Published in final edited form as:
Science. 1999 November 19; 286(5444): 1571–1577.

Genome Sequence of the Radioresistant Bacterium *Deinococcus radiodurans* R1

Owen White, **Jonathan A. Eisen**, **John F. Heidelberg**, **Erin K. Hickey**, **Jeremy D. Peterson**, **Robert J. Dodson**, **Daniel H. Haft**, **Michelle L. Gwinn**, **William C. Nelson**, **Delwood L. Richardson**, **Kelly S. Moffat**, **Haiying Qin**, **Lingxia Jiang**, **Wanda Pamphile**, **Marie Crosby**, **Mian Shen**, **Jessica J. Vamathevan**, **Peter Lam**, **Lisa McDonald**, **Terry Utterback**, **Celeste Zalewski**, **Kira S. Makarova**[†], **L. Aravind**[†], **Michael J. Daly**[◇], **Kenneth W. Minton**[◇], **Robert D. Fleischmann**, **Karen A. Ketchum**, **Karen E. Nelson**, **Steven Salzberg**, **Hamilton O. Smith**[*], **J. Craig Venter**[*], and **Claire M. Fraser**[**]

The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850

[◇]Department of Pathology, Uniformed Services University of the Health Sciences, Bethesda, MD, 20814 USA.

[†]National Center for Biotechnology Information, National Library of Medicine, Bethesda, MD, 20894 USA.

Abstract

The complete genome sequence of the radiation resistant bacterium *Deinococcus radiodurans* R1 is composed of two chromosomes (2,648,615 and 412,340 basepairs), a megaplasmid (177,466 basepairs), and a small plasmid (45,702 basepairs) yielding a total genome of 3,284,123 basepairs. Multiple components distributed on the chromosomes and megaplasmid that contribute to the ability of *D. radiodurans* to survive under conditions of starvation, oxidative stress, and high levels of DNA-damage have been identified. *D. radiodurans* represents an organism in which all systems for DNA repair, DNA damage export, desiccation and starvation recovery, and genetic redundancy are present in one cell.

Deinococcus radiodurans is a gram-positive, red-pigmented, non-motile bacterium that was originally identified as a contaminant of irradiated canned meat (1). It has been isolated worldwide from locations rich in organic nutrients including soil, animal feces, and processed meats, as well as from dry, nutrient-poor environments including weathered granite in a dry Antarctic valley, room dust, and irradiated medical instruments (2). All species in the genus *Deinococcus*, in particular *D. radiodurans*, are extremely resistant to a number of agents and conditions that damage DNA including ionizing and ultraviolet radiation, and hydrogen peroxide (3). *D. radiodurans* is the most radiation resistant organism described to date; exponentially growing cells are 200-fold more resistant to

---

Correspondence and requests for material should be addressed to C.M.F. (drdb@tigr.org).. [**]To whom correspondence should be addressed. drdb@tigr.org.
[*]Present address: Celera Genomics, 45 West Gude Drive, Rockville, MD, 20850 USA.

The sequences have been deposited in Genbank with accession numbers: AE00513, AE001825, AE001826, and AE001827 for Chromosome I, Chromosome II, the megaplasmid and the plasmid, respectively.

ionizing radiation and 20-fold more resistant to UV irradiation (as measured by survival) than *Escherichia coli* (4). This resistance may be a side effect of mechanisms that are designed to allow survival of periods of extended desiccation (5). The radiation resistance of *D. radiodurans* makes it an ideal candidate for bioremediation of sites contaminated with radiation and toxic chemicals. We selected *D. radiodurans* (type strain R1) for sequencing because, among six closely related species of radioresistant Deinococci (1); it is the only representative that is naturally transformable and therefore amenable to genetic manipulation.

