dues S408, P409, K417, K420, and Y477. This site is located on the side of the knob, maps to a single monomer, and does not overlap two adjacent monomers. This finding implies that the trimeric knob binds to three CAR molecules independently, and it agrees with a determination of the molecular mass of the Ad12 fiber knob complexed with the IgV domain of the CAR protein (*24*).

The fiber knobs of the CAR-binding serotypes display profound dissimilarity of the loops (distal DE, distal FG and HI) that surround the critical residues, which indicates a potential role for these structures in avoiding neutralizing immunity. An analogous situation may exist for the human immunodeficiency virus gp120 envelope glycoprotein (*25–27*) and the sialic acid binding site of the hemagglutinin protein of influenza viruses (*28, 29*).

Our experiments with the vector Adf.F-(TAYT-HA) demonstrate the feasibility of combined CAR-binding ablation and simultaneous retargeting to a new receptor by introduction of new targeting sequences to an adenovirus vector fiber protein. Incorporation of this concept into adenoviral gene therapy vectors will lead to the development of safer, targeted vectors capable of delivering therapeutic genes to tissues with the highest possible specificity.

Note added in proof: Bewley *et al.* (*30*) have analyzed the crystal structure of the CAR-recognizing Ad12 fiber knob complexed with the D1 domain of the CAR protein. Their study confirms that the fiber AB loop plays an essential role in the interaction with the cellular receptor protein, CAR.

References and Notes

1. T. Shenk, in *Fields Virology*, B. N. Fields, D. M. Knipe, P. M. Howley, Eds. (Lippincott-Raven, Philadelphia, 1996), vol. 2, pp. 2111–2148.
2. D. Schnurr and M. E. Dondero, *Intervirology* **36**, 79 (1993).
3. M. Horwitz, in *Virology*, B. N. Fields and D. M. Knipe, Eds. (Raven, New York, 1990), vol. 2, pp. 1679–1721.
4. L. J. Henry, D. Xia, M. E. Wilke, J. Deisenhofer, R. D. Gerard, *J. Virol.* **68**, 5239 (1994).
5. N. Louis, P. Fender, A. Barge, P. Kitts, J. Chroboczek, *J. Virol.* **68**, 4104 (1994).
6. P. W. Roelvink, I. Kovesdi, T. J. Wickham, *J. Virol.* **70**, 7614 (1996).
7. S. C. Stevenson, M. Rollence, B. White, L. Weaver, A. McClelland, *J. Virol.* **69**, 2850 (1995).
8. R. P. Tomko, R. Xu, L. Philipson, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3352 (1997).
9. J. M. Bergelson *et al.*, *Science* **275**, 1320 (1997).
10. P. W. Roelvink *et al.*, *J. Virol.* **72**, 7909 (1998).
11. M. Bai, L. Campisi, P. Freimuth, *J. Virol.* **68**, 5925 (1994).
12. H.-Y. Yeh, N. Pieniazek, D. Pieniazek, H. Gelderblom, R. B. Luftig, *Virus Res.* **33**, 179 (1994).
13. Fiber gene sequences were obtained from the GenBank/EMBL databases. The accession numbers are P03275 (F2), L19194 (F4), M18369 (F5), X74660 (F8), X74659 (F9), X73487 (F12), X74658 and X74659 (F15), X14241 (F17), X94485 (F19), Y14242 (F28), X76548 (F31), X94484 (F37), P18047 (F40 long), and P14267 (F41 long). The figures were generated with RASMOL (Glaxo R&D, UK) and the F5K structure coordinates (Entrez MMDB Id 1398, Protein Data Bank accession number 1KNB).
14. D. Xia, L. Henry, R. D. Gerard, J. Deisenhofer, *Curr. Top. Microbiol. Immunol.* **199**, 39 (1995).
15. ———, *Structure* **2**, 1259 (1994).
16. Site-directed mutagenesis was performed with the Quikchange kit (Stratagene) and the baculovirus constructs pAcSG2-F5KN, pAcSG2-F9K, and pAcSG2-F41LK (*6, 10*). All DNA manipulations and protein expression and purification were as described (*6, 10*).
17. Fiber mutants were expressed in insect cells. Soluble and insoluble fractions were analyzed by protein immunoblotting and detection with a polyclonal antibody to F5K (*4*). Assessment of solubility was made by comparing the signal in the soluble and insoluble fractions with those of wild-type F5K expression in insect cells. The F5K-specific mAbs 3D9, 3F10, and 2C9 were generated for these studies by R. Wagner (ProtoProbe Inc., Milwaukee, WI). These mAbs recognize trimeric, but not heat-denatured fiber.
18. A549 cells (~10^5 per well) were preincubated with the competitor protein at increasing concentrations for 1 hour at 37°C. Ad.CMV-βgal (*9*) was added at a multiplicity of infection of 10 and incubated for 1 hour. Cells were washed twice with Dulbecco's modified Eagle's medium (DMEM) plus 5% fetal calf serum (FCS) and incubated overnight. Galactosidase activity was determined with the Galacto Light assay kit (Tropix Inc.).
19. Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
20. P. W. Roelvink and T. J. Wickham, unpublished data.
21. I. Wilson *et al.*, *Cell* **37**, 767 (1984).
22. D. Brough *et al.*, *J. Virol.* **70**, 6497 (1996).
23. D. Einfeld *et al.*, *J. Virol.* **73**, 9130 (1999).
24. P. Freimuth *et al.*, *J. Virol.* **73**, 1392 (1999).
25. C. D. Rizzuto *et al.*, *Science* **280**, 1949 (1998).
26. P. D. Kwong *et al.*, *Nature* **393**, 648 (1998).
27. R. Wyatt *et al.*, *Nature* **393**, 705 (1998).
28. N. K. Sauter *et al.*, *Biochemistry* **31**, 9609 (1992).
29. W. Weis *et al.*, *Nature* **333**, 426 (1988).
30. M. Bewley, K. Springer, Y.-B. Zhang, P. Freimuth, J. M. Flanagan, *Science* **286**, 1579 (1999).

