GENOMIC CONFESSION OF *Deinococcus radiodurans*: IT STARTED OUT AS A RADIATION RESISTANT ORGANISM

Haïtham Sghaier[1,2], Hiroshi Mitomo[2] and Issay Narumi[1*]

[1] Research Group for Gene Resources, Japan Atomic Energy Agency, 1233 Watanuki, Takasaki, Gunma 370-1292, Japan

[2] Material Science Laboratory, Department of Biological and Chemical Engineering, Faculty of Engineering, Gunma University, 1-5-1 Tenjin, Kiryu, Gunma 376-8515, Japan

Fax: +81-27-346-9688

Email: narumi.issei@jaea.go.jp (Issay Narumi)

(Received 6 October, 2005 Accepted 16 November, 2005)

(Abstract)

Previously, analyses of ionizing-radiation-sensitive mutants of *Deinococcus radiodurans* and investigation into the genomic expression profile of the wild-type organism indicated a positive correlation between radiation resistance and desiccation tolerance phenotypes. Mainly because a terrestrial selective force for acquiring resistance to ionizing radiation was unfound, it has been assumed that the radiation resistance of *D. radiodurans* is an incidental phenotype due to its anhydrobiosis defense abilities. Here, following the stratification of the *D. radiodurans* genome, we discuss these issues from evolutionary, genomic, and experimental perspectives. Through the data collected we propose a reconciliatory model wherein radiation resistance is a unique molecular reflection of the early Earth resilience and desiccation tolerance is a mark of cells that colonized land during the Archaean epoch.

(Keywords)

desiccation tolerance, early ancestral genes, gene sharing, genome categorization, land colonization, radiation resistance

Viva Origino Vol.33 No.4 (2005) 243-257

Introduction

Deinococcus radiodurans is a nonsporing eubacteria-type organism that is characterized by extreme resistance to a variety of DNA-damaging agents, with an unusually high resistance to ionizing radiation [1]. The most striking feature of *D. radiodurans* is its capacity for repairing ionizing-radiation-induced DNA double-strand breaks. This bacterium can mend over 100 genomic DNA double-strand breaks during post-irradiation incubation, whereas just few double-strand breaks are lethal in *Escherichia coli* [2].

The ionizing-radiation resistance of *D. radiodurans* has been considered to be difficult to explain from an evolutionary point of view. Battista and his colleague [3] argued that the radiation resistance of *D. radiodurans* might be a fortuitous consequence of an evolutionary process that permitted this bacterium to cope with an environmental stress other than ionizing radiation for the following reason: (1) there is no selective advantage to being ionizing-radiation-resistant in the natural world, (2) there is no terrestrial environment that generates such a high flux of ionizing radiation. Based on their experiments showing that dehydration induces DNA damage including DNA double-strand breaks in *D. radiodurans* and that almost ionizing-radiation-sensitive mutants of this bacterium are sensitive to desiccation, Battista and his colleague hypothesized that *D. radiodurans* is an organism that has adapted to dehydration and that its DNA repair ability is a manifestation of that evolutionary adaptation (desiccation adaptation hypothesis) [3]. Transcriptome analysis revealed a significant overlap in the gene expression profiles of *D. radiodurans* cultures recovering from ionizing radiation and desiccation. During the first hour after a sublethal dose of ionizing radiation, 72 genes were upregulated in *D. radiodurans*. Thirty-three of these loci were also among a set of 73 genes expressed in *D. radiodurans* cultures recovering from desiccation [4]. This experiments clearly showed that ionizing-radiation resistance and desiccation tolerance are functionally interrelated phenomena for *D. radiodurans*. However, it does not provide any substantial evidence for the desiccation adaptation hypothesis. Recently, the ionizing-radiation-resistant fractions of two soil bacterial communities were investigated by exposing an arid soil from desert area and a nonarid soil from forest area to ionizing radiation [5]. The authors of the above investigation argued that recovery of large numbers of ionizing-radiation-resistant bacteria including new species of the genus *Deinococcus* from arid soil and not from nonarid soil provides ecological support for the desiccation adaptation hypothesis. However, all the data cannot rule out a possibility that early bacterial community acquired

ionizing-radiation resistance prior to acquisition of desiccation tolerance.

Luckily, ancient events to which *D. radiodurans* was subjected during its evolution can be theoretically reconstructed on the basis of its present-day genome sequence. We have addressed this problem by considering that the history of each open reading frame (ORF) in *D. radiodurans* could be defined based on the degree to which it reflects the whole genome evolution and the genomes with which it finds a match. Charlebois *et al*. [6] has developed a system, the NeuroGadgets Bioinformatics Web Service (NGIBWS), specializing in microbial evolutionary genomics. First, the 'phylogenetically discordant sequence' (PDS) option within NGIBWS determines if the ORF evolved concordantly (core gene pool) with the rest of the genome. Phylogenetic discordance among ORFs (discordant gene pool) in *D. radiodurans* was assessed by comparison of their rankings using a defined statistical strategy. Generally, removal of PDSs should leave sets of genes which, analyzed jointly, should yield a phylogeny that better corresponds with the phylogenetic grouping of *D. radiodurans*. Also, PDSs, as they include many true species-specific genes, are a useful genomic measure for studying the evolutionary forces that shaped the genome of *D. radiodurans* [7]. Second, the 'evolutionary scope' (ES) option is a measure of the mean distance, computed from the mean normalized BLASTP, between an ORF's genome and the genomes with which it finds a match. An ORF found only within members of the genus *Deinococcus*, for instance, would have a lower value for ES than an ORF distributed more broadly [6]. Through analysis of *D. radiodurans* genome using PDS and ES values, we have gathered evidence that contradicts the desiccation adaptation hypothesis, and we have proposed a new model. In this model, the radiation resistance of *D. radiodurans* is a unique molecular reflection of the early Earth resilience and its desiccation tolerance is a mark of cells that colonized land during the Archaean epoch.

