universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Transcriptomic analysis of *Deinococcus spp*. after low earth orbit exposure outside of the ISS for 3 years. "

verfasst von / submitted by

Natalie Özgen, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2021, Vienna 2021

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 877
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Genetik und Entwicklungsbiologie
Betreut von / Supervisor:	Doz. Dipl.-Biol. Dr. Hans-Jürgen Busse

.

Abstract

Deinococcus radiodurans exhibits metabolic and genetic features that enables the survival under extreme conditions. However, the underlying molecular mechanism of its resilience are poorly understood. To understand how it reacts to outer space exposure, dried cells were exposed to low earth orbit conditions on the outside of the ISS for 3 years in the curse of the Tanpopo mission. Subsequently, a transcriptomic analysis of the0 samples was conducted 5 and 15 hours after cell revival and compared to control samples that stayed on earth for the same amount of time.

After 5 hours of revival, several genes encoding for transporters, radical oxygen species scavengers like *SodC* as well as enzymes of the base excision repair pathway experienced highly significant upregulation. Unexpectedly, relative upregulation of the RecFOR DNA repair pathway was not a defining feature of the ISS induced stress response. Furthermore, many of the upregulated transcriptional regulators contained GGDEF domains, domains that induce the synthetization of the second messenger cyclic diGMP. This indicates that the transcriptional response was largely facilitated by second messenger mediated signalling. Interestingly, many of the uncharacterized upregulated transcripts contained intrinsically disordered protein domains which lack a 3D-structure. They are proposed to act in signaling pathways and as protective protein chaperons. Finally, cells revived for 15 hours displayed high transcriptional upregulation of peptide transporters and proteases, suggesting that elevated proteolytic activity is pivotal for cell recovery after ISS exposure.

Zusammenfassung

Deinococcus radiodurans besitzt eine Reihe von genetischen und metabolischen Besonderheiten, die das Überleben selbst unter extremen Lebensbedingungen ermöglichen. Die molekularen Mechanismen seiner Resilienz allerdings weitgehend unerforscht. Um seine Reaktion auf Weltraumstress zu erforschen, wurden während der Tanpopo Mission dehydrierte Proben 3 Jahre lang außerhalb der ISS gelagert. Danach wurde nach 5 beziehungsweise 15 Stunden Erholung eine Transkriptions-analyse durchgeführt und mit Kontrollproben verglichen, die auf der Erde verblieben.

Nach 5 Stunden Erholung fand man vor allem hochregulierte Gene vor, die für Transporter Proteine, Sauerstoffradikal-Abbau Proteine wie z.B. *SodC* und Enzyme des Basen-Exzisonsreparatur Mechanismus kodieren. Überaschenderweise wies der Organismus allerdings keine relative Überexpression der Enzyme des RecFOR DNA-Reparatur Mechanismus auf. Des Weiteren fand man Überexpression von Transkriptionsfaktoren, die GGDEF-Domänen enthalten, eine regulatorische Domäne, die zur Bildung des Signal-Moleküls zyklisches diGMP beiträgt. Dies deutet darauf hin, dass diese „second-messenger“ Signal-Moleküle vermutlich eine wichtige Rolle bei der Einleitung der transkriptionellen Stressreaktion gespielt haben. Außerdem enthielten viele der un-charakterisierten, hochregulierten Gene Domänen intrinsisch ungeordneter Proteine, welche sich durch das Fehlen einer 3D-Struktur auszeichnen und deren Aufgaben in der Weiterleitung von Signalen und dem Schutz von Proteinen vermutet werden. Zellen, welche 15 Stunden lang anwuchsen, zeichneten sich vor allem durch Hochregulierung von Peptid-transportern und Proteasen aus. Dies deutet darauf hin, dass erhöhte proteolytische Aktivität vermutlich eine tragende Rolle für die Zell-Reparatur nach der ISS-Exponierung spielt.

4. ACKNOWLEDGEMENTS

Firstly, I would like to thank Dr. Hans Jürgen-Busse, whose guidance and advice made the completion of this work possible. Furthermore, I would like to thank my colleagues Denise and Emanuel for all the fun and inspiring times we spend inside and outside the lab. Finally, I´d like to thank my parents who gifted me with nature and nurture as well as my siblings Alper and Mert and my friends who always lifted me up whenever I needed it.

TABLE OF CONTENT

1. INTRODUCTION

1.1 Survival in outer space- the biology of extremophiles

As mankind continues to explore the realms of outer space, understanding biochemical, molecular, and physiological responses of terrestrial life to outer space becomes increasingly important. Outer space represents a very challenging environment to all known lifeforms. The lack of the earth´s atmospheric shield and it´s protective magnetic field leads to a bombardment of tissues and cells by solar radiation and high energy galactic cosmic rays (GCR) [1]. Another form of stress for terrestrial organisms is the vast predominance of "nothingness" in the form of vacuum. Subjected to space vacuum and decoupled from all gravitational forces, fluids and gas instantly decompress due to a lack of ambient pressure. Due to a lack of ambient pressure, they rapidly boil and evaporate in a process called de-sublimation, leading to ruptures of tissues and cells by embolism [2] [3]. Moreover, the vacuum filled space hardly contains any molecules that could potentially contribute to life preserving physiological mechanisms of a terrestrial, carbon based life form [4]. However, some organisms were found to persevere in this hazardous environment.

Polyextremophile organisms, like the eukaryotic tardigrades [5], *Bacillus subtilis* [6] and lichens like *Xanthoria elegans* [7] developed efficient strategies to cope with extreme temperatures and desiccation on earth. Desiccation damage inflicted on macromolecules like proteins or DNA is similar to damage induced by commonly found stress sources in outer space [8]. Hence, the development of outstanding desiccation resistance has been proven to be beneficial for the survival in outer space. Tardigrades and *B. subtilis* both rely on a similar strategy, that is they enter a state of extreme resilience. Upon unfavorable environmental conditions, the gram-positive bacterium *B. subtilis* forms protective endospores and keeps this state of metabolic dormancy until the environmental conditions are more favorable. During this state of dormancy, a multilayered endospore engulfs the DNA. Furthermore, the cells facilitate the incorporation of dipicolinic acid which stabilizes DNA and proteins. Additionally, the formation of a peptidoglycan containing cortex provides additional protection against extreme temperatures and UV radiation [9] . In tardigrades, the protective state is called cryptobiosis. During cryptobiosis, all metabolic activities

come to a halt and the water content drops to 1% of the usual content. It is initiated by desiccation, oxygen deficiency and extreme temperatures [10]. The extreme resilience of this organism has been largely attributed to the buildup of the sugar trehalose, however, recent studies suggested intrinsically disordered proteins (IDP) as an important protective unit [11]. In tardigrades, this enigmatic class of proteins which fails to retain a fixed 3-D structures is proposed to form a glass-like amorphous matrix in a process called vitrification. Subsequently, this matrix presumably traps sensitive macromolecules and physically protects them from aggregation, denaturation or disintegration. In contrast to tardigrades and *B. subtilis*, *Deinococcus radiodurans* does not rely on the induction of a dormant metabolic state to achieve its outstanding capabilities to withstand detrimental surroundings. Instead, it uses a combination of extremely efficient ROS (radical oxygen species) scavengers, unique cell wall features and an efficient DNA damage repair system, as well as a set of highly redundant genes for cell sanitation and other relevant processes to alleviate damage and ensure regeneration (Slade *et al.*, 2011). The current state of knowledge about *D. radiodurans´* coping strategies of is discussed in more detail in later chapters.

Overall, studying extremophiles and understanding their responses to these detrimental stress sources serves both fundamental science as well as application-oriented purposes. In the context of astrobiology, it helps to understand theoretical problems and questions i.e., the likelihood of interplanetary transport of life or the identification of molecular footprints of potential extraterrestrial organisms. However, it also opens prospects and concepts for new biotechnological approaches. Furthermore, understanding microbial outer space survival addresses medical concerns of spacecraft missions. It can aid in the development of sterilization techniques as well as prevention of contaminating extraterrestrial landscapes and understanding possible changes in the crew´s microbiome. For example, prior studies have revealed a change in microbial stress response behavior when on board of a space craft. During space flights, bacteria have been shown to exhibit elevated growth behavior [13] and decreased susceptibility to antibiotics [14] [15]. Furthermore, studies investigating the response of *Pseudomonas aeruginosa* to space flight found an induction of certain virulence associated genes and enhanced levels of biofilm formation which could potentially cause corrosion and water contamination [16] [17]. Taken together, these observations highlight the interesting biological underpinnings

that enable survival under outer however, more investigations are needed to understand them.