31. We thank all contributing scientists and technicians.

4 June 1999; accepted 23 September 1999

Genome Sequence of the Radioresistant Bacterium *Deinococcus radiodurans* R1

Owen White,[1] Jonathan A. Eisen,[1] John F. Heidelberg,[1] Erin K. Hickey,[1] Jeremy D. Peterson,[1] Robert J. Dodson,[1] Daniel H. Haft,[1] Michelle L. Gwinn,[1] William C. Nelson,[1] Delwood L. Richardson,[1] Kelly S. Moffat,[1] Haiying Qin,[1] Lingxia Jiang,[1] Wanda Pamphile,[1] Marie Crosby,[1] Mian Shen,[1] Jessica J. Vamathevan,[1] Peter Lam,[1] Lisa McDonald,[1] Terry Utterback,[1] Celeste Zalewski,[1] Kira S. Makarova,[2] L. Aravind,[2] Michael J. Daly,[3] Kenneth W. Minton,[3] Robert D. Fleischmann,[1] Karen A. Ketchum,[1] Karen E. Nelson,[1] Steven Salzberg,[1] Hamilton O. Smith,[1]* J. Craig Venter,[1]* Claire M. Fraser,[1]†

The complete genome sequence of the radiation-resistant bacterium *Deinococcus radiodurans* R1 is composed of two chromosomes (2,648,638 and 412,348 base pairs), a megaplasmid (177,466 base pairs), and a small plasmid (45,704 base pairs), yielding a total genome of 3,284,156 base pairs. 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 amounts of DNA damage were identified. *Deinococcus 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, nonmotile 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 (UV) radiation and hydrogen peroxide (*3*). *Deinococcus radiodurans* is the most radiation-resistant organism described to

[1]The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, USA. [2]National Center for Biotechnology Information, National Library of Medicine, Bethesda, MD 20894, USA. [3]Department of Pathology, Uniformed Services University of the Health Sciences, Bethesda, MD 20814, USA.

*Present address: Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA.

