Mem. S.A.It. Vol. 90, 590

Memorie della

The "Black Box" Genome of the extremophilic bacterium Deinococcus radiodurans turning grayish

R. V. L. Zappala[1], P. G. Pascutti[2], M. Leal da Silva[3], and C. A. S. Lage[1]

[1] Laboratório de Radiações em Biologia (LaRBio), Carlos Chagas Filho Institute of Biophysics, Rio de Janeiro Federal University, Avenida Carlos Chagas Filho 373, 21941-902 Rio de Janeiro, Brazil
[2] Laboratório de Modelagem e Dinâmica Molecular, Carlos Chagas Filho Institute of Biophysics, Rio de Janeiro Federal University, Avenida Carlos Chagas Filho 373, 21941-902 Rio de Janeiro, Brazil
[3] Instituto de Biodiversidade e Sustentabilidade, NUPEM, Rio de Janeiro Federal University, Avenida Sâo José do Barreto 764, Bairro São José do Barreto, Macaé, 27965-045 Rio de Janeiro, Brazil

Abstract. In the *Deinococcaceae* group are comprised some of the known robust extremophilic bacteria. Attempts have specially focused not only on the responses against extreme doses of gamma radiation, but also to explain survival of *Deinococcus radiodurans* against other stresses, as desiccation and heat. *D. radiodurans* harbors a number of defensive mechanisms, and the expressed functions to cope with gamma radiation and desiccation revealed that some proteins were assigned to undefined functions. Therefore, it is expected that such proteins of obscure function can code for novel resistance strategies to these extreme conditions. The present contribution aims at elucidating the efforts of this team to describe the survival mechanisms of this species, covering not only the progress made by previous studies, but also predicting the roles of its exclusive proteins that may be responsible for the renowned stress resistance seen to exist in this species. This particular resistance toolbox with potentially novel and exclusive proteins was referred as the "Black Box Genome of *D. radiodurans*".

Key words. Radiation – Deinococcus radiodurans – Molecular modeling – Bioinformatics – Extremophile.

1. Introduction

In the group *Deinococcaceae* are found some of the most resistant organisms known up to date. As a matter of a fact, many members belonging to the genus *Deinococcus* are recognized by their remarkable resistance against UV and ionizing radiation (IR) and some of its members were described a decade ago living in a desert, and also surviving after exposure to thousands of Grays of gamma radiation (Rainey et al. 2005). Not only these bacteria have been characterized as being highly resistant, but also a few species belonging to the genus *Rubrobacter*, for instance (Ferreira et al. 1999).

The mechanisms through which bacteria can survive under the so-called "extreme conditions" have been reviewed in a group of papers that indicates efficient DNA repair and anti-oxidant mechanisms, protein protection, etc. (Blasius et al. 2008; Byrne et al. 2014; Cox & Battista 2005; Cox et al. 2010; Daly 2009; Krisko & Radman 2013). However, the biological singularity of *Deinococcus radiodurans* casts it as an excellent model in understanding radiation resistance mechanisms, given its impressive DNA recovery abilities and other defensive features. Moreover, it can fully repair its complete genome after only a few hours following exposure to ~7 kGy of gamma irradiation (Blasius et al. 2008).

Two chromosomes and two plasmids encode about 3200 proteins in total, out of which 400 are not shared among other species in this planet, even within the same genus (Griffiths & Gupta 2007). Considering the advanced approaches towards gene and protein annotation up to date, it is important to always check the databases in search for new extremophilic functions. Part of this review is dedicated to that end, highlighting potentially novel protein functions peered in *D. radiodurans* genome by refined modeling of some of its exclusive protein structures with similar functions already described for other organisms.

The genomic era helped unveiling particular gene functions of *D. radiodurans* since its genome has been sequenced (White et al. 1999). Afterwards, efforts were made to address the species' extremophilic behavior, by knocking-out some target genes and phenotyping the resulting mutants in terms of their responses against radiation and desiccation (Blasius et al. 2007; Daly 2009; Norais et al. 2009; Schlesinger 2007; Xu et al. 2008). Also, the functions of many genes were not properly annotated, and comparative genomics as well as other "omics" approaches have tried to elucidate such inconsistencies (Lipton et al. 2002; Makarova et al. 2007; Ott et al. 2017; Tanaka et al. 2004; Zhang et al. 2005). The species harbors a well-developed system to deal with extreme conditions, counting not only on repair mechanisms against DNA damage, but also on protection of its protein content against destruction. Moreover, antioxidant compounds provide extra shielding of membranes and DNA. Such features were discussed on a few reviews, and a transcriptome reported in 2014 provided new insights on the antioxidant role of such compounds upon gamma irradiation (Blasius et al. 2008; Luan et al. 2014; Slade & Radman 2011). This review aims at the recent findings on resistance mechanisms and at the analysis of *D. radiodurans* undescribed genes, mainly related to radiation and desiccation stresses. Altogether, these genes/functions may harbor unique roles and regulators engaged on resistance against extreme conditions on Earth, and, maybe, in other places in the Solar System.