Methods

PDS and ES values were obtained using the NeuroGadgets Inc. web service (http://www.neurogadgets.com) [6]. To determine how well correlated a gene's relationships to its orthologs in other species, PDS values were calculated with BLASTP (1e–5). A PDS is defined as an ORF that exhibits patterns of similarity relationships statistically distinguishable from those of most other ORFs in the same genome, and therefore confuse genome phylogenies. Phylogenetic discordance among ORFs was assessed by statistical comparison of their rankings. Thus, a high PDS value indicates that the gene's relationships track the genome's relationships well [7]. ES of an ORF's matches was computed as the mean distance between *D. radiodurans* and 233 complete genomes available

for analysis (206 Eubacteria, 21 Archaea, 6 Eukaryota) using the genomic distance matrix calculated at a BLASTP cutoff e-value of 1e–5. If an ORF from *D. radiodurans* does not match any other ORF from available species, then ES value of that ORF is equal to zero. In this report, an ORF's PDS and ES values are indicated within brackets as follow: [PDS, ES]. To identify laterally transferred alien genes and genomic islands (GIs) we performed analyses based on statistical parameters (http://www.fut.es/~debb/HGT/) [8], dinucleotide bias (http://www.pathogenomics.sfu.ca/islandpath) [9], and integration sites (http://129.79.232.60/islands/Dra18R.html) [10]. For proteic analysis we employed a clustering algorithm (http://www.protonet.cs.huji.ac.il./) [11]. Proteic families and domains were obtained using Superfamily [12]. The term cluster of genes was used here to describe any genomic segment of DNA containing adjacent ORFs that can be transcribed together.

Results

To probe early evolutionary history of *D. radiodurans*, PDS (from -1 to 1) against ES (from 0 to 0.77) values of 3182 ORFs were plotted (Figure 1). PDS against ES plot shows four

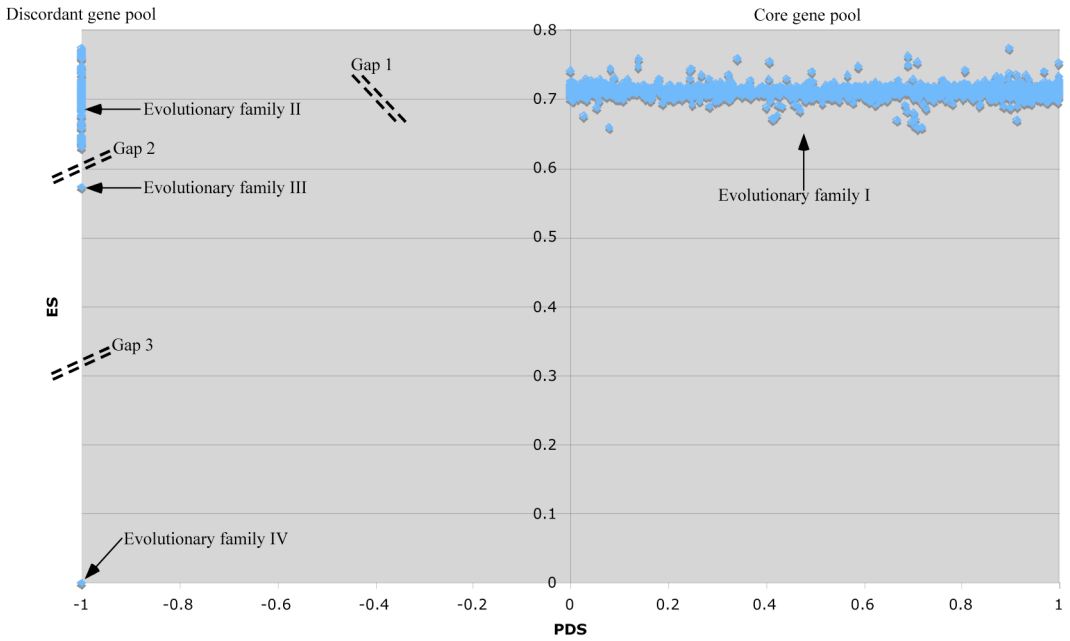

Figure 1. PDS/ES plot of *D. radiodurans* R_1.
Using PDS values ranging from –1 to 1 on the horizontal axis and ES values from 0 to 0.77 on the vertical axis 3182 ORFs from *D. radiodurans* were plotted. Plots reveal that the genome of *D. radiodurans* is characterized by the presence of four evolutionary families (I, II, III, and IV) containing 2198, 249, 63, and 672 ORFs, respectively. The first evolutionary family represents the core gene pool of *D. radiodurans*. The second, third, and fourth evolutionary families form the discordant gene pool of *D. radiodurans*. These four families are separated by three gaps reflecting that they were shaped through distinct genetic sources and pathways under different pressure forces.