†To whom correspondence should be addressed. E-mail: drdb@tigr.org

date; exponentially growing cells are 200 times as resistant to ionizing radiation and 20 times as resistant to UV irradiation (as measured by survival) as *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 (*6*). The assembled nucleotide sequence, restriction maps, Southern (DNA) hybridizations, and optical map confirm that the genome is composed of four circular molecules: chromosome I [2,648,638 base pairs (bp)], chromosome II (412,348 bp), a megaplasmid (177,466 bp), and a plasmid (45,704 bp) (Tables 1 to 5 and Figs. 1 and 2). Genes for amino acid utilization, cell envelope formation, and transporters are encoded on chromosome II, indicating that it is likely essential. Putative origins of replication were found on chromosomes I and II by correlating oligomer skew analysis (*7*) 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 (*8*) to predict coding regions was applied to the *D. radiodurans* sequence, and predicted coding regions were analyzed as previously described (*6*). The genome contains 3187 open reading frames (ORFs), with an average size of 937 bp, representing 91% of the genome (Tables 1 to 5 and http://www.tigr.org/tdb/mdb). A total of 2185 ORFs (69%) matched sequences available in public databases, of which 1674 were placed in a biological role classification scheme adopted from (*9*), and 511 matched hypothetical proteins; 1002 have no database match (Figs. 1 and 2 and Table 2). Gene families within *D. radiodurans* were identified with PSI-BLAST (*10*). A total of 1665 (52%) genes were placed into 95 families. The two largest families are the P-loop nucleotide binding proteins (*11*), 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 (*12*). 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 (*13*). The majority (143 of 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 (*14*) suggests that the common ancestor of the *Deinococcus-Thermus* group was thermophilic. Because growth at high temperature can also 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 element has statistically distinct nucleotide composition (*15*), which can be an indication of different evolutionary origins (*16*).

Table 1. General features of the *D. radiodurans* genome.

Molecule	Length	Average ORF length (bp)	Protein coding regions	GC content	Repeat-content
Chromosome I	2,648,638	913	90.8%	67.0%	1.8%
Chromosome II	412,348	1,044	93.5%	66.7%	1.4%
Megaplasmid	177,466	1,100	90.4%	63.2%	9.2%
Plasmid	45,704	928	80.9%	56.1%	13.0%
All	3,284,156	937	90.9%	66.6%	3.8%

Table 2. Predicted protein coding region sequences of the *D. radiodurans* genome.

Molecule	With a database match				No database match	Total
	Identified by database match	Putative function assigned	Function unknown	Conserved hypothetical		
Chromosome I	1812	1211	145	456	821	2633
Chromosome II	255	186	22	47	114	369
Megaplasmid	94	80	9	5	51	145
Plasmid	24	16	5	3	16	40
All	2185	1493	181	511	1002	3187

Table 3. Stable RNAs and genome coordinates of the *D. radiodurans* genome.

Stable RNAs	Genome coordinates
16*S* rRNA	84,836–86,337; 2,285,518–2,287,019; 2,467,934–2,469,435
23*S* rRNA	252,343–254,285; 2,245,319–2,246,194; 2,585,058–2,585,933
5*S* rRNA	254,392–254,515; 2,245,090–2,245,213; 2,584,828–2,584,951
tRNAs	49 (see Fig. 1 for genome locations)

Table 4. Repeat elements of the *D. radiodurans* genome.

Repeat	Length	Copies
SRE	160	84
SMR1	139	41
SMR2	114	92
SMR4	147	7
SMR5	215	38
SMR7	140	18
SMR8	131	24
SMR9	105	6
SMR10	60	6
Total		316

Table 5. Insertion elements of the *D. radiodurans* genome.

Insertion element	Length	Copies
IS2621	1322	15
IS4	1207	14
TCL121	1073	3
TCL9	1048	5
TCL23	1069	3
IS3	1304	1
IS200dr	~1700	15
AR-like	1000	2
Total		61