Understanding the biology of polyextremophiles could enable the usage of the molecular mechanisms for biotechnological advances. Overall, Space technology has often been the source for innovative technological applications on earth. Many everyday objects, such as LED lamps, Teflon, solar cells, or memory foam were originally invented by NASA for space craft related applications but have found their way into our homes. Similarly, understanding the molecular basis of the extraordinary extremophilic abilities of e.g., tardigrades, *B. subtilis* or *D. radiodurans* could endow humanity with a set of useful molecular tools which can be used to produce biofuel, to improve the longevity of pharmaceutical products or to generate multi-resistant strains for bio-leeching and medical purposes. For instance, in a previous study strains of *D. radiodurans* were generated which expressed resistance gene against the highly toxic ion Hg (II) (merA) [18]. The strains were able to effectively reduce toxic Hg (II) to less toxic elemental mercury when exposed to nuclear wastelands. Concerning medical applications, a study reported protective effects of ultrafiltrate from *D. radiodurans.* When the filtrate was intravenously injected into mice prior to the exposure to an otherwise lethal dosage of ionizing radiation, the mice were able to survive [19]. Furthermore, an exopolysaccharide component of the cell wall of *D. radiodurans,* DeinoPol, was shown to exert protective effects in human keratinocytes against oxidative stress, supposedly by acting as ROS scavenger [20]. Another possible application in medicine could be the application of novel secondary metabolites as chemotherapeutic reagents. A study conducted in 2020 demonstrated anticancer activity of *D. radiodurans* crude cell extracts in breast cancer cells. They attributed the chemotherapeutic effect to bioactive secondary metabolites [21]. Additionally, the features of *D. radiodurans* could be used to generate genetically modified crops resilient to various stress sources. Considering progression of climate change, crops resistant to desiccation and other adverse environments can be a powerful tool to fight the agricultural challenges which are posed by global warming. Although many of these applications still are in the early stages of development and need to be refined for further usage, they showcase the great innovative power that resides in this organism.

1.2 Environmental conditions of low earth orbit

Multiple studies have been conducted in a collective effort to test the boundaries of survivability in outer space. The multi-user EXPOSE facility onboard of the ISS (International Space Station) played an important role to enable the exposure of terrestrial organic samples to outer space conditions. The ISS orbits the thermosphere in an altitude ranging between 330 km to 420 km above the Earth's surface [16]. This altitude range is also known as Low Earth Orbit (LEO). The thermosphere provides a multitude of stress sources for terrestrial life. At this altitude, temperatures are extremely variable as they are highly dependent on the occurrence of solar radiation and the exact location of the spacecraft. Hence, within a 90-minute orbit around the earth, the ISS is exposed to both extremely high temperatures and low temperatures. They can reach up to 120°C when facing the sun and drop down to around -120°C in the earths´ shadow [22]. However, the temperature appears to be cold to microbes and other living organisms due to the lack of molecules in the air which could transfer the heat [23]. Furthermore, the gravitational forces of LEO ranging from 10^{-3} to 10^{-6} g can be almost equally as low as they are in deep space. Moreover, in LEO vacuum pressure reaches down to 10^{-7} to 10^{-4} Pa, presenting another potential source of stress to exposed organisms.

However, one of the most detrimental threats to all known organic molecules is posed by radiation. Without the protection of an atmospheric shield, organic matter is constantly confronted with damaging radiation of galactic and solar origin. Galactic cosmic rays are born outside the solar system through astronomical events such as supernova explosions and consist of high energy protons, helium and other high z-energy ions (HZE). Although HZE constitute only 1% of the GCR, HZE are a major concern for the viability of biological samples. The ions are able to penetrate through thick walls of shieling, depositing energy along its trajectory and leaving behind a path of destruction [24] [25]. The absorptions of energy by molecules such as proteins or nucleic acids produces highly reactive radicals for example via the radiolysis of cellular water. In microbes in particular, HZE can induce DNA double strand breakage or lead to oxidative base damage [26]. However, exposure level of cosmic ionizing radiation is highly variable depending on altitude, solar cycle, shielding. These

variables lead to enormous fluctuations, ranging from an annual dose of 1 to 10.000 gray [27] [28].The second major source of radiation originates from the center of our solar system, the sun. The sun emits hazardous ionizing radiation in the form of high energy X-rays and gamma rays. Moreover, the sun is the source of UV-C (with a spectral range of 200 to 280 nm) and UV-B (280 to 315 nm) rays which are particularly detrimental as well as UV-A (315 to 400 nm). On earth, the stratospheric ozone layers form a protective shield against UV rays, effectively absorbing UV radiation shorter than 290 nm before they can reach closer to the surface and cause damage. However, in LEO this protective layer is missing. Biomolecules are highly susceptible to ionizing radiation and only few organisms are able to withstand higher dosages for a prolonged period of time. On a molecular scale, ionizing radiation causes damage via the emission of high-energy particles which are able of displacing atomic electrons. This event can generate ROS such as hydroxyl radicals (•OH) and ionized water (H_2O^+), as well as hydrogen radical (H•), superoxide ($O_2^{\bullet -}$) and hydrogen peroxide (H_2O_2). All these oxygen containing species are highly reactive with biological macromolecules such like lipids, proteins or nucleic acid, leaving a trail of cellular destruction. In DNA, the newly emerged ROS can lead to single and double stranded breakage (SSB/DSB), cross-linking, and alteration of nucleic bases and the sugar backbone. These events can result in mutations and altered protein function and ultimately in cell death [29]. Another major site of radiation damage are lipidic structures. Especially ROS generated by water radiolysis seem to induce oxidative degradation of cellular lipids. Consequently, the disturbance of lipids can result in increased membrane permeability, can modify the function of membrane associated proteins, and impede ion gradients [30].

Finally, radiation induced ROS can damage proteins in a multitude of ways such as protein oxidation and desulfurization. Thus far over 35 oxidative protein modifications are known, such as the •OH induced oxidative cleavage of the protein backbone or amino acid side chains, carbonylation, and a covalent addition of oxidized lipid byproducts to amino acids [31]. In fact, studies suggested that in in the extremophile *D. radiodurans* not DNA but proteins are the primary targets of irradiation [32]).

1.3 The metabolic features of *D. radiodurans*

An interesting candidate for outer space related studies and applications is the exceptionally radioresistant bacterium *D. radiodurans.* The microbe can withstand an acute ionizing radiation dosage of 5.000 grays (Gy) with only miniscule loss of viability. Hence, the Gram positive non-pathogenic bacterium even exceeds the radiation resistance of tardigrades (4.000 Gy) [33]. The major detrimental effect of ionizing radiation is inflicting direct damage on macromolecules on one hand and the build-up of ROS on the other [34] [35]. Accumulated ROS have the potential to destroy important cellular components. In humans ROS are suspected to be the cause of many degenerative diseases like Alzheimer´s disease [36], Parkinson´s disease [37], autism [38] and cancer [39].

Thus far, the abilities of *D. radiodurans* to withstand high irradiation are thought to be a concomitant consequence of its ability to survive prolonged desiccation [8]. Its remarkable sturdiness is expected to stem from a combination of different features. Even though *D. radiodurans* is Gram positive, its cells wall is reminiscent of Gram negative bacteria in regards to its multilayered organization [40] [41]. Furthermore, the microbe harbors some metabolic and genetic properties which help the microbe to retain its viability even under highly adverse circumstance. Firstly, *D. radiodurans* can circumvent ROS production by inducing the glyoxylate bypass, a metabolic alternative to the tricarboxylic acid (TCA) cycle [42]. This pathway uses a smaller number of oxidation sensitive respiratory enzymes with iron-sulfur clusters, thus circumventing endogenous accumulation of metabolic ROS.