evolutionary families (I, II, III, and IV) containing 2198, 249, 63, and 672 ORFs, respectively. These four families are separated by three gaps suggesting that the selective forces shaping the evolution of *D. radiodurans* total gene pool set have not been uniform. One possibility is that these lacuna represent major evolutionary events during which many preexisting genes or part of them were transformed to produce new functions with more complex regulations, a phenomenon called "molecular tinkering/opportunism" [13]. Correlated to PDS and ES values, *D. radiodurans* genome exhibits at least two strategies related to "molecular opportunism": (i) domain duplication, differentiation, combination, and fusion, and (ii) rapid evolution and lateral gene transfer.

The first strategy is exemplified with the ancient duplication event of *mutS1* (DR1039 [0.50, 0.71]) and *mutS2* (DR1976 [0.99, 0.71]) [14]. Also, for example, *D. radiodurans* tyrosine recombinases (DR1555 [0.22, 0.71], DRA0155 [0.31, 0.71], DRB0104 [0.98, 0.71], DRC0018 [0.18, 0.71]) and topoisomerase IB (DR0690 [0.98, 0.70]) evolved by differentiation of the C terminal followed by gene fusion [15]. Moreover, *D. radiodurans* has 16 domain combination pairs, including DRA0207 [–1, 0], DRA0297 [0.57, 0.70], DR0603 [0.55, 0.71], DR2168 [0.98, 0.70], DRA0202 [0.94, 0.71], DR1103 [1, 0.71], DR0467 [0.99, 0.71], DR1417 [0.99, 0.70], DR2168 [0.98, 0.70], DR1252 [0.16, 0.71], DR2506 [0.86, 0.69], DR1247 [0.99, 0.71], DR0337 [–1, 0.69], DR0058 [0.99, 0.71] involving 27 proteic superfamilies, that were unfound in other genomes. Interestingly, DR0467 is a hypothetical protein that has DNA polymerase beta domain adjacent to Rad51 domain in its N-terminal side, nucleotidyltransferase domain as a core string, and PHP (Polymerase and Histidinol Phosphatase) domain as a C-terminal chain.

Secondly, *D. radiodurans* developed new regulatory pathways governed by *pprI* (DR0167 [–1, 0.70]), a general switch for DNA repair [16], within an extensive network of orphan genes, and a laterally acquired genomic island R_1GI1 (DR0511-DR0533) tailed with an alien genes segment (DR0534-DR0547).

The first evolutionary family, with high PDS (> 0) and ES (> 0.65) values, represents the core gene pool of *D. radiodurans*. Such a core, including the minimal set of essential genes, should be preferentially perceived as an "organismal sporadic matches set" because it represents the subset of genes shared by distinct organisms. We observed that the first family includes two categories of ORFs depending on whether ES value is inferior or superior to 0.71. Generally, ORFs with ES value superior to 0.71 (1503 ORFs) code for very ancient enzymes shared among the three primary kingdoms of all extant organisms, with the potential existence of aboriginal lateral gene transfer events. The number of genes in this subset is almost identical to that of the smallest sequenced genome of a free-living organism, *Thermoplasma acidophilum*

(1482 genes). For example, glutamine synthase (DR0451 [0.28, 0.71] and DR2033 [0.52, 0.71]), RecA (DR2340 [1, 0.71]), GTP-binding protein Obg (DR0084 [1, 0.71]), ribosomal protein S12 (DR0305 [0.99, 0.71]), (DR2128 [1, 0.71]), NAD-dependent ligase (DR2069 [1, 0.71]), and ribonuclease PH (DR1585 [0.98, 0.71]) are early ancient gene candidates based on [PDS, ES] criterion, a candidacy supported by other data [17, 18]. Proteins with ES value strictly inferior to 0.71 (695 ORFs) seem to be acquired in a later stage, including SSB (DR0099 [0.99, 0.70]), PriA (DR2606 [0.99, 0.70]), RuvA (DR1274 [0.96, 0.70]), RadA (DR1105 [1, 0.70]), RecO (DR0819 [0.05, 0.70]), RecN (DR1477 [0.83, 0.70]), and RecX (DR1310 [0.83, 0.69]) [19]. Also, the first family encloses conserved hypothetical and hypothetical proteins.

The second, third, and fourth evolutionary families constitute together the discordant gene pool of *D. radiodurans* with PDS = –1. The second family contains very diverged homologous relative to the first family, like ATP-dependent ligase (DRB0100 [–1, 0.71]) [20]. The third family does not contain any known key player implicated in *D. radiodurans* radiation resistance or desiccation tolerance, but it includes a hypothetical protein (DR0060 [–1, 0.57]) with FAD binding domain of DNA photolyase. The fourth family [–1, 0] is composed of 672 genes that orphan genes [21, 22] form the majority.