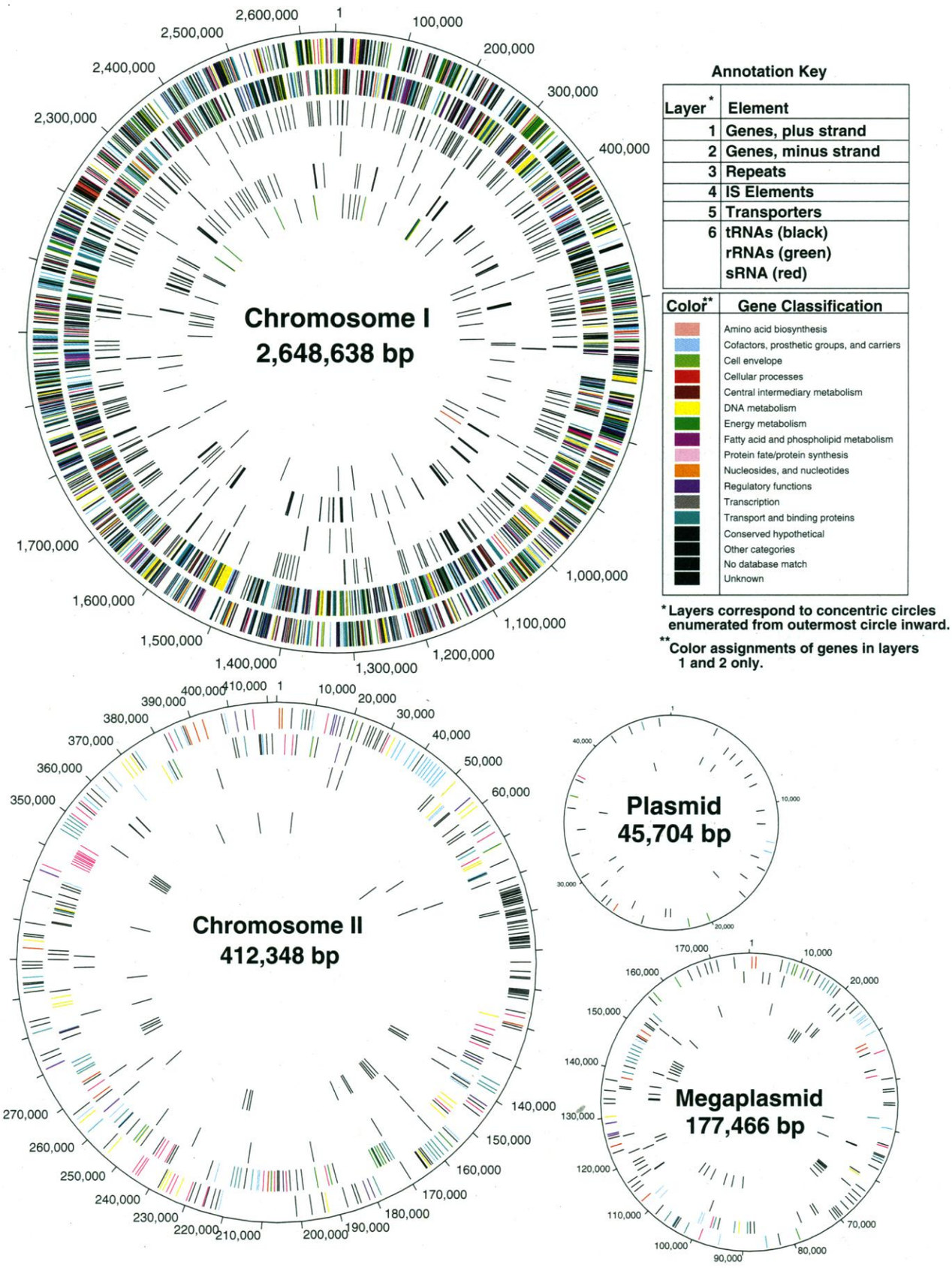

Fig. 1. A circular representation of the *D. radiodurans* genome. The locations of predicted coding regions color-coded by biological role, repeats, insertion (IS) elements, rRNA genes, tRNA genes, sRNA genes, and transporters are indicated on the four circular molecules of *D. radiodurans*.

Compositional differences are present even though some gene exchange has occurred between these molecules (Figs. 1 and 2). 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 (*17*).

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 two catalases, one of which has been shown to be induced after exposure to ionizing radiation (*18*), multiple superoxide dismutases (SOD), 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 (*19*), indicating that prevention is a component of resistance in this species.

Although 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 (*20*). For example, after the induction of hundreds of double-stranded breaks by 1.75 Mrads of ionizing radiation, in a 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 6), including nucleotide excision repair (a UvrABCD system and a UVDE system that likely correspond to the UV endonuclease α and β activities, respectively), base excision repair (nine DNA glycosylases and an apurinic-apyrimidinic endonuclease), mismatch excision repair (MutL and MutS), and various aspects of recombinational repair (for example, RecA, RuvABC, and SbcCD). Although *recA* mutants are highly radiation-sensitive (*21*), this sensitivity 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 finding is consistent with experimental studies (*22*).

