

# Genome Sizes

The genome of an organism is the complete set of genes specifying how its phenotype will develop (under a certain set of environmental conditions). In this sense, then, **diploid** organisms (like ourselves) contain two genomes, one inherited from our mother, the other from our father.

## Table of Genome Sizes (haploid)

	<b>Base pairs</b>	<b>Genes</b>	<b>Notes</b>
<a href="#">Phi-X 174</a>	5,386	10	virus of E. coli
Human <a href="#">mitochondrion</a>	16,569	37	
<a href="#">Epstein-Barr virus (EBV)</a>	172,282	80	causes mononucleosis
<a href="#">nucleomorph of Guillardia theta</a>	551,264	511	all that remains of the nuclear genome of a <a href="#">red alga</a> (eukaryote) engulfed long ago by another eukaryote
<a href="#">Mycoplasma genitalium</a>	580,073	483	three of the smallest true organisms
Ureaplasma urealyticum	751,719	652	
Mycoplasma pneumoniae	816,394	680	
<a href="#">Chlamydia trachomatis</a>	1,042,519	936	most common sexually-transmitted disease ( <b>STD</b> ) bacterium in the U.S.
<a href="#">Rickettsia prowazekii</a>	1,111,523	834	bacterium that causes epidemic typhus
<a href="#">Treponema pallidum</a>	1,138,011	1,039	bacterium that causes syphilis
Mimivirus	1,181,404	1,262	A virus (of an amoeba) with a genome larger than the six cellular organisms above

Rickettsia conorii fever	1,268,755	1,374	causes Mediterranean spotted
<a href="#">Borrelia burgdorferi</a> disease [ <a href="#">Note</a> ]	1.44 x 10 <sup>6</sup>	1,738	bacterium that causes Lyme
Aquifex aeolicus spring in Yellowstone National Park	1,551,335	1,749	bacterium isolated from a hot
Thermoplasma acidophilum1, wall	564,905	1,509	an <a href="#">archaeon</a> that lacks a cell
Campylobacter jejuni poisoning	1,641,481	1,708	frequent cause of food
Helicobacter pylori (not stress and diet)	1,667,867	1,589	chief cause of stomach ulcers
Methanococcus jannaschii look like typical bacteria but their genes are so different from those of either bacteria or eukaryotes that they are classified in a third kingdom: <a href="#">Archaea</a> .	1,664,970	1,783	These unicellular prokaryotes
Aeropyrum pernix	1,669,695	1,885	
Pyrococcus horikoshii 1, Methanobacterium	738,505	1,994	
thermoautotrophicum	1,751,377	2,008	
<a href="#">Haemophilus influenzae</a> infections	1,830,138	1,738	bacterium that causes middle ear
Thermotoga maritima	1,860,725	1,879	marine bacterium
<a href="#">Streptococcus pneumoniae</a>	2,160,837	2,236	the <a href="#">pneumococcus</a>
Archaeoglobus fulgidus	2,178,400	2,437	another member of the <a href="#">Archaea</a>

<a href="#">Neisseria meningitidis</a> 2, epidemics of meningitis in less developed countries.	184,406	2,185	Group A; causes occasional
<a href="#">Neisseria meningitidis</a> 2, of meningitis in the U.S.	272,351	2,221	Group B; the most frequent cause
<a href="#">Encephalitozoon cuniculi</a> <b>eukaryote.</b>	2,507,519	1,997	(plus 69 RNA genes); a parasitic
<a href="#">Propionibacterium acnes</a>	2,560,265	2,333	causes acne
<a href="#">Listeria monocytogenes</a> rest RNAs	2,944,528	2,926	2,853 of these encode proteins; the
<a href="#">Deinococcus radiodurans</a> bacterium noted for its resistance to radiation damage	3,284,156	3,187	on 2 chromosomes and 2 plasmids;
<a href="#">Synechocystis</a> <a href="#">cyanobacteria</a> ("blue-green algae")	3,573,470	4,003	a marine prokaryote, one of the
<a href="#">Vibrio cholerae</a>	4,033,460	3,890	in 2 chromosomes; causes cholera
<a href="#">Mycobacterium tuberculosis</a>	4,411,532	3,959	causes tuberculosis
<a href="#">Mycobacterium leprae</a>	3,268,203	1,604	causes leprosy
<a href="#">Bacillus subtilis</a>	4,214,814	4,779	another bacterium
<a href="#">E. coli</a> proteins; the rest RNAs	4,639,221	4,377	4,290 of these genes encode
<a href="#">Agrobacterium tumefaciens</a> <a href="#">transgenic plants</a> ; shares many genes with <a href="#">Sinorhizobium meliloti</a>	4,674,062	5,419	Useful vector for making
<a href="#">Salmonella enterica var Typhi</a> causes typhoid fever	4,809,037	4,395	+ 2 plasmids with 372 active genes;
<a href="#">Salmonella enterica var Typhimurium</a>	4,857,432	4,450	+ 1 plasmid with 102 active genes