In a previous DNA microarray analysis [4], 40 ORFs were implied as specific, because they were undetected after ionizing radiation, to the desiccation tolerance phenotype of *D. radiodurans*. PDS and ES values of these ORFs are listed in Table 1. Specific genes to each phenotype, obtained by the transcriptional response, included mainly genes related to metabolism and energy acquisition for radioresistance and many hypothetical proteins for anhydrobiosis. Five molecular chaperones, related to the general stress response of *D. radiodurans* [42], appeared during its response to desiccation, and only a single *hsp20* (DR1114) chaperoned this bacterium's radiation response (see Discussion Section). By PDS and ES means, plotting together radioresistance and anhydrobiosis related ORFs (data not shown)

Table 1. Functions with PDS and ES values of ORFs specific to the desiccation tolerance phenotype of *D. radiodurans* R_1.

	PDS	ES	Function
Mutational inactivation			
DR1172	–1	0	Hypothetical protein
DRB0118	0.66	0.67	Desiccation-associated protein
Transcriptional response			
DR0138	–1	0.71	Hypothetical protein
DR0145	–1	0.7	Hypothetical protein
DR0203	0.74	0.7	Putative ABC-2 type transport system permease protein
DR0253	–1	0	Hypothetical protein

DR0270*	0.92	0.72	Ribonuclease
DR0337	–1	0.69	Hypothetical protein
DR0338	–1	0	Hypothetical protein
DR0404	–1	0	Hypothetical protein
DR0422*	0.96	0.71	Trans-aconitate 2-methyltransferase
DR0475*	0.96	0.71	ABC transporter
DR0492	0.81	0.7	Conserved hypothetical
DR0567*	0.94	0.71	Threonine dehydratase
DR0606*	0.99	0.71	Chaperonin
DR0607*	1	0.71	GroEL protein
DR0621*	0.99	0.71	AAA+ ATPase
DR0661	–1	0	Hypothetical protein
DR0800	–1	0	Hypothetical protein
DR0825*	0	0.71	50S ribosomal protein L31
DR0955	–1	0	Hypothetical protein
DR0985	0.41	0.7	Heat shock protein
DR1004	0.46	0.7	Multidrug resistance protein
DR1046*	1	0.71	ATP-dependent Clp protease
DR1994	–1	0	Hypothetical protein
DR2043*	1	0.71	50S ribosomal protein L7/L12
DR2118*	0.65	0.71	Amino acid transport system ATP-binding protein
DR2142*	0.68	0.71	Putative transmembrane protein
DR2143*	0.56	0.71	Putative Band 7 family protein
DR2187*	0.68	0.71	Putative cyclopropane-fatty-acyl-phospholipid synthase
DR2218	0.37	0.7	Hypothetical protein
DR2279*	0.23	0.71	Alcohol dehydrogenase
DR2306	0.61	0.7	Transcriptional regulator, MerR family
DR2436	–1	0.71	Hypothetical protein
DR2494	0.41	0.67	DNA-directed RNA polymerase subunit omega
DR2562*	0.4	0.71	Putative 3-demethylubiquinone-9 3-methyltransferase
DR2573	–1	0	Hypothetical protein
DR2589*	0.7	0.71	Iron transport system permease protein
DRA0345	–1	0.71	Hypothetical protein
DRA0346	–1	0	Hypothetical protein
DRC0020	0.32	0.7	Putative modification methylase
DRC0021	–1	0	Hypothetical protein

This table includes two desiccation tolerance-related proteins from *D. radiodurans* that were previously analyzed by mutational inactivation [58]. Also, it contains 40 genes, without any water stress-associated ORF, that were induced in response to dry but not to ionizing radiation [4]. Here, by making two simple assumptions – that (i) ORFs with high PDS and ES values (PDS > 0, ES > 0.71) were possessed by the first microbial populations, and that (ii) before 3.05-2.78 billion years ago prokaryotes did not need to be dry tolerant as there was no land – it was possible to constrain the evolutionary distribution of desiccation tolerance-specific ORFs to the discordant gene pool of *D. radiodurans*. Furthermore, the presence of general stress-associated chaperones (DR0606, DR0607, DR0621, DR0985, DR1046) favors the perception of *D. radiodurans* dry tolerance as a general stress response.

* These ORFs are early ancestral genes (PDS > 0, ES > 0.71) in high conflict with evolutionary (most likely desiccation tolerance appeared with Terrabacteria formed by Actinobacteria, *Deinococcus*, and Cyanobacteria) and genomic (indications that desiccation tolerance is a general stress response) information if they were assumed as specific to anhydrobiosis.

reconfirmed the positive correlation between these two phenotypes [3], and contradicted the "desiccation adaptation hypothesis".

Discussion

A number of scientific elements are pertinent to the idea that affiliates life with radiation. Irradiation is widely invoked as a mechanism by which simple carbon-bearing molecules may be polymerized or otherwise reacted to form progressively more complex molecules that could ultimately synthesize life [23]. Bioastrological and paleoenvironmental data suggests that radiation was a driving force during the chemical and early biochemical evolutions [24, 25]. The early Earth was dominated by water as an oceanic lithosphere [25], and as a strong greenhouse gas [26]. If so, this environmental view implies, but does not prove, that in the early microbial population, life was irradiated, probably hot, but not desiccated.

