

Table 2. Features of a few sequenced genomes, emphasizing redundancy-related parameters (approximate values, source <http://www.cbs.dtu.dk/services/GenomeAtlas/>).

Organism	genome size C (Mbp)	coding density δ (bp/gene)	number of genes g	genomic rate R	average protein length	genomic information (Mbits)	redundancy $1-R$ (%)
<i>ΦX174</i>	0.0054	538	10	1.00	180	0.01	~0
λ bacteriophage	0.0485	683	71	0.95	216	0.09	5
<i>M. genitalium</i>	0.58	1,208	480	0.90	363	1.04	10
<i>H. pylori</i>	1.67	1,066	1,566	0.89	316	2.97	11
<i>H. influenzae</i>	1.83	1,071	1,709	0.86	307	3.15	14
<i>S. aureus</i>	2.80	1,069	2,619	0.84	299	4.70	16
<i>B. subtilis</i>	4.21	1,025	4,106	0.87	297	7.32	13
<i>M. tuberculosis</i>	4.41	1,126	3,918	0.97	364	8.56	3
<i>E. coli</i>	4.64	1,082	4,289	0.87	314	8.08	13
<i>X. fastidiosa</i>	2.52	1,238	2,034	0.78	322	3.93	22
<i>S. cerevisiae</i>	12.06	1,924	6,268	0.70	450	17.3	30
<i>C. elegans</i>	99	5,628	17,585	0.25	469	49.5	75
<i>D.melanogaster</i>	~60*	$\delta \sim 13,235$					
180 Mbp	120	$\delta \sim 8,823$	13,600	0.13	573	46.8	87
Human (old)	1,000*	$\delta \sim 30,000$					
~3,000 Mbp	2,000	$\delta \sim 20,000$	100,000?	~0.03	~300?	~180.0?	~97?
Human (update)	967*	$\delta \sim 112,500$					
~2,900 Mbp	1,933	$\delta \sim 75,000$	~25,800	~0.016	~600	~92.9	~98.4

- highly repeated sequences.