2. Exclusive functions accounting for cell survival in *D. radiodurans*

Several mechanisms described in *D. radiodurans* account for its extreme resistance against many types of biologically harmful agents. They include protein protection against denaturation (Daly et al. 2010), different chemical balance of redox ions (Daly et al. 2004), and efficient DNA repair systems (see below) (Blasius et al. 2007; Harris et al. 2004; Minton & Daly 1995), besides some morphological and biochemical features as well (Shukla et al. 2014).

Protein oxidative damage is the main cause of cell death after irradiation, as previously observed (Daly et al. 2007). Anti-oxidant agents comprise an important toolkit contributing to radiation survival. The cell itself is susceptible to radiation, but such protective agents have been described in the last 20 years as shielding the whole cellular system (Horikoshi 2011). Inspection of the cell wall structure revealed interesting features that could contribute for such resistance, as the presence of carotenoids, playing a secondary anti-oxidant role in protecting membranes, coping with IR-generated radicals (Farci et al. 2014). High intracellular Mn(II) content reinforces the classical superoxide dismutases and catalases activities described in *D. radiodurans*, thus strengthening the role of the overall protein protection system (Daly et al. 2004). Mn(II) substitution for Fe(II) as enzyme metallic cofactor appears to protect redox enzymes, thus minimizing the pro-oxidant Fenton Reaction to occur (Ghosal et al. 2005). Hence, Mn(II) influx is thought to contribute to *D. radiodurans* high resistance against IR. Moreover, molecular modeling employing bioinformatics tools revealed new connections in the Mn(II) network with predicted metal-binding proteins and DNA-damage responsive agents, generating an interactive pathway between Mn(II) complexes and DNA repair machinery (Peana et al. 2018).

The protein protection activity seems to exert an important function in this species, since recent evidence of a shielding effect was observed when *D. radiodurans* cell extracts protected *E. coli* and even human cells from the deleterious IR consequences (Daly et al. 2010). This last feature depicts how many mechanisms are involved in the remarkable radioresistance of *D. radiodurans*.

While particular proteins participate in specific antioxidant metabolism as OxyR (Yin et al. 2010), small peptides were also described as promising candidates for explaining radioresistance in *D. radiodurans* due to their ability to protect complex proteins against damage by Reactive Oxygen Species (Peana et al. 2016). *ppr*M is a putative cold shock related protein (Ohba et al. 2009; Park et al. 2016) which was recently associated to the regulation of antioxidant metabolism through the biosynthesis of deinoxanthin and concentration of metal ions (Zeng et al. 2017).

D. radiodurans lacks visible light-induced Photoreactivation Repair of UV-induced damage. Instead, an important repair activity is expected to count on a specialized endonuclease, the UvsE, a functional ortholog of the UVDE (*UV D*NA damage *E*ndonuclease) from *S. pombe* (Earl et al. 2002), which repair depends on manganese ions (Evans & Moseley 1985). Considering Base Excision Repair (BER), *D. radiodurans* seems to possess more BER-related enzymes than other known bacteria, five among those encode thermophilic Uracil DNA Glycosylases and a typical Uracil DNA N-Glycosylase (drUNG) (Pedersen et al. 2015). The repair activity has been characterized for these enzymes, and the structure of *M*ismatch-specific *U*racil DNA N-*G*lycosylases (MUG) were already determined (Blasius et al. 2007).

Bulky damages are continuously removed from *D. radiodurans* DNA by means of the Nucleotide Excision Repair (NER), extremely conserved among prokaryotes. It is accomplished by the complex formed by UvrABC proteins, which creates incisions at 5' and 3' sides of a number of structurally unrelated damages including pyrimidine-pyrimidone dimers (Blasius et al. 2008; Tanaka et al. 2005).