In the first 300,000 years after the big bang, the energy of photons lies in the megaelectronvolt-electronvolt range at temperatures between 10^{11} and 10^{4} degrees Kelvin [24]. During the formation of the Earth, the Sun was emitting an intensity of ultraviolet (UV) radiation probably 10,000 times greater than today and still four times greater at 3.5 Gyr [27]. Simulations confirmed that UV irradiation could have worked as a mutagen and a selective factor [28] leading to a relative enrichment of the ecosystem in sugar-phosphate polymers carrying nitrogenous bases as UV-protectors [29]; besides, evidence has been presented that a string of nucleotides is able to repair itself from UV damage [30]. Theoretical calculations demonstrated that DNA repair abilities of *D. radiodurans* might have been adequate to deal with a high Archaean UV flux and would be sufficient to survive in the mixed layer under surface waters [25, 31].

Cosmic rays, extremely heavy during the early stage of evolution, consist of an ionizing radiation that is higher in energy (multimegaelectronvolt levels) compared to radioactive substances [24]. Radioactive potassium, uranium and thorium constituted the basis of the radioactive heating of the Earth's interior at least 3.5 Gyr [32, 33]. With a higher intensity than cosmic radiation that emit few particles per second, radioactive minerals (thousands of millions of particles per second) have been concentrated at the Earth's surface, and accumulated accretionary coatings of carbon due to irradiation, since early Archean times [34]. Early nuclear reactors of some 3.1-4.1 Gyr were formed in a period shorter than 100 years by a process, which requires neither oxygen nor organic matter for the accumulation of uranium, based on a cycle of successive steps of weathering-erosion-transportation, and sedimentation [24]. Then, Oklo type nuclear reactors originated 2 Gyr [35] after the major Proterozoic rise in oxygen nearly at 2.3 Gyr, termed the Great Oxidation Event (GOE) [36]. It

can be calculated that about 100 million Oklo-type reactor sites could have been active in the past [24]. A typical natural nuclear reactor with a core radius of 2 meters has a total dose rate of 47.4 Gy/h [24]. At this end, it is important to notice that the genotoxic effects, notably DNA double-strand breaks, of radiations (UV, β-, x-, or γ-rays) may occur at energies above the onset of ionizing [37, 38], a process that might have been substantially higher considering the early predominance of a hot and reducing terrestrial milieu [26]. Similarly, heat stress can occur in seemingly cold environments [39]. Indeed, stress depends not only on local environmental conditions, but also on the behavior, morphology, and physiology of a particular organism [40]. The report of UNSCEAR (United Nations Scientific Committee on the Effects of Atomic Radiation) [41], adopted by a previous study [42] to suggest that the ionizing radioresistance of *D. radiodurans* cannot be an adaptation, needs a careful critical appraisal. Briefly, one "atypical" nuclear reactor with high ionizing-radiation dose rates situated beneath the ocean bed would be enough as an adaptation site of abundant radiation biochemistry in the early Earth environment.

If rates and patterns of prokaryotic evolution reflect the developmental history of the Earth's environments [43] where, literally no matter what the physical conditions are, the boundaries of life are not yet precisely defined for prokaryotes [44], radioresistant cells should existed at least 3.5 Gyr. Phylogenetically, rRNA trees place radioresistant bacteria very close to hyperthermophiles and to the root [45]. Two recombinases, represented by RecA of *D. radiodurans* and RadB of *Pyrococcus furiosus*, branch very deeply in the same tree [46]. Correlated to the earliest biomarker evidence (~3.4 Gyr) [47], the *Deinococcus-Thermus* group date back to 2.7-3.4 Gyr. Indeed, phylogenetic analysis and signature sequences support the appearance of this group before cyanobacteria (before 2.7 Gyr) [48], and before the rise in O_2 concentrations (before 2.45 Gyr) [49]. The eubacterial lineage of *Deinococcus* forms with Actinobacteria and Cyanobacteria the Terrabacteria group that colonized land 3.05-2.78 Gyr [50, 51].

Luckily, genomic sequences retain traces of extremely ancient evolutionary events, including the very first steps of life on Earth [52]. Protein sequences do contain information about their historical pasts [53, 54], and do form similar evolutionary families. Here we show that genome-wide evolutionary linkages among proteins in *D. radiodurans*, whose natural history is little known, can be inferred from bioinformatic analyses. The traditional approach for categorizing a genome is based on the codon usage [55]. This approach was used to detect alien genes in *D. radiodurans* [56]. Currently, we embraced a new approach to classify the genome of *D. radiodurans*. For each ORF, we adopted a PDS analysis value, based on each gene's relationships to its orthologs and an ES analysis value, based on genomic distance matrix [6]. If

all genes within a genome have the same phylogenetic history, usually by vertical gene transfer or very ancient lateral gene transfer events, PDS and ES values for each gene vis-à-vis the target genomes should rank similarly and have the same "evolutionary status". Contrarily, ORFs in the discordant gene pool do not have this common ranking, instead they are discordant with PDS = –1.

Hypothetical proteins that appear as a transcriptional response for anhydrobiosis, but not for ionizing radiation, and have a ES value higher than 0.71 (Table 1) are intriguing for the following reason: The plausibility that these ORFs were possessed by early microbial populations (high PDS and ES values) is distinctly at odds with the absence of any requisite for anhydrobiosis defense strategies before land colonization 3.05-2.78 Gyr.