<a href="#">Yersinia pestis</a> causes plague	4,826,100	4,052	on 1 chromosome + 3 plasmids;
Schizosaccharomyces pombe fewer genes than the five prokaryotes below.	12,462,637	4,929	Fission yeast. A <b>eukaryote</b> with
E. coli O157:H7	5.44 x 10 <sup>6</sup>	5,416	strain that is pathogenic for humans
Ralstonia solanacearum plants; 1681 of its genes on a huge plasmid	5,810,922	5,129	soil bacterium pathogenic for many
<a href="#">Pseudomonas aeruginosa</a> opportunistic infections in humans.	6.3 x 10 <sup>6</sup>	5,570	Increasingly common cause of
Streptomyces coelicolor provide us with many antibiotics	6,667,507	7,842	An <a href="#">actinomycete</a> whose relatives
Sinorhizobium meliloti Genome consists of one chromosome and 2 large plasmids.	6,691,694	6,204	The <a href="#">rhizobial symbiont</a> of alfalfa.
Saccharomyces cerevisiae	12,495,682	5,770	Budding yeast. A eukaryote.
Cyanidioschyzon merolae	16,520,305	5,331	A unicellular <a href="#">red alga</a> .
<a href="#">Plasmodium falciparum</a> most dangerous form of malaria.	22,853,764	5,268	Plus 53 RNA genes. Causes the
Thalassiosira pseudonana 40 mitochondrial genes encoding proteins	34.5 x 10 <sup>6</sup>	11,242	A <a href="#">diatom</a> . Plus 144 chloroplast and
<a href="#">Neurospora crassa</a>	38,639,769	10,082	Plus 498 RNA genes.
<a href="#">Caenorhabditis elegans</a> be sequenced.	100,258,171	~19,000	The first multicellular eukaryote to
Arabidopsis thaliana <a href="#">note.</a>	115,409,949	25,498	a flowering plant ( <a href="#">angiosperm</a> ) <a href="#">See</a>

<a href="#">Drosophila melanogaster</a>	122,653,977	13,472	the "fruit fly"
Anopheles gambiae	278,244,063	13,683	Mosquito vector of malaria.
Humans	3.3 x 10 <sup>9</sup>	20,000–25,000	<a href="#">[Link to more details.]</a>
Tetraodon nigroviridis (a pufferfish)	3.42 x 10 <sup>8</sup>	27,918	Although Tetraodon seems to have about the same number of genes as we do, it has much less <a href="#">"junk" DNA</a> so its total genome is about a tenth the size of ours.
Rice	4.3 x 10 <sup>8</sup>	~60,000	
Amphibians	10 <sup>9</sup> - 10 <sup>11</sup>	?	
Psilotum nudum	2.5 x 10 <sup>11</sup>	?	<a href="#">Note</a>

Note: The gene total for **Borrelia burgdorferi** is based on 853 genes on its single chromosome (of 910,724 base pairs) plus 430 genes on 11 of the 17 [plasmids](#) it contains.

**Arabidopsis thaliana** is a plant (in the mustard family) that has the smallest genome known in the plant kingdom and for this reason has become a favorite of plant molecular biologists. The sequences of two of its five chromosomes (#2 and #4) were published in December 1999. The others were reported in December 2000.

Even though **Psilotum nudum** (sometimes called the "whisk fern") is a far simpler plant than Arabidopsis (it has no true leaves, flowers, or fruit), it has 3000 times as much DNA. No one knows why, but 80% or more of it is [repetitive DNA](#) containing no genetic information. This is also the case for some amphibians, which contain 30 times as much DNA as we do but certainly are not 30 times as complex.

The total amount of DNA in the haploid genome is called its **C value**. The lack of a consistent relationship between the C value and the complexity of an organism (e.g., amphibians vs. mammals) is called the **C value paradox**.