Secondly, the bacterium acquired a highly economic and redundant system of recycling cellular building blocks and sources of energy. For example, it uses proteolysis of surrounding peptides and nucleic acids and subsequently imports them by employing ABC transporters [43] [44]. Furthermore, studies have shown an increase in intercellular proteolytic activities upon irradiation [45]. It is speculated that intracellular and extracellular proteolysis might be induced to recycle proteins which were damaged during irradiation or to cannibalize on cells that did not survive the stress. This efficient recycling of proteins could serve extremophiles in two ways: firstly, it could decrease biosynthetic demands, secondly they could be employed for the generation of ani-oxidative protein complexes [32]. For instance, in *D. radiodurans*

peptide-manganese complexes have been identified which are suspected to shield essential proteins from oxidation [45]. Furthermore, the bacterium seems to exploit glucose to recover from DNA damage by utilizing the pentose phosphate pathway (PPP) [42]. Studies monitoring the activity of the activity of glucose-6-phosphate dehydrogenase (G6PDH), the marker enzyme for PPP, demonstrated that *D. radiodurans* displays a PPP activity 4-fold higher than in *Escherichia coli* under normal circumstances [46] and an additional upregulation in irradiated cells [47]. These studies suggest, that the sugar processed during PPP might be converted into ROS scavenging complexes or dNTPs which are further incorporated into the DNA to mend lesions [32]. Moreover, *D. radiodurans* was shown to build up granules consisting of carbohydrates and polyphosphate which can be used as energy storage and for the production of protective manganese-phosphate complexes[48]. Taken together, it is likely that these metabolic properties act in synergy to accomplish the multi-factorial resistance that is displayed by *D. radiodurans*.

1.4 DNA damage repair pathways in *D. radiodurans*

One of the main primary targets of radiation is DNA. Like desiccation, ionizing radiation induces single strand breaks (SSB), double strand breaks (DBS) and base damage whereas UV radiation primarily generates pyrimidine dimers [49]. A dosage of 7 kGy ionizing radiation shatters the 3.25 Mb genome of *D. radiodurans* into 20 to 30kb sized fragments by introducing on average 100–150 double-strand breaks (~0.02 DSBs/Gy/genome) and approximately 10 times as many SSB with only a miniscule loss of viability [12]. Before the damaged chromosome is patched back together via complex repair mechanism, *D. radiodurans* undergoes a phase of extensive DNA degradation without any signs of reparation for 1.5 hours [50]. Although contra-intuitive at first, DNA degradation seems to be an integral part of DNA recovery. Recombination-deficient *recA* and *recFOR* mutants not only fail to reassemble of the genome, but also lack the preceding phase of DNA degradation [12] [51]. This initial phase of degradation seem to be primarily caused by exonucleolytic activity at the protruding DNA ends and by base pair excision repair (BER) before DNA synthesis and genomic reassembly is eventually initiated [52] [53].

When it comes to DNA repair, *D. radiodurans* displays an exceptional redundancy in terms of DNA repair genes. According to gene sequencing, the bacterium encodes 11 DNA glycosylases, enzymes that catalyze the removal of mutated bases during BER, and 23 genes encoding Nudix ("nucleoside diphosphate linked to some other moiety x") family hydrolases, which are involved in the degradation of oxidized nucleotides, thereby preventing their re-incorporation [54] as well as 2 UV repair pathways, namely UvrABC and UV damage endonuclease UVDE [54]. To reassemble the shattered genome and mend SSB and DBS, *D. radiodurans* developed complex pathways which are still not completely understood. Thus far, studies suggest that the bacterium uses the RecFOR pathway upon DNA damage, a pathway of homologous recombination in prokaryotes. The repair of DBS seems to rely on two phases: first "extensive synthesis-dependent strand annealing" (ESDSA) is initiated at the site of damage to create large 3´overhangs, in the second phase RecA protein-mediated double-strand break repair is induced to anneal the products from ESDSA (Fig.1 A). The main components of the current model for the repair process are the enzymes RecA (DR2340), RecF (DR1089), RecO (DR0819), RecR (DR0198), and RecJ (DR1126) [51] . At the beginning of the pathway the exonuclease RecJ and the helicase UvrD (DR1775) presumably bind sites harboring double strand breaks [32]. To subsequently produce single stranded 3´overhangs, UvrD unwinds the double helix to enable the digestion of the 5′ end by RecJ. The remaining 3´ overhang is then bound by several single strand binding proteins, which have been shown to be highly induced upon ionizing radiation and are potentially in charge of protecting the bacterium against desiccation [55]. According to the current model of ESDSA, RecFOR complex is then recruited to the single stranded overhang, which in turn initiates loading of RecA and its homologue RadA to the site of damage [32]. RecA and RadA then presumably initiate strand invasion to form a D-loop structure and prime for DNA synthesis [51]. RecA mediated D-loop formation was shown to be promoted by RecN, a cohesion like protein, under the hydrolysis of ATP [56]. The long DNA overhangs are progressively synthesized by the invasion of the single stranded DNA strand overhang into a homologous duplex DNA which serves as a template for the elongation of said overhang [57]. This process happens on multiple damaged loci upon irradiation. Afterwards, the elongated single strands dissociate from the template and re-anneal to complementary single-stranded extensions, forming long intermediate

duplex DNA -patches. These long DNA “patches” form the scaffold for the new chromosome. To reconstitute the genome, the “patches” are stitched back together via RecA dependent crossover within overlapping homologues sequences and subsequent ligation on their homologous ends, forming a new functioning circular chromosome [50] [32].

The importance of the RecFOR complex for DNA repair was highlighted in a study on *recF*, *recO*, and *recR* mutants which displayed a dramatic loss of viability after gamma ray exposure due to reduced initial DNA degradation, abolished DNA synthesis and insufficient genome restauration [51]. Furthermore, RecA was found to be a key enzyme for homologous recombination as mutants are highly sensitive to oxidative damage [58], and displayed reduced viability upon exposure to ionization and UV radiation [59] [60]. In *D. radiodurans,* the *recA* locus (DR2340) consists of a polycistronic operon containing the upstream damage-inducible protein CinA together with a 2′-5′ RNA ligase (LigT)[61]. Usually expressed at constitutively low levels, RecA is highly expressed in *D. radiodurans* together with LigT upon DNA damage induced by radiation [43] [62] [63].

During the DNA repair by ESDSA, initiation of strand invasion is performed by DNA Pol III and subsequent elongation is either carried out by DNA Pol III or taken over by DNA Pol I in case of unrepaired damage sites in the template [12]. Overall, the contribution of Pol I to global DNA synthesis is negligible in D. radiodurans as synthesis is predominantly performed by Pol III [12]. Besides Pol I and Pol II *D. radiodurans* possesses a third polymerase called Pol X (DR0467). The enzyme is Mn^{2+}-dependent and equipped with a 3′-5′ exonuclease activity modulated by a stem-loop structure [64] and a 5′-deoxyribose phosphate lyase activity required for BER [65].

Apart from recA-dependent DNA damage repair pathways, recA-independent pathways have been proposed as well. In fact, one-third of the DS induced by ionizing radiation seem to be repaired by inducing recA-independent single strand annealing (SSA) reactions [58] [50]. SSA is a repair pathway that includes initial DNA breakage followed by exonucleolytic DNA degradation on the ends and annealing of the exposed complementary ssDNA overhangs [66]. In *D. radiodurans*, SSA was shown to be dependent on two deinococcal proteins called DdrA (“DNA damage response”) (DR0423) and DdrB. DdrA is highly induced upon radiation and desiccation stress [67],

and has been proposed to shield 3´ssDNA overhangs from degradation [68]. In fact, a study demonstrated that this protein protects 3´ single strand overhangs from degradation by E.coli exonuclease I [68], which indicates that DdrA might preserve the integrity of DNA fragments which are used for recombination. The protein proteins might have versatile functions, as it has been proposed to be involved in ESDSA as well: it could promote the recycling of RecA by enhancing the longevity of its DNA binding substrate [69]. DrdB (DR0070) displays single strand binding protein behavior with strand-annealing properties and is induced by irradiation [70]. Furthermore, recent studies revealed that DdrB facilitates the high fidelity annealing of ssDNA into a fully formed duplex by proofreading for complete strand complimentary [71]

Another specific repair protein unique to the *Deinococcaceae* species is PprA (DRA0346) which stands for pleiotropic protein promoting DNA repair protein [72]. PrpA can bind to dsDNA damaged by strand breaks, stimulates DNA end-joining reactions catalyzed by DNA ligases and is highly expressed after subjecting cells to ionizing radiation and desiccation [43] [73]. Furthermore, it has been shown to inhibit *E.coli* exonuclease III activity, suggesting a protective role against excessive DNA degradation on DNA ends [72].