The *D. radiodurans* genome sequence was determined by the random whole genome shotgun method as described previously (7). The assembled nucleotide sequence, restriction maps, Southern hybridizations, and optical map confirm that the genome is composed of four circular molecules: chromosome I (2,648,615 basepair (bp)), chromosome II (412,340 bp), a megaplasmid (177,466 bp) and a plasmid (45,702 bp) (Table 1, Fig. 1). Genes for amino acid utilization, cell envelope formation, and transporters are encoded on chromosome II indicating it is likely essential. Putative origins of replication were found on chromosomes I and II by correlating oligomer skew analysis (8) with the presence of certain genes (*dnaA* and *dnaN* for chromosome I and *parA* for chromosome II). We were unable to identify a likely origin of replication for the megaplasmid or the plasmid either by oligomer skew analysis or by sequence similarity with predicted origins of the chromosomes.

A statistical analysis program (9) to predict coding regions was applied to the *D. radiodurans* sequence and predicted coding regions were analyzed as previously described (7). The genome contains 3,193 open reading frames (ORFs), with an average size of 937 bp, representing 91% of the genome (Tables 1 and http://www.tigr.org/tdb/mdb). A total of 2,192 ORFs (69%) matched sequences available in public databases, of which 1,678 were placed in a biological role classification scheme adopted from (10), and 514 matched hypothetical proteins; 1,001 have no database match (Fig. 1 and Table 2). Gene families within *D. radiodurans* were identified using PSI-BLAST (11). A total of 1,665 (52%) genes were placed into 95 families. The two largest families are the P-loop nucleotide binding proteins (12), with 120 representatives, and the helix-turn-helix family (of DNA binding proteins) with 72 members.

Phylogenetic studies of highly conserved genes have suggested that the *Deinococci* are most closely related to the *Thermus* genus and that these two lineages form a eubacterial phylum (13). To determine the extent of this relationship, we compared the 175 currently available *Thermus thermophilus* proteins against *D. radiodurans* and against all other complete genome sequences (14). The majority (143/175) are most similar to a *D. radiodurans* protein indicating that the *Thermus* and *Deionoccous* lineages share extensive similarity throughout their genomes and are even more closely related than was previously suggested. The observation that all members of the *Thermus* genus are thermophilic and some *Deinococci* are slightly thermophilic (15) suggests that the common ancestor of the *Deinococcus/ Thermus* group was thermophilic. Since growth at high temperature also can cause extensive damage to cellular components, the extreme resistance of the Deinococci may have originated through modification of systems that evolved for resistance to heat.

Of the proteins in *D. radiodurans* that are the most similar to *T. thermophilus* proteins, all except one are encoded by genes on chromosome I. Thus it is possible that only chromosome I shares a common ancestry with the *Thermus* lineage and the smaller genetic elements may have been acquired separately. This possibility is supported by the finding that each genetic elements has statistically distinct nucleotide composition (16) which can be an indication of different evolutionary origins (17). Compositional differences are present even though some gene exchange has occurred between these molecules (Fig. 1). The possibility of acquisition by horizontal transfer of the smaller elements is consistent with the observation that *D. radiodurans* is one of the most transformable species known (19).

The ability to survive the potentially damaging effects of ionizing and ultraviolet irradiation and desiccation can be the result of three mechanisms: prevention, tolerance, and repair. Scavenging oxygen radicals is an important component of prevention because oxygen radicals are a key intermediate in the damage to cells caused by ionizing and UV radiation, and desiccation. Several such prevention genes are present in the *D. radiodurans* genome including catalase, which has been shown to be induced following exposure to ionizing radiation (20), multiple superoxide dismutases, and a homolog of the DPS protein in *E. coli*. Catalase and SOD mutants of *D. radiodurans* are more sensitive to ionizing radiation than the wild type (21) indicating that prevention is a component of resistance in this species.