PDS and ES analysis yielded estimates of *D. radiodurans* evolutionary families, favoring the hypothesis that the radiation resistance phenotype is related to more ancient metabolic pathways including the energy acquisition pathway (DR0702 [0.69, 0.74]; DR0970 [0.92, 0.71]; DR0971 [0.97, 0.71]; DR1019 [0.01, 0.71]; DR2195 [0.12, 0.71]; DR2206 [0.72, 0.71]; DR2594 [0.004, 0.72] [57]. This analysis shows that ORFs hypothesizable as desiccation-specific are, with only one exception (DRB0118) [58], in the discordant gene pool of *D. radiodurans*. Functionally, the desiccation tolerance phenotype can be explained by the phenomenon of gene sharing [59]: Gene products contributing to the radioresistance of *D. radiodurans* were recruited to serve an additional function that is dry tolerance. In the terminology of Gould and Vrba [60]: (i) the genome of *D. radiodurans* illustrates adaptations for ionizing radiation serving as exaptations for desiccation, which explains the positive correlation between the two phenotypes, and probably during land colonization (ii) some of these preadaptations becoming modified, in the discordant gene pool, by selection into adaptations for the desiccation tolerance phenotype, which might explain the higher number of unknown proteins in response to desiccation (51 proteins) versus 38 unknown protein in response to ionizing radiation.

Conclusion

Given that the "desiccation adaptation hypothesis" claims are founded on an unfounded evidence for a terrestrial ionizing-radiation natural source, and given that a relevant literature on biophysical, geological, evolutionary, and experimental data favor more plausible alternatives, it was prudent to ask: Which was acquired first by *D. radiodurans* progenitors, radiation resistance or desiccation tolerance? Assuming that *D. radiodurans* genome should harbor a clear-cut answer for this question, we used two evolutionary parameters (PDS and ES values) to stratify its genome into a core gene pool and a discordant gene pool. Then, we

adopted a simple argument: Early ancestral genes cannot be specific to drought stress as the need for the desiccation tolerance phenotype appeared with Terrabacteria. Following the desiccation of *D. radiodurans* cells, the evolutionary status of transcriptionally activated genes, the absence of water stress-related genes, the presence of molecular chaperones have put forward an explanation of *D. radiodurans* desiccation endurance phenotype as a general stress response. Contrarily, preliminary evidence from the evolution of the early Earth's atmosphere coupled with the concept of coevolution of organismal phenotypes and environmental niches fits the view that *D. radiodurans* radiation resistance is a primitive phenotype.

While this study is expected to energize evolutionary discussions about the genetic fundaments of *D. radiodurans* radiation resistance and desiccation tolerance; PDS and ES formulations might guide further experimental analyses such as molecular genetics and protein biochemistry.

References

1. Narumi, I. Unlocking radiation resistance mechanisms: still a long way to go, Trends Microbiol. 11, 422–425 (2003).
2. Krasin, F. and Hutchinson, F. Repair of DNA double-strand breaks in *Escherichia coli*, which requires *recA* function and the presence of a duplicate genome, J. Mol. Biol. 116, 81–98 (1977).
3. Mattimore, V. and Battista, J. R. Radioresistance of *Deinococcus radiodurans*: functions necessary to survive ionizing radiation are also necessary to survive prolonged desiccation, J. Bacteriol. 178, 633–637 (1996).
4. Tanaka, M., Earl, A. M., Howell, H. A., Park, M. J., Eisen, J. A., Peterson, S. N. and Battista, J. R. Analysis of *Deinococcus radiodurans*'s transcriptional response to ionizing radiation and desiccation reveals novel proteins that contribute to extreme radioresistance, Genetics 168, 21–33 (2004).
5. Rainey, F. A., Ray, K., Ferreira, M., Gatz, B. Z., Nobre, M. F., Bagaley, D., Rash, B. A., Park, M. J., Earl, A. M., Shank, N. C., Small, A. M., Henk, M. C., Battista, J. R., Kampfer, P. and da Costa, M. S. Extensive diversity of ionizing-radiation-resistant bacteria recovered from Sonoran Desert soil and description of nine new species of the genus *Deinococcus* obtained from a single soil sample, Appl. Environ. Microbiol. 71, 5225–5235 (2005).
6. Charlebois, R. L., Clarke, G. D., Beiko, R. G. and St Jean, A. Characterization of species-specific genes using a flexible, web-based querying system, FEMS Microbiol. Lett. 225, 213–220 (2003).
7. Clarke, G. D. P., Beiko, R. G., Ragan, M. A. and Charlebois, R. L. Inferring genome trees by using a filter to eliminate phylogenetically discordant sequences and a distance matrix

based on mean normalized BLASTP scores, J. Bacteriol. 184, 2072–2080 (2002).