Overall, *D. radiodurans* is 30-fold and 1.000-fold more resistant to ionization than *E.coli* and humans, respectively [74] [75]. Nevertheless, recent genetic analyses indicate that its resistance does not stem from extraordinary DNA repair capabilities [42] [76]. In fact, its DNA repair system seems even less complex than those found in *E.coli* or *Shewanella oneidensis* [77] [42]. Additionally, many of the most severe repair deficient mutants such as *polA* [78] or, in parts, *recA* [79] and *recO* [80] have been rescued with the respective *E.coli* genes. These findings indicate, that its remarkable resistance is not caused by exceptionally good DNA damage repair pathways but due to the level of protein protection and ROS scavenging it is exerting, although the existence of unknown special deinococcal DNA repair genes cannot be excluded [32]. Hence, the most well studied means of ROS scavenging and protein protection in *D. radiodurans* will be discussed in the next paragraph.

1.5 Antioxidative scavengers in *D. radiodurans*

The most common oxidative modification of proteins is carbonylation. It alters enzymatic activities and protein interactions, which subsequently can lead to cell death. For example, oxidation of DNA repair proteins can result in error prone activities which causes DNA mutations [81]. One of the arguably most prevalent sources of oxidative damage are ROS which are generated during desiccation, radiation and other stress sources. To prevent protein oxidation, *D. radiodurans* employs an elaborate and highly redundant antioxidant defense system in order to cleanse cells from ROS. The primary ROS relevant for damage control are hydroxyl radicals (OH·), superoxide radicals ($O_2{\cdot}^-$), and hydrogen peroxide (H_2O_2). Especially hydroxyl radicals which are caused by water hydrolysis upon ionizing radiation and by the Fenton or Haber-Weiss reaction are highly reactive to all macromolecules of the cell [45]. To ward of this offence, *D. radiodurans* encodes enzymatic ROS scavengers like superoxide dismutase, peroxidases and Dps (DNA-binding proteins from starved cells) proteins on one hand. ON the other hand, it produces divalent manganese complexes and carotenoids for non-enzymatic scavenging on the other. To remove superoxide radicals from the cells *D. radiodurans* harbors four superoxide dismutases (Mn-dependent DR1279 and Cu/Zn-dependent DR1546, DRA0202, and DR0644) [42]. Furthermore, enzymes like catalases and peroxidases remove H_2O_2. The bacterium encodes several types of catalases, namely katE catalases DR1998, DRA0259 and the eukaryotic-type DRA0146 catalase, a cytochrome c peroxidase (DRA0301) and an iron-dependent peroxidase (DRA0145) [42].

The most efficient ROS scavengers in *D. radiodurans* however, are manganese complexes. Across all species of bacteria, the element manganese seems to play an important role against oxidative damage. It has been shown that a high manganese-to-iron ratio correlates with elevated resistance against ionizing radiation and desiccation [77] [82] as well as decreased levels of oxidative protein damage [81]. Additionally, studies have provided evidence for a protective effect of manganese complexes on human cell lines against radiation induced cell death, opening up the possibilities of exciting applications in biotechnology and medicine [45]. With 0.2 to 4 mM per cell, *D. radiodurans* possess extraordinary amounts of intracellular manganese and a high intracellular manganese-to-iron ratio [77], deeming it well armed against radiation and ROS-induced damage. In their fight against ROS species divalent manganese (Mn^{2+}) ions can team up with different molecular entities. For

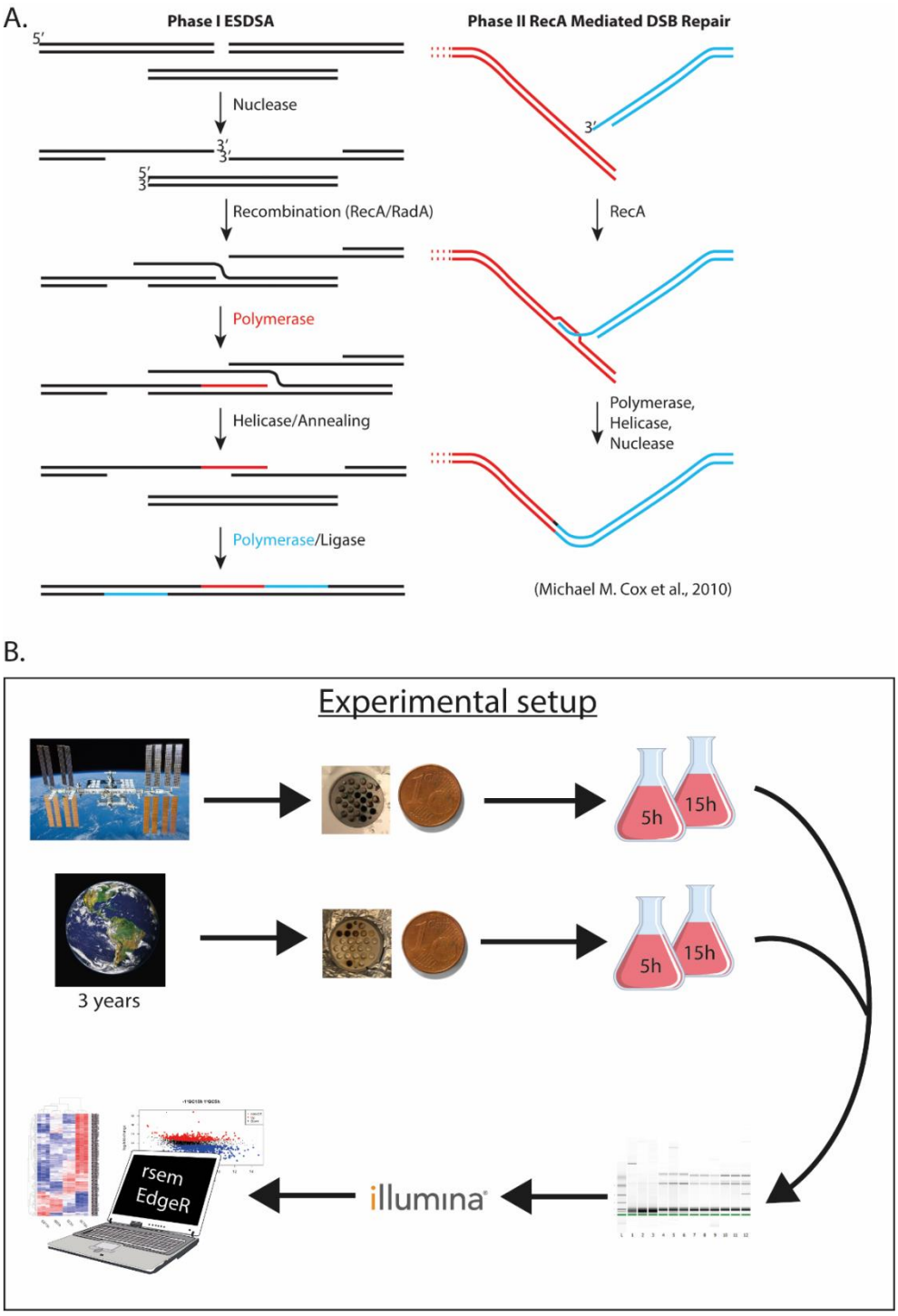


Figure 1. DNA damage repair in D. radiodurans and experimental setup

(A) Genome reconstruction happens in two phases. First ESDSA (extended synthesis DNA strand annealing) is induced whereby sites of double strand breaks get recessed and overhangs undergo multiple rounds of recA and helicase depending strand extension before annealing to a complementary overhang to form a DNA duplex intermediate. In the second phase the intermediate DNA double strands undergo DSB break repair to form a new circular chromosome. Dehydrated cell pellets were stored in aluminum plates and exposed to low earth orbit onboard of the ISS on the EXPOSE facility for 3 years during the Tanpopo mission in. Control pellets stayed on earth. After receiving the samples, cells were revived for 5 and 15 hours respectively and RNA was extracted. Then, illumina sequencing was performed and transcriptomic expression was evaluated using the software programs rsem and EdgeR.

instance, they are able to complex with orthophosphates, peptides, free amino acid and nucleosides in order to hunt down long lived $O_2 \cdot -$ and H_2O_2 or short lived (OH·) ROS, thereby preventing interconversion and propagation [45]. The majority of manganese forms complexes with orthophosphates and peptides, which scavenge $O_2 \cdot^-$ and H_2O_2 [45]. Some Mn^{2+} also bind nucleosides containing two carbonyl groups such as uridine or adenosine which scavenge $O2 \cdot -$ [45]. However, it has been proposed that nucleosides, amino acids, peptides, and other small organic metabolites are also able to work individually without manganese, scavenging OH· efficiently but not $O_2 \cdot^-$ or H_2O_2 [83 45]. To gather up complexation partners for building manganese scavengers, *D. radiodurans* could accesses granules containing polyphosphates [76 45]. Furthermore, its inability to use nucleosides or TCA cycle products as carbon source may be beneficial for the accumulation of amino acids and nucleosides for ROS scavenging [45]. Additionally, irradiation leads to high levels of nucleolytic [84 85 12] and proteolytic activity [45], further filling up the pool of potential small-molecule partners for manganese complexes aiding in the protection against ROS.