Whereas prevention and tolerance mechanisms likely contribute to the resistance of *D. radiodurans*, the main component of its resistance is a highly efficient DNA repair system (22). For example, after the induction of hundreds of double-strand breaks by 1.75 Mrads of ionizing radiation, in little over 24 hours, most cells restore the genome without rearrangement or increased mutation frequency. Only a limited amount is known about the molecular mechanisms of repair in this species. Analysis of the genome sequence of *D. radiodurans* identifies a nearly full suite of potential DNA repair activities (Table 2) including nucleotide excision repair (a UvrABCD and a UVDE system that likely correspond to the UV endonuclease α and β activities, respectively), base excision repair (9 DNA glycosylases and an AP endonuclease), mismatch excision repair (MutL and MutS) and various aspects of recombinational repair (for example, RecA, RuvABC, SbcCD). It is important to note that while *recA* mutants are highly radiation sensitive (6), this may be due to *recA* based transcriptional regulation as in *E. coli*, and not *recA* based recombination. The only major repair processes for which homologs are not present are alkylation transfer and photoreactivation; this is consistent with experimental studies (23).

Essentially all the DNA repair genes identified in *D. radiodurans* have functional homologs in other prokaryotic species, suggesting that this complement of genes alone is not sufficient to explain the organism's extreme resistance. However, *D. radiodurans* displays a high level of redundancy in DNA repair genes. No other species studied to date encodes as many DNA glycosylases, MutY-Nths, and UvrAs; of the bacteria, only *B. subtilis* has both UvrABCD and UVDE pathways for nucleotide excision repair, only *E. coli* has both uracil DNA glycosylase and a G:U glycosylase (24), and no species has two different 8-oxo-guanine glycosylases (25). *D. radiodurans* also encodes 23 genes with a signature sequence of the Nudix family of nucleoside-triphosphate pyrophosphorylases (26), which is more than any other prokaryote. Some members of the Nudix family (for example, MutT) limit mutations

by hydrolyzing oxidized products of nucleotide metabolism that are mutagenic when misincorporated into the genome. Thus the extra Nudix family members may be partly responsible for *D. radiodurans'* unique capability to resist the induction of mutations by a broad range of mutagenic agents (27). Many of the extra copies of genes in *D. radiodurans* (for example, the Nudix family, MutY-Nth, SodC) are the result of very recent gene duplication events (Fig. 1)(18).

The polyploid nature of *D. radiodurans* (with logarithmically growing cells containing 4 to 10 genome equivalents) is likely an important component of its efficient homologous recombination-based repair of DNA double-strand breaks. Another important component may be the presence of DNA repeat elements scattered throughout the genome (Fig. 1). These repeats satisfy several expected requirements for involvement in recombinational repair including that they are intergenic, they are ubiquitous in the chromosomes and the megaplasmid, and they occur at a frequency that is comparable to the number of double stranded DNA breaks that can be tolerated by *D. radiodurans*. A possible function of the repeats may be in regulating DNA degradation after damage. DNA degradation following the introduction of double stranded breaks is an integral part of the DNA repair process in *D. radiodurans*; however, the extent of DNA degradation appears to be limited by an inhibitory protein (IrrI) that is activated shortly after DNA damage (28). A binding activity from soluble cell extracts, with specificity for the genomic repeat sequences, was identified experimentally (29). The binding of this factor to the repeats may prevent exhaustive chromosomal degradation after radiation exposure. Binding activity increased to a maximum three hours after DNA damage, continued at that level to seven hours, and then decreased gradually to uninduced levels after 24 hours.

A unique mechanism may contribute to *D. radiodurans*' resistance to DNA damage; this organism transports damaged nucleotides out of the cell (30) which potentially prevents their reincorporation into the genome (4). The presence of two UvrA homologs in *D. radiodurans* may in part explain this unique export activity. Many UvrA homologs (including the UvrA1 of *D. radiodurans* (31)) are involved in the recognition of DNA damage for nucleotide excision repair. It has been proposed that some UvrA homologs may have an additional role in the export of DNA damage because they are closely related to ABC transporter proteins and because UvrA serves as a site for the attachment of nucleotide excision repair to the cell membrane in *E. coli* (24). UvrA2 may be involved in the export process in *D. radiodurans* (possibly as a component of a nucleotide transporter complex) because it is most closely related to the DrrC protein of *Streptomyces peucetius* (http://www.tigr.org/~jeisen/UvrA/UvrA.html) that probably functions to transport antibiotic daunorubicin out of the cell.

