9.14: E. coli as a model system

Every surface of your body harbors a flourishing microbial ecosystem. This is particularly true of the gastrointestinal system, which runs from your mouth and esophagus (with a detour to the nose), through the stomach, into the small and large intestine and the colon\textsuperscript{291}. Each of these regions supports its own unique microbial community (known as a microbiome). These environments differ in terms of a number of properties, including differences in pH and O\textsubscript{2} levels. Near the mouth and esophagus O\textsubscript{2} levels are high and microbes can use aerobic (O\textsubscript{2} dependent) respiration to extract energy from food. Moving through the system O\textsubscript{2} levels decrease until anaerobic (without O\textsubscript{2}) mechanisms are necessary. At different positions along the length of the gastrointestinal tract microbes with different ecological preferences and adaptabilities are found.

One challenge associated with characterizing the exact complexity of the microbiome present at various locations is that often the organisms present are dependent upon one another for growth; when isolated from one another they do not grow. The standard way to count bacteria is to grow them in the lab using plates of growth media. Samples are diluted so that single bacteria land (in isolation from one another) on the plate. When they grow and divide, they form macroscopic colonies and it is possible to count the number of "colony forming units" (CFUs) per original sample volume. This provides a measure of the number of individual bacteria present. If an organism cannot form a colony under the assay conditions, it will appear to be absent from the population. But as we have just mentioned some bacteria are totally dependent on others and therefore do not grow in isolation. To avoid this issue, newer molecular methods use DNA sequence analyses to identify which organisms are present without having to grow them\textsuperscript{292}. The result of this type of analysis reveals the true complexity of the microbial ecosystems living on and within us\textsuperscript{293}.

For our purposes, we will focus on one well known, but relatively minor member of this microbial community, \textit{Escherichia coli}\textsuperscript{294}. \textit{E. coli} is a member of the Enterobacteriaceae family of bacteria and is found in the colon of birds and mammals\textsuperscript{295}. \textit{E. coli} is what is known as a facultative aerobe, it can survive in both anaerobic and an aerobic...
environment. This flexibility, as well as *E. coli*’s generally non-fastidious nutrient requirements make it easy to grow in the laboratory. Moreover, the commonly used laboratory strain of *E. coli*, known as K12, does not cause disease in humans. That said, there are other strains of *E. coli*, such as *E. coli* O157:H7 that is pathogenic (disease-causing). *E. coli* O157:H7 contains 1,387 genes not found in the *E. coli* K12. It is estimated that the two *E. coli* strains diverged from a common ancestor ~4 million years ago. The details of what makes *E. coli* O157:H7 pathogenic is a fascinating topic, but beyond our scope here.

**Adaptive behavior and gene networks (the lac response):** Lactose is a disaccharide (a sugar) composed of D-galactose and D-glucose. It is synthesized, biologically, exclusively by female mammals. Mammals use lactose in milk as a source of calories (energy) for infants. One reason (it is thought) is that lactose is not easily digested by most microbes. The lactose synthesis system is derived from an evolutionary modification of an ancestral gene that encodes the enzyme lysozyme. Through duplication and mutation, a gene encoding the protein α-lactoalbumin was generated. α-lactoalbumin is expressed only in mammary glands, where it forms a complex with a ubiquitously expressed protein, galactosyltransferase, to form the protein lactose synthase.

*E. coli* is capable of metabolizing lactose, but only when there are no better (easier) sugars to eat. If glucose or other compounds are present in the environment the genes required to metabolize lactose are turned off. Two genes are required for *E. coli* to metabolize lactose. The first encodes lactose permease. Lactose, being large and highly hydrophilic cannot pass through the *E. coli* cell membrane. Lactose permease is a membrane protein that allows lactose to enter the cell, moving down its concentration gradient. The second gene involved in lactose utilization encodes the enzyme β-galactosidase, which splits lactose into D-galactose and D-glucose, both of which can be metabolized by proteins expressed constitutively (that is, all of the time) within the cell. So how exactly does this system work? How are the lactose utilization genes turned off in the absence of lactose and how are they turned on when lactose is present and energy is needed. The answers illustrate general principles of the interaction networks controlling gene expression.

In *E. coli*, like many bacteria, multiple genes are organized into what are known as operons. In an operon, a single regulatory region controls the expression of multiple genes. It is also common in bacteria that multiple genes involved in a single metabolic pathway are located in the same operon (the same region of the DNA). A powerful approach to the study of genes is to look for relevant mutant phenotypes. As we said, wild type (that is, normal) *E. coli* can grow on lactose as their sole energy sources. So to understand lactose utilization, we can look mutant *E. coli* that cannot grow on lactose. To make the screen for such mutations more relevant, we first check to make sure that the mutant can grow on glucose. Why? Because we are not really interested (in this case) in mutations in genes that disrupt standard metabolism, for example the ability to use glucose; we seek to understand the genes involved in the specific process of lactose metabolism. Such an analysis revealed a number of distinct classes of mutations:. some led to an inability to respond to the presence of lactose in the medium, others led to the de-repression, that is the constant expression of two genes involved in the ability to metabolize lactose, lactose permease and β-galactosidase. In these mutant strains both genes were expressed where or not lactose was present. By mapping (using the Hfr system, see above) where these mutations are in the genome of *E. coli*, and a number of other experiments, the following model was generated.