A set of enzymes belonging to the Recombinase (Rec) family account for a significant amount of DNA repair (Blasius et al. 2008). RecA, for instance, has been described as a key protein for Homologous Recombination Repair (HRR) and extended Synthesis-dependent Strand Annealing (SSA), both implicated in full sealing of DNA strand breaks (Blasius et al. 2008).

Even though the abovementioned mechanisms exist in other bacteria, the way them seem to operate in *D. radiodurans* are not observed in other species. Additionally, exclusive *D. radiodurans* proteins PprA and DdrB seem to be vital in restoration of DNA integrity because they appear to drive other enzymes such as DNA ligase and DNA-binding activities, respectively (Narumi et al. 2004; Norais et al. 2009). Radiation-induced responses regulators include PprI, DdrO and DdrA, with DdrA playing a central regulatory agent in the cell recovery process (Liu et al. 2003a; Tanaka et al. 2004; Zhang et al. 2005).

3. The *D. radiodurans* way of facing extremes

After exposure to gamma radiation, many genes were overexpressed, and others were expressed upon exposure to both radiation and desiccation (Liu et al. 2003b; Luan et al. 2014; Tanaka et al. 2004). Several new members of the *Ddr* gene family were thus detected and their functions were described later (Selvam et al. 2013). For instance, the *ddr*B gene product seems to have a role as a single-strand DNA binding protein, an important group of elements in the DNA metabolism. Genomic re-annotation and studies on protein structural determination were conducted since then for a number of *D. radiodurans* genes (Norais et al. 2009; Tanaka et al. 2004).

At least 32 genes were seen to be induced as belonging to a common pathway acting against desiccation and IR stresses, but the genetic sequences of only a few of these have significantly matched to proteins that putatively participate in DNA repair in other species. The majority of these genes are likely to encode proteins of unknown function that could participate in repair or damage-prevention mechanisms ensuing resistance against many cellular stresses (Tanaka et al. 2004). A group of induced genes detected in previous reports were functionally evaluated through determination of phenotypes from knocked-out mutants. These data were obtained from gene expression analyses and later validated through loss-of-function experiments, in which the correspondent mutant strain was exposed to UV or gamma radiation and its sensitivity determined (Harris et al. 2004; Norais et al. 2009; Selvam et al. 2013; Xu et al. 2010).

Such proteins were assigned to belong to recovery responses as DNA repair, and other related metabolic functions that can enlighten new concepts about novel strategies in extremophilic species. For instance, an ATP-dependent DNA ligase encoded by the *DR_2069* gene is probably the main DNA ligase activity during the response to IR together with a rare 3'→5' RNA ligase function to recover cellular RNAs (Liu et al. 2003a). Notwithstanding the presence of similar repair proteins related to DNA repair in *E. coli* and *D. radiodurans*, the collected data suggests the way in which each species responds to IR stress to be completely different (Makarova et al. 2001). *D. radiodurans* harbors a new kind of regulon distinct from that observed in *E. coli* that is activated during damage response and the most probable transcriptional regulators are encoded by *DR_0171* and *DR_2574* gene products, with a RecA-like control of expression (Liu et al. 2003a).

IR-induced proteins linked with DNA repair, and oxidative stress alleviation responses were shown to be over-expressed after exposure to IR. Enzymes involved in protein expression and protection, and housekeeping functions also rose during this period. Altogether, IR-induced changes in *D. radiodurans* are represented by a metabolic shift towards enhancing the levels and activating of DNA repair proteins and the recycling of several damaged peptides (Basu & Apte 2012; Joshi et al. 2004; Zhang et al. 2005).

The survival against other harsh conditions such as high-energy UV (Paulino-Lima et al. 2010, 2011) and exposure to vacuum (Bauermeister et al. 2011) is also being explored to reveal singular features of this species. Once any repair process needs to be fed by a fuel source, a study on metabolic shifts upon stresses was conducted prior to the Japanese Tanpopo space mission, reporting a specific tendency for the enhancement of energy producing enzymes and abundances of metabolites following UVC and vacuum exposure, besides expected changes in gene expression activators, DNA damage response agents and antioxidant proteins (Ott et al. 2017).