Recovery from extreme conditions may also require an increase in the *de novo* synthesis or import of precursors to: i) regenerate new complex molecules that have been damaged, and ii) provide a source of alternative energy when environmental conditions (such as desiccation) are accompanied by a reduction in nutrients. Chromosome II and the megaplasmid contain sets of specialized genes that likely play a role in these types of cellular responses following exposure to extreme physiological conditions (Fig. 1).

A number of genes found on chromosome II and the megaplasmid may provide the cell with non-carbohydrate, nitrogenous precursors for protein production. One source of such compounds could be from the proteins of cells that did not survive the stress condition. Together chromosome II and the megaplasmid encode two of the three candidate hemolysins in the genome and four of the nine extracellular proteases. Chromosome II and the megaplasmid also have operons for three ABC transport systems likely to import amino acids: one with homology to the branch-chain amino acid transporter, livFGHK, a second that may import peptide fragments, and a third with broad substrate specificity, perhaps importing proline and glycine-betaine. These proteins may work in concert with an alanine/glycine permease to supply amino acids from the environment. Several genes on chromosome II encode proteins that are involved in production of ammonia through the action of urease; these include xanthine permease and xanthine dehydrogenase that produce urate, and an ABC transporter with specificity for urea. Ammonia represents the key intermediate for assimilation of nitrogen into amino acids. Typically in bacteria the first step of ammonia assimilation into amino acids occurs via glutamine synthetase, an enzyme that converts ammonia and glutamate to glutamine. A potential source of glutamate may be through the degradation of 4-aminobutyrate and histidine via pathways encoded by genes on chromosome II. This pathway for ammonia utilization is consistent with experimental evidence where *D. radiodurans* is not able to grow on minimal media containing ammonia as the sole nitrogen source but will grow on minimal media supplemented with the amino acids cysteine, glutamine, and histidine.

Other proteins encoded on chromosome II and the megaplsmid are involved in generation of cellular energy and may play important roles in the recovery of *D. radiodurans* from prolonged periods of desiccation and/or starvation. Several proteins encoded on chromosome II are involved in fatty acid degradation and, in *E. coli*, act to convert fatty acids to the energy source acetyl-CoA, usually after other carbon sources have been exhausted. The only carbohydrate transporting phosphoenolpyruvate:phosphotransferase system in *D. radiodurans* is specific for fructose and is encoded on the megaplasmid. The energy for fructose uptake is provided by phosphoenolpyruvate, an intermediate in glycolysis. Transport and phosphorylation of fructose will promote a feed forward metabolic loop that may be used to generate ATP and NADH as *D. radiodurans* recovers from desiccation or other forms of cellular stress.

The megaplasmid contains genes that likely participate in restoration of damaged DNA by synthesis of dNTPs (the classIb ribonucleotide reductase and a cofactor thioredoxin) or dNTP precursors (for example, a periplasmic alkaline phosphatase that generates orthophosphate). The megaplasmid also encodes an extracellular nuclease, highly specific for single stranded DNA and RNA, which is immediately released and activated after radiation exposure (32). The 5′-mononucleotides so generated may be imported via a putative purine permease found on chromosome II.

Analysis of the *D. radiodurans* genome reveals that nearly 30% of the total number of genes encoding proteins with regulatory functions (39/140) including transcription factors, response regulators, and kinases are found on chromosome II and the megaplsmid. The role of these regulatory proteins is not known; however, their localization on the smaller genetic

elements suggests that they may be involved in regulating expression of genes for stress responses. This segregation of potential stress recovery genes may reflect the fact that the smaller genetic elements were acquired from other species or that they are under different regulatory controls as compared to the genes on chromosome I.