Essentially all of 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 amount of redundancy in DNA repair genes. No other species studied to date encodes as many DNA glycosylases, MutY-Nths, and UvrAs; of the bacteria for which complete genomes are published, only *Bacillus subtilis* has both UvrABCD and UVDE pathways for nucleotide excision repair, only *E. coli* has both uracil DNA glycosylase and a G:U glycosylase (*23*), and no species has two different 8-oxoguanine glycosylases (*24*). *Deinococcus radiodurans* also encodes 23 genes with a signature sequence of the Nudix family of nucleoside triphosphate pyrophosphorylases (*25*), which is more than any other prokaryote. Some members of the Nudix family (for example, MutT of *E. coli*) 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*'

Table 6. 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 that this process may not be present (*34*)
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
Ung2	Uracil DNA glycosylase? (*35*)
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 (*36*)?
MutY-Nth-3	Second FAPY glycosylase from (*36*)?
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 that this gene functions differently than in *E. coli*
SbcCD	Homology to Rad50/MRE11 suggests a role in DSB repair (*37*)
Recombination	
RecA	Recombinase; may also regulate transcription of other genes
Resolution	
RuvABC	Likely redundant to the RecG pathway (*38*)
RecG	Likely redundant to RuvABC pathway (*38*)
DNA polymerases	
PolA	Repair replication polymerase (*39*)
PolC	Chromosomal replication polymerase
PolX	DNA polymerase of unknown function (PolX family)
Ligation	
DnlJ	Ligation activity is required for all excision and recombinational repair pathways
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 or DNA repair (or both); 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

