

= Escherichia coli =

Escherichia coli (/ ɛʃəˈrɪkiə ˈkɒli / ; also known as *E. coli*) is a gram @-@ negative , facultatively anaerobic , rod @-@ shaped bacterium of the genus *Escherichia* that is commonly found in the lower intestine of warm @-@ blooded organisms (endotherms) . Most *E. coli* strains are harmless , but some serotypes can cause serious food poisoning in their hosts , and are occasionally responsible for product recalls due to food contamination . The harmless strains are part of the normal flora of the gut , and can benefit their hosts by producing vitamin K2 , and preventing colonization of the intestine with pathogenic bacteria . *E. coli* is expelled into the environment within fecal matter . The bacterium grows massively in fresh fecal matter under aerobic conditions for 3 days , but its numbers decline slowly afterwards .

E. coli and other facultative anaerobes constitute about 0 @.@ 1 % of gut flora , and fecal ? oral transmission is the major route through which pathogenic strains of the bacterium cause disease . Cells are able to survive outside the body for a limited amount of time , which makes them potential indicator organisms to test environmental samples for fecal contamination . A growing body of research , though , has examined environmentally persistent *E. coli* which can survive for extended periods outside of a host .

The bacterium can be grown and cultured easily and inexpensively in a laboratory setting , and has been intensively investigated for over 60 years . *E. coli* is a chemoheterotroph whose chemically defined medium must include a source of carbon and energy . Organic growth factors included in chemically defined medium used to grow *E. coli* includes glucose , ammonium phosphate , mono basic , sodium chloride , magnesium sulfate , potassium phosphate , dibasic , and water . The exact chemical composition is known for media that is considered chemically defined medium . *E. coli* is the most widely studied prokaryotic model organism , and an important species in the fields of biotechnology and microbiology , where it has served as the host organism for the majority of work with recombinant DNA . Under favorable conditions , it takes only 20 minutes to reproduce .

= = Biology and biochemistry = =

= = = Type and morphology = = =

E. coli is a gram @-@ negative , facultative anaerobic (that makes ATP by aerobic respiration if oxygen is present , but is capable of switching to fermentation or anaerobic respiration if oxygen is absent) and nonsporulating bacterium . Cells are typically rod @-@ shaped , and are about 2 @.@ 0 micrometers (?m) long and 0 @.@ 25 ? 1 @.@ 0 ?m in diameter , with a cell volume of 0 @.@ 6 ? 0 @.@ 7 ?m³ .

E. coli stains gram @-@ negative because its cell wall is composed of a thin peptidoglycan layer and an outer membrane . During the staining process , *E. coli* picks up the color of the counterstain safranin and stains pink . The outer membrane surrounding the cell wall provides a barrier to certain antibiotics such that *E. coli* is not damaged by penicillin .

Strains that possess flagella are motile . The flagella have a peritrichous arrangement .

= = = Metabolism = = =

E. coli can live on a wide variety of substrates and uses mixed @-@ acid fermentation in anaerobic conditions , producing lactate , succinate , ethanol , acetate , and carbon dioxide . Since many pathways in mixed @-@ acid fermentation produce hydrogen gas , these pathways require the levels of hydrogen to be low , as is the case when *E. coli* lives together with hydrogen @-@ consuming organisms , such as methanogens or sulphate @-@ reducing bacteria .

= = = Culture growth = = =

Optimum growth of *E. coli* occurs at 37 ° C (98 @. @ 6 ° F) , but some laboratory strains can multiply at temperatures of up to 49 ° C (120 @. @ 2 ° F) . Growth can be driven by aerobic or anaerobic respiration , using a large variety of redox pairs , including the oxidation of pyruvic acid , formic acid , hydrogen , and amino acids , and the reduction of substrates such as oxygen , nitrate , fumarate , dimethyl sulfoxide , and trimethylamine N @- @ oxide . *E. coli* is classified as a facultative anaerobe . It uses oxygen when it is present and available . It can however , continue to grow in the absence of oxygen using fermentation or anaerobic respiration . The ability to be able to continue growing in the absence of oxygen is an advantage to bacteria because their survival is increased in environments where water predominates .