Acknowledgments

Supported by the U.S. Department of Energy, Office of Biological and Environmental Reasearch, Cooperative Agreement # DE-FC02-95ER61962. We thank E. V. Koonin for thoughtful contribution to this manuscript.

REFERENCES AND NOTES

1. Anderson A, Nordan H, Cain R, Parrish G, Duggan D. Food Technology. 1956; 10:575–578.
2. Masters CI, Murray RG, Moseley BE, Minton KW. J Gen Microbiol. 1991; 137:1459–69. [PubMed: 1683397] Counsell TJ, Murray RGE. Int J Syst Bacteriol. 1986; 36:202.
3. Minton KW. Mol Microbiol. 1994; 13:9. [PubMed: 7984097]
4. Battista JR. Annu Rev Microbiol. 1997; 51:203. [PubMed: 9343349]
5. Mattimore V, Battista JR. J Bacteriol. 1996; 178:633. [PubMed: 8550493]
6. Daly MJ, Ouyang L, Fuchs P, Minton KW. J Bacteriol. 1994; 176:3508. [PubMed: 8206827]
7. Nelson KE, et al. Nature. 1999; 399:323. [PubMed: 10360571]
8. Salzberg SL, Salzberg AJ, Kerlavage AR, Tomb JF. Gene. 1998; 217:57. [PubMed: 9795135]
9. Salzberg SL, Delcher AL, Kasif S, White O. Nucleic Acids Res. 1998; 26:544. [PubMed: 9421513]
10. Riley M. Microbiol Rev. 1993; 57:862. [PubMed: 7508076]
11. Protein sequences were filtered for low complexity [Wootton JC, Federhen S. Methods Enzymol. 1996; 266:554. [PubMed: 8743706]], grouped by single-linkage clusters in pairwise comparisons by PSI-BLAST [Altschul SF, et al. Nucleic Acids Res. 1997; 25:3389. [PubMed: 9254694]]) using GROUPER [Koonin EV, Mushegian AR, Galperin MY, Walker DR. Mol Microbiol. 1997; 25:619. [PubMed: 9379893]]. Resulting multiple alignments of gene families were used to find additional family members in D. radiodurans and related families were joined into superfamilies. Genes with multiple domains were placed into more than one family.
12. Wolf YI, Brenner SE, Bash PA, Koonin EV. Genome Res. 1999; 9:17. [PubMed: 9927481]
13. Hensel R, Demharter W, Kandler O, Kroppenstedt RM, Stackebrandt E. Int. J. Syst. Bacteriol. 1986; 36:444.Weisburg WG, Giovannoni SG, Woese CR. Syst. Appl. Microbiol. 1989; 11:128. [PubMed: 11542160] Eisen JA. J Mol Evol. 1995; 41:1105. [PubMed: 8587109]
14. Searches were done as in (7).
15. Ferreira AC, et al. Int J Syst Bacteriol. 1997; 47:939. [PubMed: 9336890]
16. The distribution of the 64 trinucleotides was computed for the megaplasmid and the two chromosomes. The composition of the megaplasmid and chromosome II were compared, using the X^2 test, to each other and to similarly sized samples of chromosome I chosen uniformly at random. These tests indicate that all three elements have significantly different trinculeotide composition ($p<0.01$). Although the size of the plasmid is too small to obtain statistically meaningful comparisons of trinucleotide composition, its GC content is significantly different from the remaining genome (56% vs. 67%) (Table 1).
17. Lawrence JG, Ochman H. J Mol Evol. 1997; 44:383. [PubMed: 9089078]
18. A gene is considered recently duplicated if its greatest degree of similarity is to another gene in the D. radiodurans genome relative to genes from other completed genomes.
19. Tirgari S, Moseley BEB. J Gen Microbiol. 1980; 119:287.
20. Tanaka A, Hirano H, Kikuchi M, Kitayama S, Watanabe H. Radiat Environ Biophys. 1996; 35:95. [PubMed: 8792456]
21. Markillie LM, Varnum SM, Hradecky P, Wong KK. J Bacteriol. 1999; 181:666. [PubMed: 9882685]
22. Setlow JK, Duggan DE. Biochim Biophys Acta. 1964; 87:664. [PubMed: 14220696]