Prior studies demonstrated that the protective abilities of manganese complexes can be expanded to other species. For instance, Escherichia coli cells, human Jurkat T cells and murine cells were administered protein-free *D. radiodurans* cell extracts enriched for radioprotective metabolites *ex vivo*. Intriguingly, the ultrafiltrate was able to shield the cells from otherwise lethal doses of ionizing radiation [45].

In a study conducted in 2016 the radioprotective capacity of manganese-peptides on eukaryotic cells was investigated by administering synthetically derived Mn2+-decapeptide complexes to female mice and exposing them to 9.5 Gy ionizing radiation [19]. Compared to untreated mice who displayed a lethality rate of 63%, all treated mice survived the exposure and displayed no acute radiation syndrome.

Another strategy of damage control against oxidative stress and radiation is to sanitize the cell quickly and efficiently from and oxidized nucleotides and dysfunctional proteins. To clean the cell from potentially toxic or mutagenic nucleotides and proteins, *D. radiodurans* can access an armada of proteases, hydrolases and nucleosidases and transporter proteins to first break down damaged macromolecules and then export the peptides.

As mentioned previously, initial DNA degradation in the process of DNA synthesis is an important part of cell regeneration. To break down damaged nucleotides, *D. radiodurans* employs nucleosidases (nucleoside monophosphate phosphohydrolases) [86] and Nudix hydrolases [87] [54]. After degradation of the damaged molecules, cells then export the molecules via UvrA homologs which are closely related to ABC transporter proteins to prevent reincorporation of damaged nucleotides, the [88]. Furthermore, cells are also cleaned from oxidatively damaged, dysfunctional proteins by proteolytic removal and subsequent re-synthesis [89]. For the degradation export of damaged proteins on the other hand, *D. radiodurans* possesses 3 hemolysin genes, 10 secreted subtilisin-like proteases and amino acid ABC transporters which are highly induced after radiation exposure [44] [43] [90]. Interestingly, like E.coli and yeast, *D. radiodurans* displays a preference for which proteins to degrade first [91]. Some proteins like the chaperons GroEL and Dnk, as well as TCA cycle enzymes aconitase and citrate synthase are rapidly degraded but are also the first to be resynthesized [92] [47]. Others like RNA polymerase subunits β (RpoB) and β′ (RpoC) seemed to be spared from degradation [92]. Overall, studies have shown that proteolytic activities increase in cells subjected to radiation and are triggered by aconitase which may act as a oxidative stress sensor [93]. In a nutshell, the enhanced proteolytic activity seems to serve multiple purposes in the irradiated cell: first, it seems to cleanse the cell of off malfunctioning proteins. Secondly, it seems to provide an exceptionally efficient way to generate building blocks for ROS scavenging manganese-peptide complexes and other proteins that are in high demand by cannibalizing on the proteins of neighboring dead cells or on damaged intracellular proteins [44] [43].

1.6 The Tanpopo mission- experimental set up and goals

To elucidate one of the most thrilling mysteries of humankind -the question about the origin of life- the Tanpopo mission was initiated. The main goal of the mission was to assess the possibility of the panspermia theory [94], which states that organic compounds or even life could exist elsewhere in the universe and could be distributed by meteorites, asteroids comets, dust particles and other means of travel. Ultimately, it also proposes the possibility that seeds of terrestrial life originated from an

extraterrestrial source. To test the possibility of interplanetary transfer of life, one of the first questions to pose is: can lifeforms retain their viability in the harsh environment of outer space, and if so, how? During the Tanpopo mission dehydrated *D. radiodurans, Deinococcus aerius* and *Deinococcus aetherius* cells were exposed to LEO conditions outside of the ISS for two and three years respectively [95] (Fig.1 B), to subsequently assess their viability. Therefore, the dried cells were kept in aluminum plates (25 x 3 mm) and subjected to LEO at the outside of Exposure Facility in the Japanese Experimental Module of the ISS. As control, equally treated samples were kept on earth for the same amount of time. Back on earth, succeeding investigations of the samples revealed that cells suffered a considerable amount of DSB. However, pellets of 500 µm thickness or greater were able to retain their viability (Kawaguchi *et al.*, 2020). Nevertheless, the molecular basis of this extreme resistance remains elusive. In this study I investigated the transcriptional response of *Deinococcus spp*. exposed to LEO conditions. The aim of this investigation was to identify unknown potential multi-resistance inducing genes and molecular pathways which are upregulated under outer space stress exposure. To explore the early gene expression upon ISS exposure, cells were harvested after 5 hours of revival in TGB medium for the subsequent RNA extraction. Furthermore, some cells remained in TGB for 15 hours in order to investigate the long-term effects of space exposure on gene expression. Samples were designated as *D. aerius* TR0125, however the transcriptome shared much more similarities with reference genome of *D. radiodurans* strain R1 (genome assembly ASM856v1) than to *D. aerius* (genome assembly ASM289737v1), hence this genome was used as reference for the transcriptomic analysis (Table.1). Initially, the extracted RNA of samples kept on the ISS was highly degraded, whereas RNA from ground control samples displayed moderate RNA degradation. However, after 15 hours of revival the extracted RNA displayed a better quality, demonstrating that cells were able to regenerate even after the high amount of stress they were subjected to. Furthermore, the transcriptomic analysis revealed differential expression of more than 1300 genes after 5 and 15 hours. Interestingly, classical DNA damage response genes encoding for pathways like RecFOR were not significantly affected. Instead, both early and late transcriptomic response resulted in the upregulations of transporter encoding genes. While early transcriptomic response experienced especially pronounced upregulation

of ion-channel encoding genes ISS15h cells primarily upregulated genes encoding amino-acid transporter genes. Furthermore, the elevated expression of isocitrate lyase ISS5h cells indicated that the glyoxylate bypass might have been more active. In contrast to cells which stayed on earth, some genes with implications in the process of ROS scavenging have been upregulated as well as a considerable number of glycosidases involved in the BER process. Puzzlingly, I found highly significant upregulation of genes related to transposition. This upregulation could have been induced passively due downregulation or destruction of elements which regulate transposition or because of the need for fitness selection [96]. In other words: transposition could have been a way of introducing possible beneficial mutations. Moreover, results suggest that many proteins carrying the enigmatic family of intrinsically disordered proteins (IDP) experienced transcriptional upregulation. This family of proteins has been proposed to encapsulate macromolecules in order to protect them from desiccation damage [11]. Lastly, according to the analysis, GGDEF domain containing messenger proteins seemed to be the main conductor of the transcriptomic concerto. Overall, the study managed to elucidate interesting features which could have helped the cells to survive the hostile environment of outer space. Additionally, bearing the limitations of transcriptional analysis in mind, it proposes interesting candidate genes for future investigations to find out what makes the members of *Deinococcus spp*. so exceptionally resilient.

2. RESULTS AND DISCUSSION

2.1 Growth kinetics and bioinformatic quality control of samples after LEO exposure

Organisms and Spacecraft alike are exposed to detrimental high levels of radiation such as non-ionizing UV radiation or ionizing radiation caused by sporadic solar particle events and constant galactic cosmic radiation when exposed to outer space conditions. To elucidate how polyextremophilic *Deinococcus spp.* can withstand adverse outer space-like conditions, dry samples were exposed to LEO conditions on the outside of the spacecraft of the International Space Station (ISS) for 3 years in the course of the Tanpopo mission. The transcriptome of these cells was then compared to dehydrated “ground control” (GC) cells kept on earth for the same time span. Therefore, the first batch of cells was harvested and subjected to RNA extraction after 5 hours of revival in TGB (ISS5h and GC5h) to investigate the early transcriptional response to the detrimental environmental condition posed by LEO. This point was chosen because the cells were presumably in the process of regeneration as they still exhibited outer space related damage as demonstrated by the fragmented RNA according to the Bioanalyzer (Fig.D, Table 1). However, a low RIN (RNA Integrity number) value can alter the results of sequencing and compromise downstream analysis [97] therefore, taking an earlier sample was deemed too risky. Hence, a 5-hour period was considered suitable to investigate of early stress related gene expression. Subsequently, another batch was taken after 15 hours of revival to elucidate the long-term effects on gene expression (ISS15h and GC15h). The second point of time was chosen based on the growth curve performed on cells that were exposed on the ISS for 2 years (Fig.2 A, B). The curve revealed that cells started to enter the exponential growth phase after 15 hours. Especially after 20 hours of growth, samples displayed a very high standard deviation curve. The standard variation was caused by single cell contaminations whose morphology was smaller than the tetrads of Deinococcus spp. cells. Sample 2 displayed the strongest contamination (Fig.2 B).