= = = Cell cycle = = =

The bacterial cell cycle is divided into three stages . The B period occurs between the completion of cell division and the beginning of DNA replication . The C period encompasses the time it takes to replicate the chromosomal DNA . The D period refers to the stage between the conclusion of DNA replication and the end of cell division . The doubling rate of *E. coli* is higher when more nutrients are available . However , the length of the C and D periods do not change , even when the doubling time becomes less than the sum of the C and D periods . At the fastest growth rates , replication begins before the previous round of replication has completed , resulting in multiple replication forks along the DNA and overlapping cell cycles .

Unlike eukaryotes , prokaryotes do not rely upon either changes in gene expression or changes in protein synthesis to control the cell cycle . This probably explains why they do not have similar proteins to those used by eukaryotes to control their cell cycle , such as cdk1 . This has led to research on what the control mechanism is in prokaryotes . Recent evidence suggests that it may be membrane- or lipid @- @ based .

= = = Genetic adaptation = = =

E. coli and related bacteria possess the ability to transfer DNA via bacterial conjugation or transduction , which allows genetic material to spread horizontally through an existing population . The process of transduction , which uses the bacterial virus called a bacteriophage , is where the spread of the gene encoding for the Shiga toxin from the *Shigella* bacteria to *E. coli* helped produce *E. coli* O157 : H7 , the Shiga toxin producing strain of *E. coli* .

= = Diversity = =

Escherichia coli encompasses an enormous population of bacteria that exhibit a very high degree of both genetic and phenotypic diversity . Genome sequencing of a large number of isolates of *E. coli* and related bacteria shows that a taxonomic reclassification would be desirable . However , this has not been done , largely due to its medical importance , and *E. coli* remains one of the most diverse bacterial species : only 20 % of the genes in a typical *E. coli* genome is shared among all strains .

In fact , from the evolutionary point of view , the members of genus *Shigella* (*S. dysenteriae* , *S. flexneri* , *S. boydii* , and *S. sonnei*) should be classified as *E. coli* strains , a phenomenon termed taxa in disguise . Similarly , other strains of *E. coli* (e.g. the K @- @ 12 strain commonly used in recombinant DNA work) are sufficiently different that they would merit reclassification .

A strain is a subgroup within the species that has unique characteristics that distinguish it from other strains . These differences are often detectable only at the molecular level ; however , they may result in changes to the physiology or lifecycle of the bacterium . For example , a strain may gain pathogenic capacity , the ability to use a unique carbon source , the ability to take upon a particular ecological niche , or the ability to resist antimicrobial agents . Different strains of *E. coli* are often host @- @ specific , making it possible to determine the source of fecal contamination in environmental samples . For example , knowing which *E. coli* strains are present in a water sample

allows researchers to make assumptions about whether the contamination originated from a human , another mammal , or a bird .

== Serotypes ==

A common subdivision system of *E. coli* , but not based on evolutionary relatedness , is by serotype , which is based on major surface antigens (O antigen : part of lipopolysaccharide layer ; H : flagellin ; K antigen : capsule) , e.g. O157 : H7) . It is , however , common to cite only the serogroup , i.e. the O @-@ antigen . At present , about 190 serogroups are known . The common laboratory strain has a mutation that prevents the formation of an O @-@ antigen and is thus not typeable .

== Genome plasticity and evolution ==

Like all lifeforms , new strains of *E. coli* evolve through the natural biological processes of mutation , gene duplication , and horizontal gene transfer ; in particular , 18 % of the genome of the laboratory strain MG1655 was horizontally acquired since the divergence from *Salmonella* . *E. coli* K @-@ 12 and *E. coli* B strains are the most frequently used varieties for laboratory purposes . Some strains develop traits that can be harmful to a host animal . These virulent strains typically cause a bout of diarrhea that is unpleasant in healthy adults and is often lethal to children in the developing world . More virulent strains , such as O157 : H7 , cause serious illness or death in the elderly , the very young , or the immunocompromised .