23. Moseley BE, Evans DM. J Gen Microbiol. 1983; 129:2437–45. [PubMed: 6415229] Tempest PR, Moseley BE. Mol Gen Genet. 1980; 179:191. [PubMed: 6935492]
24. Eisen, JA. Ph.D. Thesis. Stanford University; 1999.
25. Bauche C, Laval J. J Bacteriol. 1999; 181:262. [PubMed: 9864339]
26. Bessman MJ, Frick DN, O'Handley SF. J Biol Chem. 1996; 271:25059. [PubMed: 8810257]
27. Sweet DM, Moseley BE. Mutat Res. 1976; 34:175. [PubMed: 765808]
28. Udupa KS, O'Cain PA, Mattimore V, Battista JR. J Bacteriol. 1994; 176:7439. [PubMed: 8002565]
29. An electrophoretic mobility-shift assay identified a DNA binding activity from soluble cell extracts specific for the D. radiodurans genomic repeat. A truncated DNA ligand tracking an inverted repeat suggests the protein responsible for binding acts as a dimer. Binding was demonstrated in recA- cells.
30. Boling ME, Setlow JK. Biochim Biophys Acta. 1966; 123:26. [PubMed: 5964047]
31. Agostini HJ, Carroll JD, Minton KW. J Bacteriol. 1996; 178:6759. [PubMed: 8955293]
32. The predicted molecular weight and isoelectric point are consistent with a protein that increases in abundance after irradiation [Tanaka A, Hirano H, Kikuchi M, Kitayama S, Watanabe H. Radiat Environ Biophys. 1996; 35:95. [PubMed: 8792456]]. This protein is likely the nuclease reported in [Mitchel RE. Biochim Biophys Acta. 1980; 621:138. [PubMed: 6243490]
33. Lin CS, Wang HC, Wong TY, Liu JK. Biochem Mol Biol Int. 1998; 45:651. [PubMed: 9713687]
34. Mun C, Del Rowe J, Sandigursky M, Minton KW, Franklin WA. Radiat Res. 1994; 138:282. [PubMed: 8183999]
35. Sharples GJ, Leach DRF. Mol Microbiol. 1995; 17:1215. [PubMed: 8594339]
36. Kitayama S, Kohoroku M, Takagi A, Itoh H. Mutat Res. 1997; 385:151. [PubMed: 9447236]
37. Gutman PD, Fuchs P, Ouyang L, Minton KW. J Bacteriol. 1993; 175:3581. [PubMed: 8501062]

