lactosa o no. Algunas causas de esta mutación son que el represor no se puede juntar con el inductor o que si se pueden juntar pero el inductor no puede hacer el cambio conformacional necesario para que este se desprenda del operador. Aparte pueden ocurrir otras mutaciones con *lacI*, estas mutaciones incluyen *lacI* tight binding, esta es muy parecida a la anterior pero afecta a otro dominio de la proteína *lac* represora. El PCR también podría servir para ver si hay alguna delección del operón *lac* y o que este no tenga un promotor y no se puede activar que sea una mutación más grande y haya desaparecido todo el operón. Después de hacer el PCR, en laboratorio, o *in silico* sería muy importante ver si tiene el operón *lac* o si tiene una mutación el *lacI* o una mutación Y- (cuando no hay permeasa y la lactosa no puede entrar a la célula).

4. Si tuviera que utilizar una cepa de referencia, ¿Cuál de las tres escogería? ¿Por qué? (20 pts)

Si tuviera que utilizar una cepa de referencia escogería o la cepa 1 o la cepa 3. Ya que en los experimentos dados estas presentan resultados consistentes. En ambas cepas se presentan resultados consistentes de lactosa positivo para el control y agregando el inductor (IPTG). Al ver que estas cepas presentan resultados consistentes creo que estas serían buenas para utilizar como referencia. Estas cepas no están tan mutadas o expresan los genes necesarios para la metabolización de la lactosa más consistentemente. No me gustaría utilizar la cepa 2 para cepa de referencia ya que esta no presenta un resultado de lactosa positivo consistente. Esto nos indica que esta cepa o no tiene los genes o no están expresados. Al descartar la cepa 2 como cepa de referencia y tener como opciones la 1 y la 3. Me decido en la cepa 3. Decidí que la cepa 3 es una buena cepa de referencia ya que como había mencionado anteriormente esta presenta resultados positivos y consistentes tanto para el experimento 1 como para el experimento 2. Esta cepa tiene los genes lactosa permeasa (*lacY*) y la B-galactosidasa (*lacZ*) expresados para los 2 experimentos y por lo tanto presentan un resultado de lactosa positivo.

Kristen Brandt
171482

5. Para comprobar su hipótesis busque 2 secuencias que cumplan con las características de la cepa de referencia elegida y una con el fenotipo de la cepa 2 y realice *in silico* lo planteado en el inciso 3. Deje descrito que secuencias utilizó (Núm. De Acceso, tamaño (bp), base de datos donde la obtuvo) y resultados obtenido. (20 pts)

Positivo: K-12 GM4792 Lac+

- <https://www.ncbi.nlm.nih.gov/nuccore/CP011342.2>

NZ_CP011342
4622342 bp

Negativo: K-12 GM4792 Lac-

- <https://www.ncbi.nlm.nih.gov/bioproject/224131>
- CP011343
- 4621656 bp

En la siguiente imagen se pueden ver los estadísticos de cada una de las dos secuencias utilizadas. En la primera columna de Value y % of total se presentan los datos de K-12 GM4792 Lac+ mientras que en la segunda columna de Value y % of total se presentan los datos de K-12 GM4792 Lac-.

Genome statistics				
Attribute	Value ^b	% of Total ^{a,b}	Value ^c	% of Total ^{a,c}
Genome size (bp)	4,622,342	100.00	4,621,656	100.00
DNA coding (bp)	3,888,159	84.12	3,873,721	83.82
DNA G + C (bp)	2,348,605	50.81	2,348,022	50.80
DNA scaffolds	1		1	0.00
Total genes	4,144	100.00	4,117	100.00
Protein coding genes	4,061	98.00	4,043	98.20
RNA genes	83	2.00	74	1.80
Pseudo genes	0	0.00	0	0.00
Genes in internal clusters	2,036	49.13	2,027	49.23
Genes with function prediction	3,922	94.64	3,900	94.73
Genes assigned to COGs	3,592	88.45	3,580	88.55
Genes with Pfam domains	3,838	92.62	3,818	92.74
Genes with signal peptides	410	9.89	408	9.91
Genes with transmembrane helices	1,058	25.53	1,048	25.46
CRISPR repeats	2		2	

Literatura Citada:

Kristen Brandt
171482

Catabolite repression and tryptophan operon. (2001). Retrieved October 25, 2020, from <https://www.dartmouth.edu/~bio23/lect4.html>

Griffiths, A. (1970, January 01). Positive and negative control. Retrieved October 25, 2020, from <https://www.ncbi.nlm.nih.gov/books/NBK22091/>

Libretexts. (2020, August 15). 9.14: E. coli as a model system. Retrieved October 25, 2020, from [https://bio.libretexts.org/Bookshelves/Cell_and_Molecular_Biology/Book:_Biofundamentals_\(Klymkowsky_and_Cooper\)/09:_Genomes,_genes,_and_regulatory_networks/9.14:_E._coli_as_a_model_system](https://bio.libretexts.org/Bookshelves/Cell_and_Molecular_Biology/Book:_Biofundamentals_(Klymkowsky_and_Cooper)/09:_Genomes,_genes,_and_regulatory_networks/9.14:_E._coli_as_a_model_system)

Yaratha, G., Perloff, S., & Changala, K. (2017, October 4). Lactose vs Non-Lactose Fermenting *E. coli*: Epidemiology, Clinical Outcomes, and Resistance. Retrieved October 25, 2020, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5631330/>

Zhang, Y., Zhang, Y., Zhu, B., Zhang, B., Ni, C., Zhang, D., . . . Lin, K. (2015, December 10). Genome sequences of two closely related strains of *Escherichia coli* K-12 GM4792. Retrieved October 25, 2020, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4675052/>