8. Garcia-Vallve, S., Guzman, E., Montero, M. A. and Romeu, A. HGT-DB: a database of putative horizontally transferred genes in prokaryotic complete genomes, Nucl. Acids Res. 31, 187–189 (2003).
9. Hsiao, W., Wan, I., Jones, S. J. and Brinkman, F. S. IslandPath: aiding detection of genomic islands in prokaryotes, Bioinformatics 19, 418–420 (2003).
10. Mantri, Y. and Williams, K. P. Islander: a database of integrative islands in prokaryotic genomes, the associated integrases and their DNA site specificities, Nucleic Acids Res. 32, D55–58 (2004).
11. Kaplan, N., Sasson, O., Inbar, U., Friedlich, M., Fromer, M., Fleischer, H., Portugaly, E., Linial, N. and Linial, M. ProtoNet 4.0: a hierarchical classification of one million protein sequences, Nucl. Acids Res. 33, D216–218 (2005).
12. Gough, J., Karplus, K., Hughey, R. and Chothia, C. Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure, J. Mol. Biol. 313, 903–919 (2001).
13. Jacob, F. Evolution and tinkering, Science 196, 1161–1166 (1977).
14. Eisen, J. A. A phylogenomic study of the MutS family of proteins, Nucl. Acids Res. 26, 4291–4300 (1998).
15. Krogh, B. O. and Shuman, S. A poxvirus-like type IB topoisomerase family in bacteria, Proc. Natl. Acad. Sci. USA 99, 1853–1858 (2002).
16. Hua, Y., Narumi, I., Gao, G., Tian, B., Satoh, K., Kitayama, S. and Shen, B. PprI: a general switch responsible for extreme radioresistance of *Deinococcus radiodurans*, Biochem. Biophys. Res. Commun. 306, 354–360 (2003).
17. Kumada, Y., Benson, D. R., Hillemann, D., Hosted, T. J., Rochefort, D. A., Thompson, C. J., Wohlleben, W. and Tateno, Y. Evolution of the glutamine synthetase gene, one of the oldest existing and functioning genes. Proc. Natl. Acad. Sci. USA 90, 3009–3013 (1993).
18. Anantharaman, V., Koonin, E. V. and Aravind, L. Comparative genomics and evolution of proteins involved in RNA metabolism, Nucl. Acids Res. 30, 1427–1464 (2002).
19. Eisen, J. A. and Hanawalt, P. C. A phylogenomic study of DNA repair genes, proteins, and processes, Mutat. Res. 435, 171–213 (1999).
20. Liu, Y., Zhou, J., Omelchenko, M. V., Beliaev, A. S., Venkateswaran, A., Stair, J., Wu, L., Thompson, D. K., Xu, D., Rogozin, I. B., Gaidamakova, E. K., Zhai, M., Makarova, K. S., Koonin, E. V. and Daly, M. J. Transcriptome dynamics of *Deinococcus radiodurans* recovering from ionizing radiation, Proc. Natl. Acad. Sci. USA 100, 4191–4196 (2003).
21. Fukuchi, S. and Nishikawa, K. Estimation of the number of authentic orphan genes in

bacterial genomes, DNA Res. 11, 219–231, 311–313 (2004).

22. Siew, N., Azaria, Y. and Fischer, D. The ORFanage: an ORFan database, Nucl. Acids Res. 32, D281–283 (2004).
23. Bernstein, M. P., Dworkin, J. P., Sandford, S. A., Cooper, G. W. and Allamandola, L. J. Racemic amino acids from the ultraviolet photolysis of interstellar ice analogues, Nature 416, 401–403 (2002).
24. Draganic, I. G., Draganic, Z. D. and Adloff, J.-P. Radiation and radioactivity on Earth and beyond, CRC Press, 1993.
25. Cockell, C. S. Ultraviolet radiation and the photobiology of earth's early oceans, Orig. Life Evol. Biosph. 30, 467–499 (2000).
26. Nisbet, E. G. and Sleep, N. H. The habitat and nature of early life, Nature 409, 1083–1091 (2001).
27. Canuto, V. M., Levine, J. S., Augustsson, T. R. and Imhoff, C. L. UV-radiation from the young Sun and oxygen and ozone levels in the prebiological paleoatmosphere, Nature 296, 816–820 (1982).
28. Rothschild, L. J. and Cockell, C. S. Radiation: microbial evolution, ecology, and relevance to mars missions, Mutat. Res. 430, 281–291 (1999).
29. Mulkidjanian, A. Y., Cherepanov, D. A. and Galperin, M. Y. Survival of the fittest before the beginning of life: selection of the first oligonucleotide-like polymers by UV light, BMC Evol. Biol. 3, 12 (2003) doi:10.1186/1471-2148-3-12.
30. Chinnapen, D. J.-F. and Sen, D. A deoxyribozyme that harnesses light to repair thymine dimers in DNA, Proc. Natl. Acad. Sci. USA 101, 65–69 (2004).
31. Garcia-Pichel, F. and Bebout, B. M. The penetration of ultraviolet radiation into shallow water sediments: high exposure for photosynthetic communities, Mar. Ecol. Prog. Ser. 131, 257–262 (1996).
32. Murthy, V. R., Westrenen, W. V. and Fei, Y. Experimental evidence that potassium is a substantial radioactive heat source in planetary cores, Nature 423, 163–165 (2003).
33. Gessmann, C. K. and Wood, B. J. Potassium in the Earth's core? Earth Planetary Sci. Lett. 200, 63–78 (2002).
34. Parnell, J. Mineral radioactivity in sands as a mechanism for fixation of organic carbon on the early Earth, Orig. Life Evol. Biosph. 34, 533–547 (2004).
35. Naudet, R. The Oklo nuclear reactors: 1800 Million years ago, Interdisciplinary Sci. Rev. 1, 72–84 (1976).
36. Holland, H. D. Volcanic gases, black smokers, and the Great Oxidation Event, Geochim. Cosmochim. Acta 21, 3811–3826 (2002).
37. Folkard, M., Prise, K. M., Turner, C. J. and Michael, B. D. The production of single strand and double-strand breaks in DNA in aqueous solution by vacuum UV photons below 10 eV, Radiat. Prot. Dosimetry 99, 147–149 (2002).
38. Boudaiffa, B., Cloutier, P., Hunting, D., Huels,