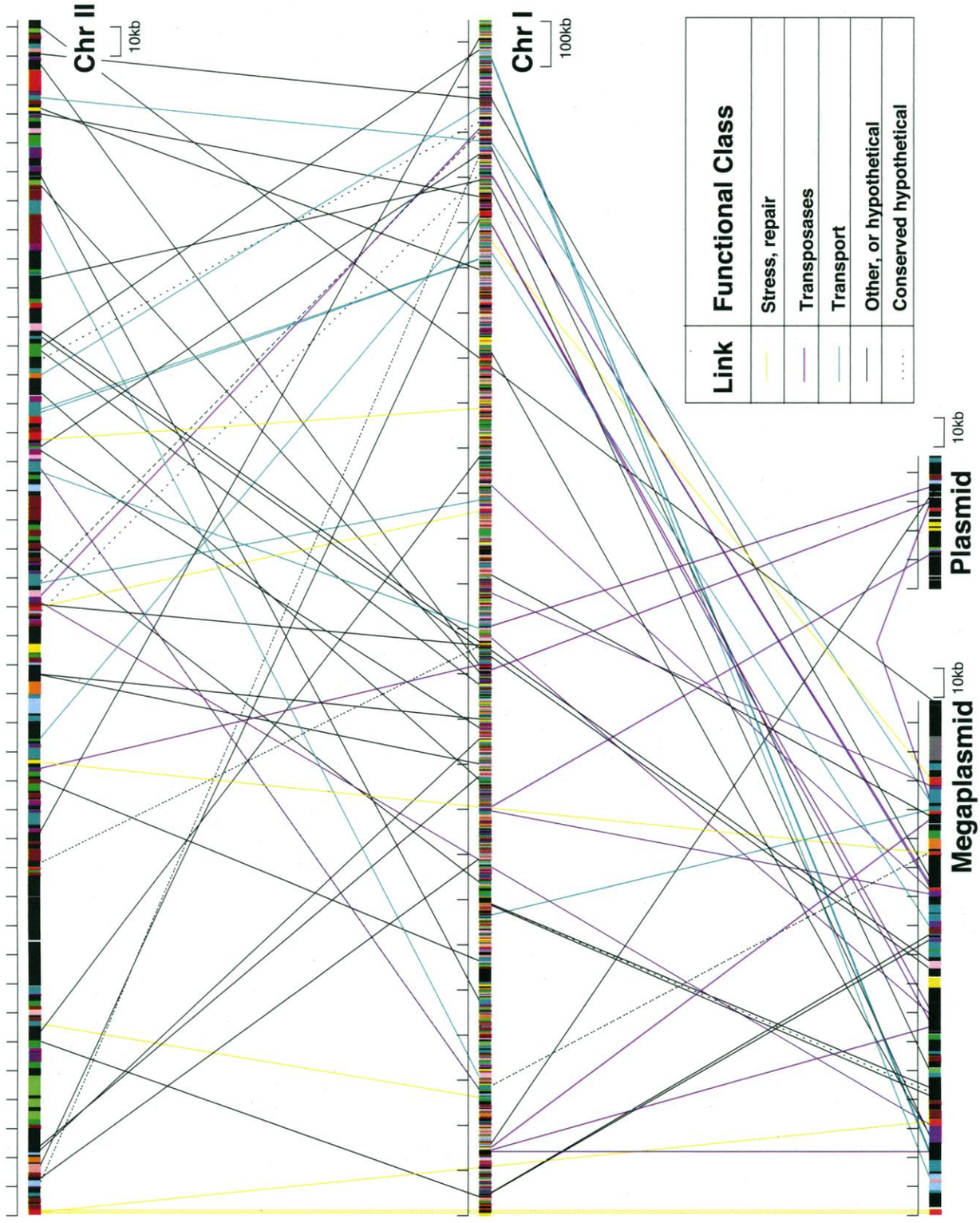

Fig. 2. Linear representation of the *D. radiodurans* genome. Recent gene duplications between molecules (27) are linked by lines; the color of each link corresponds to particular functional categories. Recent duplications within molecules are not shown. Genes duplicated between molecules are linked by lines. Circular and linear molecules are not drawn to scale.

unique capability to resist the induction of mutations by a broad range of mutagenic agents (*26*). Many of the extra copies of genes in *D. radiodurans* (for example, the Nudix family, MutY-Nth, and SodC) are the result of very recent gene duplication events (Fig. 2) (*27*).

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 and Table 4). 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 after the introduction of double-strand 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 3 hours after DNA damage, continued at that level to 7 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), which probably functions to transport antibiotic daunorubicin out of the cell (*32*).

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 after exposure to extreme physiological conditions (Figs. 1 and 2).

A number of genes found on chromosome II and the megaplasmid may provide the cell with noncarbohydrate, 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 proteins include xanthine permease and xanthine dehydrogenase, which 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 through 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 by means of pathways encoded by genes on chromosome II. This pathway for ammonia utilization is consistent with experimental evidence in which *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 megaplasmid are involved in generation of cellular energy and may play important roles in the recovery of *D. radiodurans* from prolonged periods of desiccation or starvation (or both). 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–coenzyme A, 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 adenosine triphosphate and the reduced form of nicotinamide adenine dinucleotide 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 deoxynucleotide triphosphates (dNTPs) (the class Ib 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 (*33*). The 5′-mononucleotides so generated may be imported by means of 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 of 140) including transcription factors, response regulators, and kinases are found on chromosome II and the megaplasmid. 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 with the genes on chromosome I.

References and Notes

1. A. Anderson, H. Nordan, R. Cain, G. Parrish, D. Duggan, *Food Technol.* **10**, 575 (1956).
2. C. I. Masters, R. G. Murray, B. E. Moseley, K. W. Minton, *J. Gen. Microbiol.* **137**, 1459 (1991); T. J. Counsell and R. G. E. Murray, *Int. J. Syst. Bacteriol.* **36**, 202 (1986).
3. K. W. Minton, *Mol. Microbiol.* **13**, 9 (1994).
4. J. R. Battista, *Annu. Rev. Microbiol.* **51**, 203 (1997).
5. V. Mattimore and J. R. Battista, *J. Bacteriol.* **178**, 633 (1996).
6. K. E. Nelson *et al.*, *Nature* **399**, 323 (1999).
7. S. L. Salzberg, A. J. Salzberg, A. R. Kerlavage, J. F. Tomb, *Gene* **217**, 57 (1998).
8. S. L. Salzberg, A. L. Delcher, S. Kasif, O. White, *Nucleic Acids Res.* **26**, 544 (1998).
9. M. Riley, *Microbiol. Rev.* **57**, 862 (1993).
10. Protein sequences were filtered for low complexity [J. C. Wootton and S. Federhen, *Methods Enzymol.* **266**, 554 (1996)] and grouped by single-linkage clusters in pairwise comparisons by PSI-BLAST [S. F. Altschul *et al.*, *Nucleic Acids Res.* **25**, 3389 (1997)] with the use of GROUPER [E. V. Koonin, A. R. Mushegian, M. Y. Galperin, D. R. Walker, *Mol. Microbiol.* **25**, 619 (1997)]. 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.
11. Y. I. Wolf, S. E. Brenner, P. A. Bash, E. V. Koonin, *Genome Res.* **9**, 17 (1999).
12. R. Hensel, W. Demharter, O. Kandler, R. M. Kroppenstedt, E. Stackebrandt, *Int. J. Syst. Bacteriol.* **36**, 444 (1986); W. G. Weisburg, S. G. Giovannoni, C. R. Woese,