4. Sheding light on the proteins of unknown function

Deinococci species remain as an isolated Thermo-Deino phylogenetic group, despite the expressive number of completed genome sequencing of bacterial genomes (Schulz et al. 2017) (Figure 1). It means that in *D. radiodurans* genome there are so many unique genes that it cannot be grouped together with any relative bacterial group. This fact highlights the need for a comprehensive advance on stress resistance in this species, which would explain more of the robustness of *Deinococci* species. Some genes with no homology to any other known similar can only be detected when *D. radiodurans* cells were submitted to various stressing conditions. The approach envisaged by this group intended to provide functional predictions for exclusive genes of *D. radiodurans*, capable of act upon exposure to beyond-limits of biological resistance, as the submis-

Table 1. Function prediction for a group of 20 gene products overexpressed after exposure to stressful conditions in *D. radiodurans* that were ascribed to "unknown function" in the genome of the species. Their functions were predicted by comparison with analog structures present in proteins of known function in other species.

Simple Modeling	Complex Modeling
DR0438 – DNA Ligand	DR0025 – Acylase
DR0491 – HSP31	DR0052 – Aldolase TIM barrel
DR1263 – N-glycosidase	DR0219 – Isomerase
DR1314 – Photosystem Subunit H	DR0227 – YmcC Protein
DR1370 – Outer-membrane lipoprotein carrier protein (IolA)	DR0893 – ATP:cobalamin adenosyltransferase
	DR1143 – Heat Resistant ATPase
DR2073 – Shikimate kinase	DR1264 – Acetyl transferase
DR2441 – N-acetyl transferase	DR1315 – Transcriptional Regulator
	DR1654 – 3-oxoacyl reductase
	DR2143 – Stomatin
	DR2309 – Thioredoxin
	DR2414 – Guanylate kinase
	DR2563 – Chitinase

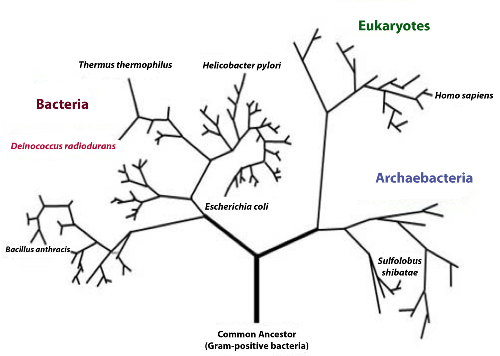

Fig. 1. A summarized, typical phylogenetic distribution of major biological groups with the *Deinococcus* species placed, up to date, as an individual group with no genetic affiliation with other bacterial groups.

sion to extraterrestrial-simulated or real conditions.

To evaluate that, our group queried these specific sequences for multiple alignments using Blastp[1] considering E-values below the cutoff of 1×10^{-4} and protein identity above 40%. Protein sequence alignments that did not match any other species' sequences under those criteria were tagged as exclusive of *D. radiodurans*.

Gene products listed in Table 1 indicate that several protein sequences remain absolutely novel, except for *DR_1654* and *DR_2309* which were related to existing matches in the Blastp database, although alignments referred to hypothetical proteins of *Deinococcus apachensis* and *Deinococcus wulumugiensis*. Proteins belonging to those *taxa* frequently align with *D. radiodurans* proteins, which corroborate to the hypothesis of a close phylogenetic group. This also raises a question concerning the effectiveness of alignments and sequencing against truly exclusive sequences of this species. If new generation sequencing and databases curation continue enlarging for a decade more, would those exclusive targets start to align to other species? This is a perspective to have in mind, when considering the unique genetics of a given species or group.

These data show that *D. radiodurans* harbors a genetic pool of novel functions for future studies and trials, useful for prospecting biotechnological tools and possibly enhancement of radioresistance in other species, ultimately mammals.

[1] `https://blast.ncbi.nlm.nih.gov/Blast.cgi`

5. Concluding remarks

D. radiodurans has a powerful defense toolbox against stresses that was unveiled through sequence alignments, taxonomy, and bioinformatic analyses. They appeared highly induced by IR and desiccation stresses, but with functions awaiting elucidation. Additionally, a group of genes do not seem to change their expression at all. A few genes were not seen to be expressed under the tested stresses of radiation and desiccation. They possibly remain unique to the lifestyle of *D. radiodurans*, exerting novel functions on yet non-described resistance mechanisms, to stresses other than IR or desiccation. In order to characterize these proteins, new approaches should be considered. Solving the tridimensional structure of proteins of unknown function *in silico* can save time for experimental characterization of their possible roles. Moreover, investigation on phylogenetic relationships may bring new thoughts on the origin of several metabolic features found in *D. radiodurans*. Work in progress of this group is currently exploring the remaining annotated gene sequences exclusive to *D. radiodurans*, belonging to what we addressed as the "black box genome", for their potential role on extremophilic features of *D. radiodurans*, operating, for instance, under extraterrestrial-simulated or real conditions as in the case of experiments conducted at the International Space Station.