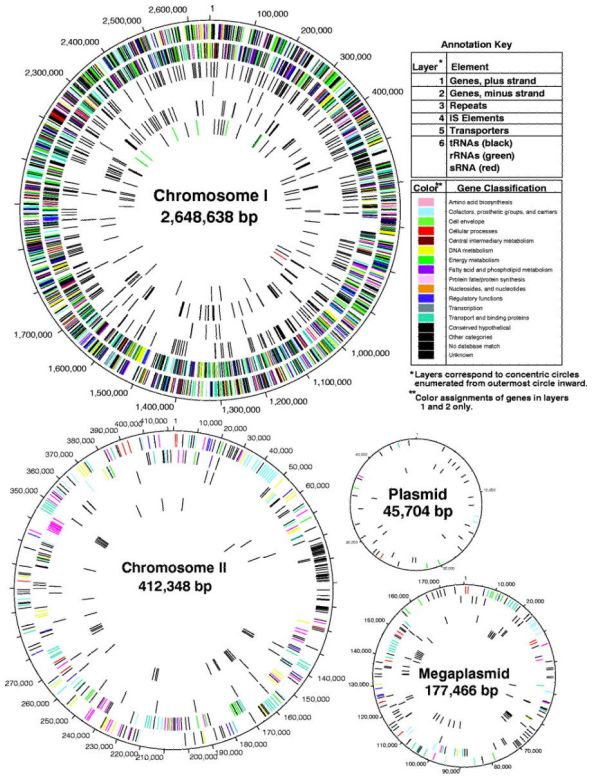

Figure 1. The *D. radiodurans* genome

A. Circular representation. The location of predicted coding regions color-coded by biological role, repeats, insertion elements, rRNA genes, tRNA genes, sRNA genes, and transporters are indicated on the four circular molecules of *D. radiodurans*. B. Linear representation of the genome. Recent gene duplications between molecules (18) are linked by lines; the color of each link corresponds to particular functional categories. Recent duplications within molecules are not shown. Genes duplicated between molecules are linked by lines. Circular and linear molecules are not drawn to scale.

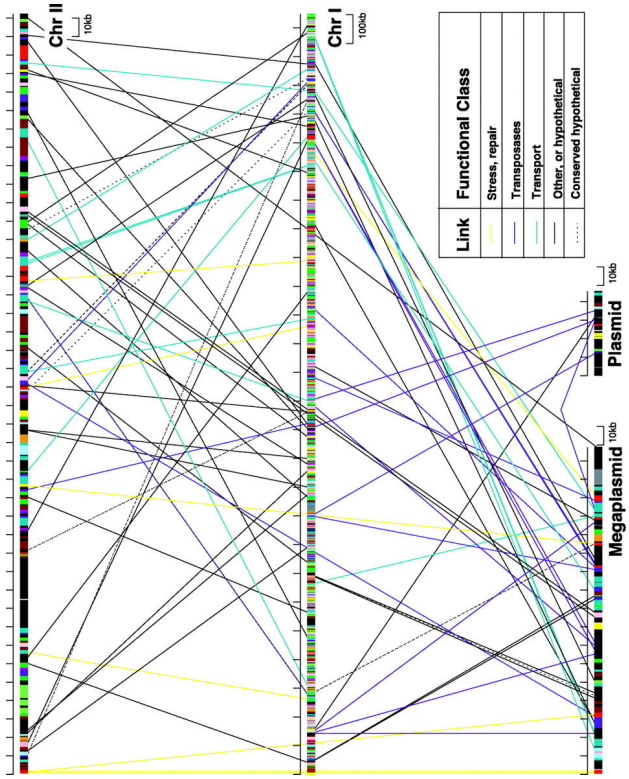

Figure 2.

Table 1

D. radiodurans Genome Features.

GENERAL FEATURES					
Molecule	**Length**	**Average ORF Length (bp)**	**Protein Coding Regions**	**GC-Content**	**Repeat-Content**
Chromosome I	2,648,615	740	90.8%	67.0%	1.8%
Chromosome II	412,340	846	93.5%	66.7%	1.4%
Megaplasmid	177,466	802	90.4%	63.2%	9.2%
Plasmid	45,702	257	80.9%	56.1%	13.0%
All	3,284,123	748	90.9%	66.6%	3.8%

PREDICTED PROTEIN CODING REGION SEQUENCES						
	With a database match					
Molecule	**Identified by database match**	**Putative function assigned**	**Function Unknown**	**Conserved hypothetical**	**No database match**	**Total**
Chromosome I	1817	1195	164	458	817	2634
Chromosome II	254	186	21	47	115	369
Megaplasmid	94	80	9	5	51	145
Plasmid	24	19	2	3	17	41
All	2192	1482	196	514	1001	3193