M. A. and Sanche, L. Resonant formation of DNA strand breaks by low-energy (3 to 20 eV) electrons, Science 287, 1658–1660 (2000).

39. Kingsolver, J. G. and Watt, W. B. Thermoregulatory strategies in *Colias* butterflies: Thermal stress and the limits to adaptation in temporally varying environments, American Naturalist 121, 32–55 (1983).
40. Bartholomew, G. A. The roles of physiology and behaviour in the maintenance of homeostasis in the desert environment, Symp. Soc. Exp. Biol. 18, 7–29 (1964).
41. United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR). Ionizing radiation: Sources and biological effects, New York, NY: United Nations, 1982.
42. Makarova, K. S., Aravind, L., Wolf, Y. I., Tatusov, R. L., Minton, K. W., Koonin, E. V. and Daly, M. J. Genome of the extremely radiation-resistant bacterium *Deinococcus radiodurans* viewed from the perspective of comparative genomics, Microbiol. Mol. Biol. Rev. 65, 44–79 (2001).
43. Knoll, A. H. and Bauld, J. The evolution of ecological tolerance in prokaryotes, Trans. R. Soc. Edinb. Earth Sci. 80, 209–223 (1989).
44. Rothschild, L. J. and Mancinelli, R. L. Life in extreme environments, Nature 409, 1092–1101 (2001).
45. Woese, C. Bacterial evolution, Microbiol. Rev. 51, 221–271 (1987).
46. DiRuggiero, J., Brown, J. R., Bogert, A. P. and Robb, F. T. DNA repair systems in archaea: mementos from the last universal common ancestor? J. Mol. Evol. 49, 474–484 (1999).
47. Schopf, J. W. Microfossils of the Early Archean Apex chert: new evidence of the antiquity of life, Science 260, 640–646 (1993).
48. Brocks, J. J., Logan, G. A., Buick, R. and Summons, R. E. Archean molecular fossils and the early rise of Eukaryotes, Science 285, 1033–1036 (1999).
49. Bekker, A., Holland, H. D., Wang, P. L., Rumble, D., Stein, H. J., Hannah, J. L., Coetzee, L. L. and Beukes, N. J. Dating the rise of atmospheric oxygen, Nature 427, 117–120 (2004).
50. Battistuzzi, F. U., Feijao, A. and Hedges, S. B. A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land, BMC Evol. Biol. 4, 44 (2004). doi:10.1186/1471-2148-4-44.
51. Bern, M. and Goldberg, D. Automatic selection of representative proteins for bacterial phylogeny, BMC Evol. Biol. 5, 34 (2005). doi:10.1186/1471-2148-5-34.
52. Galtier, N., Tourasse, N. and Gouy, M. A nonhyperthermophilic common ancestor to extant life forms, Science 283, 220–221 (1999).
53. Pauling, L. and Zuckerkandl, E. Chemical paleogenetics: Molecular restoration studies of extinct forms of life, Acta Chem. Scand. 17, S9–S16 (1963).

54. Liberles, D. A. and Wayne, M. L. Tracking adaptive evolutionary events in genomic sequences, Genome Biol. 3, 1018.1–1018.4 (2002).

55. Karlin, S. and Mrazek, J. Predicted highly expressed genes of diverse prokaryotic genomes, J. Bacteriol. 182, 5238–5250 (2000).

56. Karlin, S. and Mrazek, J. Predicted highly expressed and putative alien genes of *Deinococcus radiodurans* and implications for resistance to ionizing radiation damage, Proc. Natl. Acad. Sci. USA 98, 5240–5245 (2001).

57. Kyrpides, N., Overbeek, R. and Ouzounis, C. Universal protein families and the functional content of the last universal common ancestor, J. Mol. Evol. 49, 413–423 (1999).

58. Battista, J. R., Park, M. J. and McLemore, A. E. Inactivation of two homologues of proteins presumed to be involved in the desiccation tolerance of plants sensitizes *Deinococcus radiodurans* R1 to desiccation, Cryobiology 43, 133–139 (2001).

59. Piatigorsky, J., O'Brien, W. E., Norman, B. L., Kalumuck, K., Wistow, G. J., Borras, T., Nickerson, J. M. and Wawrousek, E. F. Gene sharing by delta-crystallin and argininosuccinate lyase, Proc. Nat. Acad. Sci. USA 85, 3479–3483 (1988).

60. Gould, S. J. and Vrba. E. S. Exaptation - a missing term in the science of form, Paleobiology 8, 4–15 (1982).