Syst. Appl. Microbiol. **11**, 128 (1989); J. A. Eisen, *J. Mol. Evol.* **41**, 1105 (1995).
13. Searches were done as in (7).
14. A. C. Ferreira *et al.*, *Int. J. Syst. Bacteriol.* **47**, 939 (1997).
15. The distribution of the 64 trinucleotides was computed for the megaplasmid and the two chromosomes. The compositions of the megaplasmid and chromosome II were compared with the χ^2 test, with each other, and with similarly sized samples of chromosome I chosen uniformly at random. These tests indicate that all three elements have significantly different trinucleotide 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% compared with 67%) (Tables 1 to 5).
16. J. G. Lawrence and H. Ochman, *J. Mol. Evol.* **44**, 383 (1997).
17. S. Tirgari and B. E. B. Moseley, *J. Gen. Microbiol.* **119**, 287 (1980).
18. A. Tanaka, H. Hirano, M. Kikuchi, S. Kitayama, H. Watanabe, *Radiat. Environ. Biophys.* **35**, 95 (1996).
19. L. M. Markillie, S. M. Varnum, P. Hradecky, K. K. Wong, *J. Bacteriol.* **181**, 666 (1999).
20. J. K. Setlow and D. E. Duggan, *Biochim. Biophys. Acta* **87**, 664 (1964).
21. M. J. Daly, L. Ouyang, P. Fuchs, K. W. Minton, *J. Bacteriol.* **176**, 3508 (1994).
22. B. E. Moseley and D. M. Evans, *J. Gen. Microbiol.* **129**, 2437 (1983); P. R. Tempest and B. E. Moseley, *Mol. Gen. Genet.* **179**, 191 (1980).
23. J. A. Eisen, thesis, Stanford University, Stanford, CA (1999).
24. C. Bauche and J. Laval, *J. Bacteriol.* **181**, 262 (1999).
25. M. J. Bessman, D. N. Frick, S. F. O'Handley, *J. Biol. Chem.* **271**, 25,059 (1996).
26. D. M. Sweet and B. E. Moseley, *Mutat. Res.* **34**, 175 (1976).
27. 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.
28. K. S. Udupa, P. A. O'Cain, V. Mattimore, J. R. Battista, *J. Bacteriol.* **176**, 7439 (1994).
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 that the protein responsible for binding acts as a dimer. Binding was demonstrated in *recA*$^-$ cells.
30. M. E. Boling and J. K. Setlow, *Biochim. Biophys. Acta* **123**, 26 (1966).
31. H. J. Agostini, J. D. Carroll, K. W. Minton, *J. Bacteriol.* **178**, 6759 (1996).
32. K. Furuya and C. R. Hutchinson, *FEMS Microbiol. Lett.* **168**, 243 (1998).
33. The predicted molecular weight and isoelectric point are consistent with a protein that increases in abundance after irradiation (*18*). This protein is likely the nuclease reported by R. E. Mitchel [*Biochim. Biophys. Acta* **621**, 138 (1980)].
34. C. S. Lin, H. C. Wang, T. Y. Wong, J. K. Liu, *Biochem. Mol. Biol. Int.* **45**, 651 (1998).
35. M. Sandigursky and W. A. Franklin, *Curr. Biol.* **9**, 531 (1999).
36. C. Mun, J. Del Rowe, M. Sandigursky, K. W. Minton, W. A. Franklin, *Radiat. Res.* **138**, 282 (1994).
37. G. J. Sharples and D. R. F. Leach, *Mol. Microbiol.* **17**, 1215 (1995).
38. S. Kitayama, M. Kohoroku, A. Takagi, H. Itoh, *Mutat. Res.* **385**, 151 (1997).
39. P. D. Gutman, P. Fuchs, L. Ouyang, K. W. Minton, *J. Bacteriol.* **175**, 3581 (1993).
40. Supported by the U.S. Department of Energy, Office of Biological and Environmental Research, Cooperative Agreement DE-FC02-95ER61962. We thank E. V. Koonin for thoughtful contributions to this manuscript. Correspondence and requests for material should be addressed to C.M.F. (E-mail: drdb@tigr.org). 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.