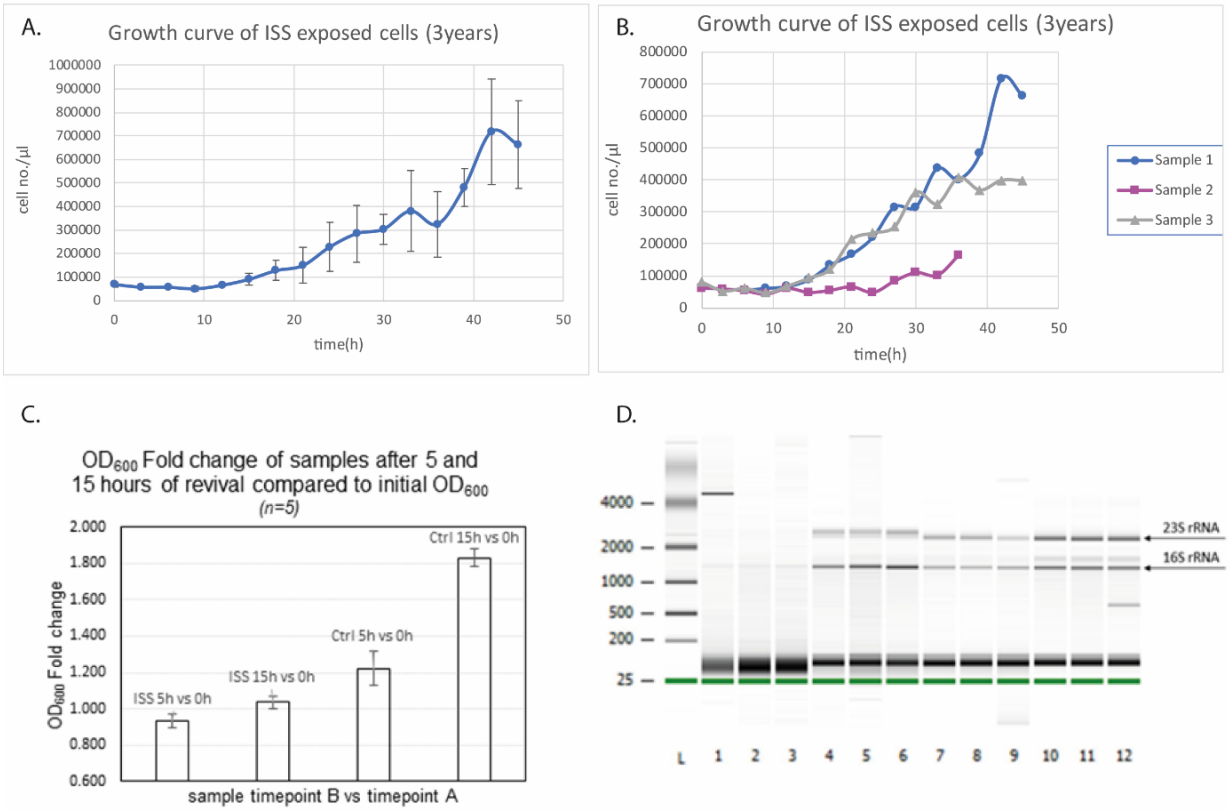


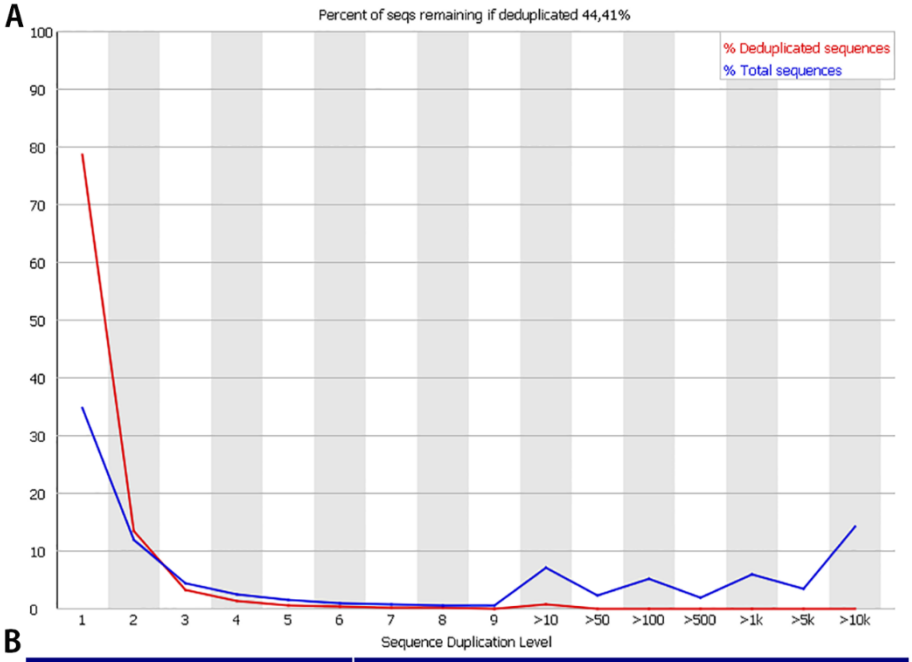
Figure 2. Growth curve and kinetics and RNA quality of ISS exposed samples and ground control samples harvested after 5 and 15 hours of revival.

Upon exposure of *Deinococcus spp*. samples to ISS LEO conditions and terrestrial conditions in a dried state for 3 years, growth habits of ISS exposed cells were assessed by counting cells every 3 hours for 45 hours. (A) shows the average cell number at certain points of time. Standard deviation is depicted as black. Graphs in (B) shows results from individual samples. Prior to extraction, OD_{600} measurements via spectrophotometry were performed to assess growth kinetics (C). Fold changes were calculated using the initial OD_{600} values after inoculation of TGB 0h (time point A) and values of the same culture after 5 and 15 hours of recovery (time point B), respectively. Values are depicted in logarithmic scaling. (D) For extraction, cells were harvested after 5 and 15 hours of recovery in TGB medium and RNA integrity and concentrations were measured using the Agilent RNA 6000 Nano Kit. Bands for 16S and 23S rRNA are indicated by arrows. RNA integrity was measured in triplicates. Lanes 1 to 3 depict RNA retrieved from ISS exposed samples after 5 hours and 2 to 6 after 15 hours. Lanes 7 to 9 show ground control samples after 5 hours and 10 to 12 after 15 hours of revival.

Overall, the growth curve showed that the stressed cells started their exponential growth phase after approximately 15 hours. By investigating gene expression of cells after 15 hours, we hoped to elucidate long term effects of exposure to outer space related stress on the transcriptome as well as the kinetics of the transcriptional response.

Unfortunately, a full growth curve could not be created for control cells and for cells exposed for 3 years due to a very limited number of samples. Hence, the growth curve of the 2-year exposed samples was used as an orientation.

A



B

Measure	Value
Filename	88573_ATTACTCGGGCTCTGA_CD4K4ANXX_1_20190621B_20190621.bam
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	10286424
Sequences flagged as poor quality	0
Sequence length	50
%GC	61

Figure 3. FastQC report of preliminary spike in experiment

To evaluate the quality of the RNA preparation and illumina sequencing, raw reads were controlled via the FastQC tool. Panel (A) depicts duplication rate of reads, table (B) lists the type of sequencing mode as well as GC content and sequence length.