STABLE RNAs	GENOME COORDINATES
16S rRNA	84,836-86,337 2,286,997-2,285,496 2,469,417-2,467,916
23SrRNA	252,340-254,282 2,246,172-2,245,297 2,585,914-2,585,039
5s rRNA	254,389-254,512 2,245,191-2,245,068 2,584,932-2,584,809
tRNAs	49 (see Fig. 1 for genome locations)

REPEAT ELEMENTS INSERTION ELEMENTS		
Repeat	**Length**	**Copies**
SRE	160	43
SMR1	139	40
SMR2	114	84
SMR4	147	7
SMR5	215	28
SMR7	140	16
SMR8	131	20
SMR9	105	7
SMR10	60	6
		251

INSERTION ELEMENTS		
Insertion element	**Length**	**Copies**
IS2621	1322	14
IS4	1207	9
TCL121	1073	3
TCL9	1048	5
TCL23	1069	2
IS3	1304	1
IS200dr	~1700	8
AR-like	1000	2
		51

Table 2

DNA repair genes and pathways encoded by *D. radiodurans*.

Pathway **Genes in *D. radiodurans***	**Predicted Biochemical Activities and Comments**
Nucleotide Excision Repair	
UvrABCD	Corresponds to UV endonuclease α; *uvrA=mtcAB, uvrD =irrB* (31).
Transcription Repair Coupling	
MFD	Experiments suggest this process may not be present (33).
UV Excision Repair	
UVDE	Corresponds to UV endonuclease β (*uvsCDE*).
Base Excision Repair	
AlkA	3-methyl-guanine glycosylase.
MPG-3MG	3-methyl-guanine glycosylase.
Ung	Uracil DNA glycosylase.
Mug	G:U mismatch glycosylase.
MutM-Fpg	FAPY and 8-oxo-guanine DNA glycosylase.
MutY-Nth-1	Likely a G:A glycosylase because most similar to MutYs.
MutY-Nth-2	Thymine glycol glycosylase from (34)?
MutY-Nth-3	Second FAPY glycosylase from (34)?
MutY-Nth-4	Unknown.
AP Endonuclease	
Xth	May also be an exonuclease.
Mismatch Excision Repair	
MutLS	Absence of MutH suggests different strand recognition system than *E. coli*.
Recombinational Repair	
Initiation	
RecFJNRQ	Nearly complete RecF pathway (RecO missing).
RecD	Absence of RecB and RecC orthologs suggests this gene functions differently than in *E. coli*.
SbcCD	Homology to Rad50/MRE11 suggests a role in DSB repair (35).
Recombination	
RecA	Recombinase. May also regulate transcription of other genes.
Resolution	
RuvABC	Likely redundant to the RecG pathway (36).
RecG	Likely redundant to RuvABC pathway (36).
DNA polymerases	
PolA	Repair replication polymerase (37).
PolC	Chromosomal replication polymerase.
PolX	DNA polymerase of unknown function (PolX family).
Pol	DNA polymerase of unknown function (similar to spo1 phage).
Ligation	
DnlJ	Ligation activity is required for all excision and recombinational repair pathways

Pathway **Genes in *D. radiodurans***	**Predicted Biochemical Activities and Comments**
dNTP pools, cleanup	
MutT and Nudix family	dNTP cleanup. More copies than any other prokaryote.
NrdEFI	Ribonucleotide reductase.
NrdX	Ribonucleotide reductase.
Induction	
LexA	Transcription repressor, possibly for SOS response.
Other	
RadA/SMS	DNA damage response?
HepA	Likely role in transcription and/or DNA repair. Member of SNF2 family *(*24*)*.
MutS2	Possible role in recognizing mismatches but not likely involved in mismatch repair *(*24*)*.
XseA	Exonuclease VII subunit (but XseB is absent).
UvrA2	Export of damaged DNA?
Extracellular nucleases	Degradation of exported DNA?
SSB	Single strand DNA binding protein.
CinA	May recruit RecA to cell membrane.
XerD	Site-specific recombinase.