3 August 1999; accepted 22 October 1999

Species Diversity and Invasion Resistance in a Marine Ecosystem

John J. Stachowicz,[1]* Robert B. Whitlatch,[1] Richard W. Osman[2]

Theory predicts that systems that are more diverse should be more resistant to exotic species, but experimental tests are needed to verify this. In experimental communities of sessile marine invertebrates, increased species richness significantly decreased invasion success, apparently because species-rich communities more completely and efficiently used available space, the limiting resource in this system. Declining biodiversity thus facilitates invasion in this system, potentially accelerating the loss of biodiversity and the homogenization of the world's biota.

Along with habitat modification, the intentional or accidental introduction of new species by humans is a leading cause of the global biodiversity crisis (*1*). Because biological invasions can dramatically alter community composition and ecosystem function (*2–4*) and cause considerable economic damage (*5*), there is substantial interest in understanding why and how successful invasions occur. Although all systems do not appear to be equally invasible (*3, 6, 7*), factors determining the susceptibility of a community to invasion remain unclear. Theory predicts that species-rich communities should be less susceptible to invasion because of a more complete utilization of resources (*6, 8, 9*), but data in support of this prediction have been elusive (*10*). Some observational studies do support a positive relation between biodiversity and invasion resistance (*7, 8*), but others do not (*11*). However, the large number of uncontrolled factors in these studies makes interpreting these findings difficult; manipulative experiments are needed to assess the effect of species richness on invasion success more directly. Studies of terrestrial grasslands and aquatic microbial communities in laboratory microcosms have demonstrated that species-rich communities are more resistant to being invaded by additional species than are species-poor communities (*12*). However, no studies have investigated this relation by using exotic species that currently pose an invasion threat to natural systems, and few studies offer evidence for the mechanisms underlying these patterns.

A growing number of marine invertebrates have been introduced to the coastal waters of southern New England (*13*). In some habitats, these species have invaded successfully and reduced the abundance of native species, whereas in others, they have been unsuccessful and the native community remains unchanged (*14*). Some of these invaders have become locally dominant space holders, including the colonial ascidian *Botrylloides diegensis*, native to the Pacific Ocean (*13*). By introducing invasive species recruits (<1 week old) into experimentally assembled epifaunal communities with varying numbers of native species, we tested the effects of native community species richness on the ability of *Botrylloides* to invade coastal habitats.

Experimental communities were composed of zero to four native species (Fig. 1). This range of diversity treatments was selected because most of the space in undisturbed areas that were equal in size to our communities (100 cm^2) was occupied by three to four species. Each community consisted of 25 2-cm-by-2-cm tiles that fit on tracks bolted to a larger substrate (10 cm by 10 cm). Native sessile invertebrates were cultured on tiles in the field by allowing individuals to settle on tiles, then these tiles were "gardened" weekly to remove all other species except the target species. Once a single tile was covered by an individual or colony of a native species, a number of such tiles were arranged to produce communities with the desired species richness (Fig. 1). Five *Botrylloides* recruits were interspersed throughout each community so that there was only one individual in each row and column of the five-by-five grid of tiles that composed the community (Fig. 1). The remaining 20 tiles in each community were covered by native species. In multispecies communities, available space was divided equally among native species, and the

[1]Department of Marine Sciences, University of Connecticut, 1084 Shennecossett Road, Groton, CT 06340, USA. [2]Academy of Natural Sciences, Estuarine Research Center, St. Leonard, MD 20685, USA.

*To whom correspondence should be addressed. E-mail: jstach@uconnvm.uconn.edu