To assess growth kinetics of the 3 years exposed and control samples, OD_{600} values of samples right after inoculation (0h) and after 5 and 15 hours (5h, 15h) of revival in TGB were measured (Fig 2, C). Cells exposed on the ISS for 3 years experienced an initial decrease of OD_{600} values displayed by an average fold change of 0.9 after 5 hours of growth. After 15 hours of growth, OD_{600} values exhibited miniscule changes, indicated by a 1.04-fold increase compared to initial OD_{600} values measured at 0h. On the other hand, cells kept under terrestrial conditions showed an average fold change of 1.2 upon 5 hours of revival and an almost 2-fold increase after 15 hours of growth in TGB. This indicates that control cells started exponential growth after 15 hours. Furthermore, RIN values and concentrations of extracted RNA were compared between different samples and points of time. While OD_{600} levels remained similar in ISS samples even after 15 hours of growth, RNA concentrations increased by a factor

of 3 (Table 1). The initial average RIN value of 2.6 in exposed cells, which indicates highly degraded RNA, increased to 6. Samples kept under terrestrial conditions experienced a 4-fold increase in extracted RNA when comparing extractions performed after 5 and 15 hours, respectively. Their RIN value increased from 5.1 to 7.3. Overall, the OD_{600} measurements indicate that cells kept on the ISS experienced a longer lag phase and enter the logarithmic phase later than cells kept under terrestrial conditions. The increase of RNA concentration and RIN value during the lag phase show that cells likely use this elongated lag phase for the recovery the cells. This is consistent with prior studies in which cells were shown to undergo a recovery phase during which growth stagnates after they were experiencing high oxidative stress [12]. This stagnation phase, during which degradation of DNA is actively induced by the cell, is necessary for genomic restauration before the cell can enter growth phase [12]. The elongated lag phase and the higher RNA degradation levels of cells kept on the ISS hint on the fact that these cells experienced higher damage than control cells which resided on earth. However, whether the high RNA degradation levels of the cells were a direct consequence of the high UV and radiation exposure and other outer space related stress factors or whether degradation was actively induced by the cells for genomic recovery was not evaluated during this study.

2.2 Taxonomic classification of samples

Next, we assessed differential gene expression by sequencing the extracted RNA. The samples were denoted as *Deinococcus aerius* TR0125, a member of the *Deinococcus* genus with a GC content of 68 mol% and a draft genome length of 4,524,446 bp [98]. To test the protocol for successful RNA extraction and rRNA depletion, a preliminary test extraction was performed on samples kept inside the ISS space craft for two years. Subsequent Illumina sequencing using a sequencing length of 50 bp resulted in 5191978 total reads with good sequencing quality and high but tolerable duplication levels (Fig. 3 A, B). Henceforth, the protocol was also applied on the „outside"

novo Transcriptome via the Trinity pipeline. The de novo transcriptome (blue box) was aligned to the reference genome (red box) of *D.aerius* (A) and *D.radiodurans*(B). Images show cut outs from the complete alignment as the complete alignment is too large. Colored blocks indicate the presence of the same sequence in both genomes. If the outline of a block is a straight like, no mismatches between sequence were found (b.) If the outline has a dent, mismatches have been found between sequences with each dent representing a mismatch (a.). Colored blocks facing downwards from midline indicate an inverted sequence in the de novo transcriptome compared to the reference genome. Numbers stand for base pairs. (C) Represents the alignment of reads (grey boxes) extracted from ISS15h samples (upper panel) to the reference 16S rRNA gene locus DR_r01 of *D.radiodurans* (lower panel). Colored base pairs in reads indicate mismatch, grey area indicates no mismatches.

samples which were later used for downstream analysis. The software tool Trinity [99] was used for a de-novo assembly of the transcriptome, in order to generate a draft genome scaffold using the reads from sequencing. This scaffold

can later be used for alignment and subsequent quantification of aligned reads in order to identify gene expression. Curiously, the scaffold genome generated by Trinity showed poor alignment to the published draft genome of *D. aerius* (NZ_BFAG01000001.1) but very good alignment to the published genome of *D. radiodurans* (NC_001263.1, NC_001264.1, NC_000959.1, NC_000958.1) (Fig.4 A, B) when using the "progressiveMauve" algorithm for alignment [100]. To investigate the taxonomy of the samples further, the genome-wide Average Nucleotide Identity (aANI) was evaluated. This analysis is considered as a robust method to assess genomic relatedness [101]. Generally, when comparing two genomes of interest, gANI values of 96.5 are categorized as minimum threshold to assign a genome pair to the same species. ANI has been proposed to strongly correlate with the 16S rRNA gene sequence identity [102]. Comparison of the assembled Trinity Fasta file with the reference genome of *D. radiodurans* revealed a gANI value of 99.81 whereas the value for the *D. aerius* genome was 79.74 (Table 2). To compare with a third *Deinococcus* member,

the draft genome of *Deinococcus apachensis* (ARLS01000001.1) [103], a close relative of *D. aerius,* was used. The pairing of *D.apachensis* and *D. aerius* revealed the closest relation with gANI value of 89.9, whereas the comparison with the Trinity

Parameters	***ISS 5h (n=3)***	***ISS 15h (n=3)***	***GC 5h (n=3)***	***GC 15h (n=3)***
RNA (ng/µl)	112.7	338	389	1494.7
RIN	2.6	6	5.1	7.3
total reads	12685747.67	12584140	16711091.67	12831363
Reads aligned to R1 (%)	52.7	67	80.5	80
reads aligned to TR0125 (%)	14	16.5	25.6	18.2

Table 1. RNA extraction and alignment statistics

RNA concentration was measured using nanodrop after Trizol extraction. RIN value to assess RNA quality was evaluated using bioanalyzer. Raw reads were aligned to reference genomes of R1 (*Deinococcus radiodurans*) and TR0125 (*Deinococcus aerius*) using the bowtie2 function in RSEM and statistics were retrieved using the "flagstat" command of the tool samtools. All values depicted are the average of 3 replicates.

assembly revealed a more distant relation (gANI of 80) and the comparison with *D. radiodurans* the furthest with a value of 62. To further validate the use of the genome of *D. radiodurans* for downstream gene expression analysis, quality trimmed reads of the final samples were aligned to the genome of *D. aerius* and *D. radiodurans* and alignment rates were quantified using the samtools “flagstat” command. Around 80% of the control reads and around 50% to 69%, depending on the time point, of the ISS exposed reads aligned to the genome of *D. radiodurans* (Table1). The lower alignment rate of ISS exposed samples could be due to contamination, or due to the experimental conditions causing higher mutation rates and RNA high degradation after high oxidative stress exposure. Moreover, considering the observation that the alignment rate increased after 15 hours, it can be hypothesized that the lower rate is partially due to damage related reasons which were alleviated over time. Another explanation for the increase is that *D. radiodurans* was gradually overgrowing potential contaminations. Using the Blast aligning software [104] to identify the origin of the first 54 unmapped reads from the sample ISS5h_R1 revealed that the majority of the reads could not be assigned to any species. Most of the unmapped reads with known origin were deriving from genomic, non-coding regions from *D. radiodurans*, which suggests that DNA contamination was present and/or hints at miRNA interactions. However, investigating possible RNAi pathways

Species	***Trinity gANI***	***D. aerius gANI***
D. aerius	79.74	
D. radiodurans	99.81	61.99
D. apachensis	80.39	89.9

Table 2. Comparisons of gANI values of *Deinococcus spp. and the de-novo Transcriptome*
The de-novo Transcriptome that was assembled by the Trinity software based on RNA sequencing of the spike-in experiment was compared to the reference genomes of *D.aerius, D.radiodurans* and *D.apachensis.* Reference genome of *D.aerius* was further compared to *D.radiodurans* and *D.apachensi.*

was beyond the scope of this project. Furthermore, small numbers of the unmapped reads derived from contaminating species e.g., *Bacillus subtilis* and *Cutibacterium acnes.* Alignment rates of adaptor trimmed reads to the genome of *D. aerius* were overall lower and ranged from 20 to 25% in ground control sample and 14 to 16.5% in ISS exposed samples (Table 2).