The genes encoding lactose permease (*lacY*) and β-galactosidase (*lacZ*) are part of an operon, known as the *lac* operon. This operon is regulated by two distinct factors. The first is the product of a constitutively active gene, *lacI*, which encodes a polypeptide that assembles into a tetrameric protein that acts as a transcriptional repressor. In a typical cell there are ~10 lac repressor proteins present. The lac repressor protein binds to sites in the promoter of the *lac* operon.
When bound to these sites the repressor protein blocks transcription (expression) of the \(lac\) operon. The repressor’s binding sites within the \(lac\) operon promoter appear to be its only functionally significant binding sites in the entire \(E. coli\) genome. The second regulatory element in the system is known as the activator site. It can bind the catabolyte activator protein (or CAP), which is encoded by a gene located outside of the \(lac\) operon. The DNA binding activity of CAP is regulated by the binding of a co-factor, cyclic adenosine monophosphate (cAMP). cAMP accumulates in the cell when nutrients, specifically free energy delivering nutrients (like glucose) are low. Its presence acts as a signal that the cell needs energy. In the absence of cAMP, CAP does not bind to or activate expression of the \(lac\) operon, but in its presence (that is, when energy is needed), the CAP-cAMP protein is active, binds to a site in the \(lac\) operon promoter, recruits and activates RNA polymerase, leading to the synthesis of lactose permease and \(\beta\)-galactosidase RNAs and proteins. However, even if energy levels are low (and cAMP levels are high), the \(lac\) operon is inactive in the absence of lactose because of the binding of the \(lac\) repressor protein to sites (labelled \(0_1\), \(0_2\), and \(0_3\)) in \(lac\) the regulatory region of the operon.

So what happens when lactose appears in the cell’s environment? Well, obviously nothing, since the cells are expressing the \(lac\) repressor, so no lactose permease is present, and lactose cannot enter the cell without it. But that prediction assumes that, at the molecular level, the system works perfectly and deterministically. This is not the case, however, the system is stochastic, that is subject to the effects of random processes - it is noisy and probabilistic.Given the small number of \(lac\) repressor molecules per cell (~10), there is a small but significant chance that, at random, the \(lac\) operon of a particular cell will be free of bound repressor. If this occurs under conditions in which CAP is active then, if lactose is present, we see the effect of positive feedback loop\(^{299}\). When lactose is added, those cells that have, by chance, expressed both the lactose permease and of \(\beta\)-galactosidase (a small percentage of the total cell population) will respond: lactose will enter these cells (since the permease is present) and, since \(\beta\)-galactosidase is also present, it will be converted to allolactone (a reaction catalyzed by \(\beta\)-galactosidase. Allolactone binds to, and inhibits the activity of the \(lac\) repressor protein. In the presence of allolactone the repressor no longer inhibits \(lac\) operon expression and there is further increase (~1000 fold) in the rate of expression of lactose permease and \(\beta\)-galactosidase. \(\beta\)-galactosidase also catalyzes the hydrolysis of lactose into D-galactosidase and D-glucose, which are then used to drive cellular metabolism. Through this process, the cell goes from essentially no expression of the \(lac\) operon to full expression, and with full expression becomes able to metabolize lactose. At the same time, those cells that did not (by chance) express lactose permease and \(\beta\)-galactosidase will not be able to metabolizing lactose at all. So even though all of the \(E. coli\) cells present in a culture may be genetically identical, they can express different phenotypes due to the stochastic nature of gene expression. An example of such behavior is presented in the PhET gene expression basics applet\(^{300}\). In the case of the \(lac\) system, over time the noisy nature of gene expression lead to more and more cells activating their copy of the \(lac\) operon. Once “on”, as long as lactose is present in the system, its entry into the cell and its conversion into allolactone will keep the \(lac\) repressor protein in an inactive state and allow continued expression of the \(lac\) operon.

What happens if lactose disappears from the environment, what determines how long it takes for the cells to return to the state in which they no longer express the \(lac\) operon? The answer is determined by the effects of cell division and regulatory processes. In the absence of lactose the allolactone concentration falls and the \(lac\) repressor protein returns to its active state and inhibits expression of the \(lac\) operon. No new lactose permease and \(\beta\)-galactosidase will be synthesized and their concentrations will fall due to degradation. At the same time, and again because their synthesis has stopped, with each cell division the concentration of the lactose permease and \(\beta\)-galactosidase will decreases by ~50%. With time the proteins will be diluted (and degraded) and so the cells return to the initial state, that is, with the \(lac\) operon off and no copies of either lactose permease or \(\beta\)-galactosidase present.
References

291 The gut microbiome: scourge, sentinel or spectator?: http://www.journaloforalmicrobiology...dly/9367/19922


294 virtual lab on E. coli: http://virtuallaboratory.colorado.edu...uced/Coli.html

295 The Evolutionary Ecology of Escherichia coli: http://www.americanascientist.org/iss...erichia-coli/1

296 Enterohemorrhagic E. coli (EHEC) pathogenesis: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3417627/


298 The basic experimental approach involves a technique known as replica plating: see https://en.wikipedia.org/wiki/Replica_plating


Contributors

- Michael W. Klymkowsky (University of Colorado Boulder) and Melanie M. Cooper (Michigan State University) with significant contributions by Emina Begovic & some editorial assistance of Rebecca Klymkowsky.