The genera *Escherichia* and *Salmonella* diverged around 102 million years ago (credibility interval : 57 ? 176 mya) which coincides with the divergence of their hosts : the former being found in mammals and the latter in birds and reptiles . This was followed by a split of the *Escherichia* ancestor into five species (*E. albertii* , *E. coli* , *E. fergusonii* , *E. hermannii* , and *E. vulneris*) . The last *E. coli* ancestor split between 20 and 30 million years ago .

The long @-@ term evolution experiments using *E. coli* , begun by Richard Lenski in 1988 , have allowed direct observation of major evolutionary shifts in the laboratory . In this experiment , one population of *E. coli* unexpectedly evolved the ability to aerobically metabolize citrate , which is extremely rare in *E. coli* . As the inability to grow aerobically is normally used as a diagnostic criterion with which to differentiate *E. coli* from other , closely related bacteria , such as *Salmonella* , this innovation may mark a speciation event observed in the laboratory .

== Neotype strain ==

E. coli is the type species of the genus (*Escherichia*) and in turn *Escherichia* is the type genus of the family Enterobacteriaceae , where the family name does not stem from the genus *Enterobacter* + " i " (sic .) + " aceae " , but from " enterobacterium " + " aceae " (enterobacterium being not a genus , but an alternative trivial name to enteric bacterium) .

The original strain described by Escherich is believed to be lost , consequently a new type strain (neotype) was chosen as a representative : the neotype strain is U5 / 41T , also known under the deposit names DSM 30083 , ATCC 11775 , and NCTC 9001 , which is pathogenic to chickens and has an O1 : K1 : H7 serotype . However , in most studies , either O157 : H7 , K @-@ 12 MG1655 , or K @-@ 12 W3110 were used as a representative *E. coli* . The genome of the type strain has only lately been sequenced . Particularly the use of whole genome sequences yields highly supported phylogenies . Based on such data , five subspecies of *E. coli* were distinguished .

The link between phylogenetic distance (" relatedness ") and pathology is small , e.g. the O157 : H7 serotype strains , which form a clade (" an exclusive group ") ? group E below ? are all enterohaemorrhagic strains (EHEC) , but not all EHEC strains are closely related . In fact , four different species of *Shigella* are nested among *E. coli* strains (vide supra) , while *E. albertii* and *E. fergusonii* are outside of this group . Indeed , all *Shigella* species were placed within a single subspecies of *E. coli* in a phylogenomic study that included the type strain , and for this reason an

according reclassification is difficult . All commonly used research strains of *E. coli* belong to group A and are derived mainly from Clifton 's K @-@ 12 strain (? ? F ? ; O16) and to a lesser degree from d 'Herelle 's *Bacillus coli* strain (B strain) (O7) .

= = Genomics = =

The first complete DNA sequence of an *E. coli* genome (laboratory strain K @-@ 12 derivative MG1655) was published in 1997 . It was found to be a circular DNA molecule 4 @.@ 6 million base pairs in length , containing 4288 annotated protein @-@ coding genes (organized into 2584 operons) , seven ribosomal RNA (rRNA) operons , and 86 transfer RNA (tRNA) genes . Despite having been the subject of intensive genetic analysis for about 40 years , a large number of these genes were previously unknown . The coding density was found to be very high , with a mean distance between genes of only 118 base pairs . The genome was observed to contain a significant number of transposable genetic elements , repeat elements , cryptic prophages , and bacteriophage remnants .

Today , several hundred complete genomic sequences of *Escherichia* and *Shigella* species are available . The genome sequence of the type strain of *E. coli* has been added to this collection not before 2014 . Comparison of these sequences shows a remarkable amount of diversity ; only about 20 % of each genome represents sequences present in every one of the isolates , while around 80 % of each genome can vary among isolates . Each individual genome contains between 4 @,@ 000 and 5 @,@ 500 genes , but the total number of different genes among all of the sequenced *E. coli* strains (the pangenome) exceeds 16 @,@ 000 . This very large variety of component genes has been interpreted to mean that two @-@ thirds of the *E. coli* pangenome originated in other species and arrived through the process of horizontal gene transfer .

= = Gene nomenclature = =