References

Basu, B., & Apte, S. K. 2012, Mol. Cell. Proteomics, 11, M111.011734
Bauermeister, A., et al. 2011, Microb. Ecol., 61, 715
Blasius, M., et al. 2007, BMC Molecular Biol., 8, 69
Blasius, M., Hübscher, U. & Sommer, S., 2008, Crit. Rev. Biochem. Mol. Biol., 43, 221
Byrne, R. T., et al. 2014, Elife, 3, e01322
Cox, M. M., & Battista, J. R. 2005, Nat. Rev. Microbiol., 3, 882
Cox, M. M., Keck, J. L., & Battista, J. R. 2010, PLoS Genet., 6, e1000815
Daly, M. J. 2009, Nat. Rev. Microbiol., 7, 237
Daly, M. J., et al. 2004, Science, 306, 1025
Daly, M. J., et al. 2007, PLoS Biol., 5, e92
Daly, M. J., et al. 2010, PLoS One, 5, e12570
Earl, A. M., et al. 2002, Journal of Bacteriology, 184, 1003
Evans, D. M., & Moseley, B. E. 1985, Mutat. Res., 145, 119
Farci, D., et al. 2014, Biochim. Biophys. Acta, 1838, 1978
Ferreira, A. C., et al. 1999, Extremophiles, 3, 235
Ghosal, D., et al. 2005, FEMS Microbiol. Rev., 29, 361
Griffiths, E., & Gupta, R. S. 2007, Int. Microbiol., 10, 201
Harris, D. R., et al. 2004, PLoS Biol., 2, e304
Horikoshi, K. 2011, Extremophiles Handbook (Springer, Tokyo)
Joshi, B., et al. 2004, Biochem. Biophys. Res. Commun., 320, 1112
Krisko, A., & Radman, M. 2013, Cold Spring Harb. Perspect. Biol., 5, a012765
Lipton, M.S., et al. 2002, Proc. Natl. Acad. Sci. USA, 99, 11049
Liu, Y., et al. 2003a, Proc. Natl. Acad. Sci. USA, 100, 4191
Liu, Y., et al. 2003b, Proc. Natl. Acad. Sci. USA, 100, 4191
Luan, H., et al. 2014, PLoS One, 9, e85649
Makarova, K. S., et al. 2001, Microbiol. Mol. Biol. Rev., 65, 44
Makarova, K. S., et al. 2007, PLoS One, 2, e955
Minton, K. W., & Daly, M. J. 1995, Bioessays, 17, 457
Narumi, I., et al. 2004, Mol. Microbiol., 54. 278
Norais, C. A., et al. 2009, J. Biol. Chem., 284. 21402
Ohba, H., et al. 2009, Extremophiles, 13. 471
Ott, E., et al. 2017, PLoS One, 12, e0189381
Park, S. H., et al. 2016, Frontiers in Microbiology, 7, 2124
Paulino-Lima, I. G., et al. 2010, Planetary and Space Science, 58, 1180
Paulino-Lima, I. G., et al. 2011, Astrobiology, 11, 875
Peana, M., et al. 2016, J. Inorg. Biochem., 164, 49
Peana, M., et al. 2018, J. Trace Elem. Med. Biol., 50, 465

Pedersen, H. L., et al. 2015, Acta Crystallogr. D Biol. Crystallogr., 71, 2137
Rainey, F. A., et al. 2005, Appl. Environ. Microbiol., 71, 5225
Schlesinger, D. J. 2007, FEMS Microbiol. Lett., 274, 342
Schulz, F., et al. 2017, Microbiome, 5, 140
Selvam, K., et al. 2013, PLoS One, 8, e69007
Shukla, S. K., et al. 2014, Curr. Microbiol., 68, 247
Slade, D., & Radman, M. 2011, Microbiol. Mol. Biol. Rev., 75, 133
Tanaka, M., et al. 2004, Genetics, 168, 21
Tanaka, M., et al. 2005, J. Bacteriol., 187, 3693
White, O., et al. 1999, Science, 286, 1571
Xu, G., et al. 2008, J. Bacteriol., 190, 2624
Xu, G., et al. 2010, DNA Repair (Amst), 9, 805
Yin, L., et al. 2010, J. Microbiol., 48, 473
Zeng, Y., et al. 2017, Indian J. Microbiol., 57, 316
Zhang, C., et al. 2005, Proteomics, 5, 138