The commonly used gold standard to infer taxonomic relations is the evaluation of sequence similarities of the 16S rRNA gene. For this method, the current threshold to claim that two samples might belong to the same species is a sequence similarity of 98.65% [105]. To inspect the samples in terms of their relatedness, the software integrative genomic viewer (IGV) [106] was used to locate and extract the consensus sequence of the 1501 bp long 16S rRNA gene (DR_r01) of samples exposed to ISS and terrestrial condition. Finally, the BLAST software package [107] was used to search for sequence similarities. The consensus sequence of 16S rRNA of ISS exposed samples after 5 hours of revival showed an overall sequence similarity of 98.46% with *D. radiodurans* R1 and after 15h a similarity of 97.86%. The control samples displayed a sequence similarity of 99.47% after 5 hours and 99.73% after 15 hours of revival. The comparison to *D. aerius* however, did not appear in the top 20 results ordered according to sequence similarities indicating that there were few similarities between the samples and the genome of *D. aerius*. Nevertheless, the consensus sequence similarity of the ISS exposed samples was too low to claim that they might be *D. radiodurans*. A closer look into the origin of some of the reads that aligned to the reference 16S rRNA gene revealed that reads which harbored mismatches most likely derived from a contamination according to BLAST results. For example, reads containing mismatches displayed in Fig.4 C. were 100% identical to sequences found in *B. subtilis* or *Enterococcus sp*. However, they aligned because the alignment algorithm allowed a certain number of mismatches. Almost all reads with mismatches that aligned to the 16S rRNA genome displayed a 100% sequence similarity with common contaminations such as *Bacillus sp, Escherichia coli, Enterococcus faecalis* and in one instance to an uncultured bacterium according to BLAST, indicating that these sequences most likely derived from a contamination. When determining the consensus sequence, IGV considers all the reads which aligned to the reference (i.e., also reads with mismatches). If two or more different nucleotides align to the same locus, the algorithm will choose the nucleotide whose frequency is greater than 50%

and greater than twice the number of the second most frequent nucleotide [108]. Hence, if the frequency of a “foreign” read was higher than that of a read with 100% similarity with *D. radiodurans,* the consensus sequence would contain artificially introduced “point mutations”. This led to the generation of a chimeric 16S rRNA contig and an artificial downscaling of the sequence similarity during the BLAST search. Overall, for almost every read containing a mismatch that aligned to the reference there was a read that shared a 100% similarity with the reference. After manually removing reads that had a 100% sequence similarity to contaminating strains and <99% similarity to representatives of the *Deinococcus* genus, the newly compiled 16S rRNA contigs for the samples displayed high similarities to *D. radiodurans* R1 ranging from 99.75 to 99.87% for the R1 reference genome with the accession number CP015081.1 and NR_074411.1, respectively.

Overall, this observation exemplifies the challenges of identifying microbes, especially in complex communities or when working with non-model organisms and what effect even small numbers of “point mutations” can have on the outcome of sequencing data analysis. A critical element of taxonomic evaluation though sequencing seems to be the software which should be conservative and stringent regarding mismatches to avoid the generation of chimeric 16S rRNA and corrupt taxonomic assessment. Therefore, sensitive methods and rigorous assessment of the data is needed to verify claims. Since the corrected 16S rRNA had a high similarity with the corresponding gene in the reference genome *of D. radiodurans,* the samples are referred to as *D. radiodurans* during downstream analysis for simplification. Furthermore, as it displayed the highest similarity, the reference genome of *D. radiodurans* was used to identify differentially expressed genes. However, the possibility that the samples actually stemmed froma close relative cannot be ruled out.

2.3 RNA expression value normalization and clustering of samples to infer similarities and differences.

To elucidate differential gene expression, transcripts were fist quantified by using the software package RSEM [99] in order to generate raw counts by aligning the reads to

the reference genome of *D. radiodurans*. The RSEM equivalent to raw counts is denoted as expected counts. A critical factor of finding biologically meaningful gene expression in RNA-seq data is the type of normalization raw counts are subjected to. By normalization we can adjust for biases in sequencing depth, library size and compositional biases. Since we were primarily interested in making comparisons between samples and not between genes within a specific sample, we decided to perform the normalization of expected counts using TMM normalization (Trimmed Mean of M-values) as described in Robinson and Oshlack [109] using the Bioconductor tool EdgeR. The underlying hypothesis of edgeR is that most genes are not differentially expressed. This method of normalization is specifically recommended for inter-sample comparisons and has been shown to give good results for both simulated and real data [110] [111]. In edgeR, normalization is achieved by computing a normalization factor (norm.factor). The standard method of computing the norm.factor is by calculating the weighted trimmed mean of the log ratios between two samples (the M value). The normalization factor is then multiplied with the library size in order to unify library sizes and minimize composition biases. Composition bias in RNA-seq is generated if for instance few genes are very highly expressed due to treatment and are therefore monopolizing the sequencing procedure, leaving less reads to be be distributed among the remaining genes. Hereby, the monopolization by these "greedy" genes artificially lowers the read counts of the remaining genes, since there is a limited number of possible reads that have to be distributed amongst all the genes in the sample. The monopolization by these genes would thereby lead to detection of false positive differential gene expression when performing inter-sample fold-change comparisons. For this study, the unnormalized library sizes were not similar between the samples (Fig.5 A). ISS5h samples displayed the smallest library sizes whereas the ones of GC15h were almost twice as big. This was probably because ISS5h had a low total raw read count and the lowest alignment rate of all the samples (Table1). Calculating the norm.factors for all samples using the standard TMM method (Fig.5 B)

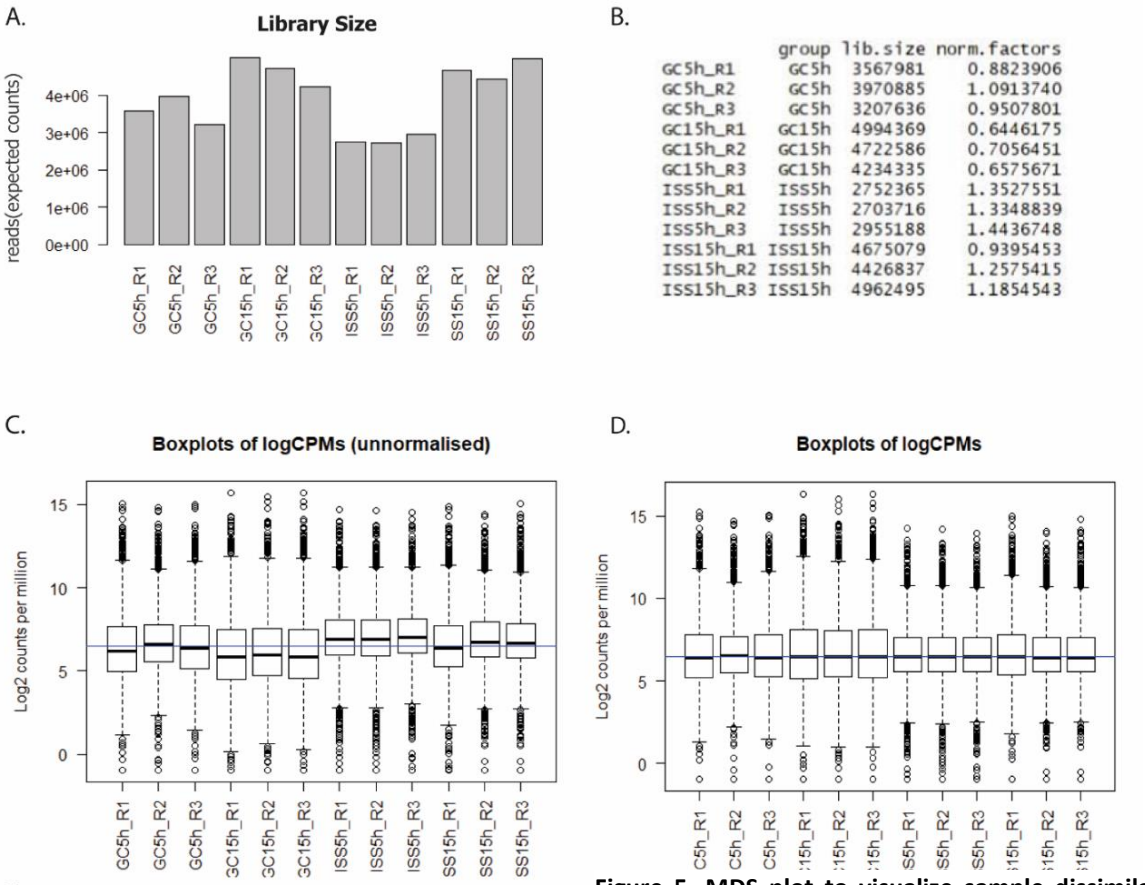


Figure 5. MDS plot to visualize sample dissimilarities after TMM normalization of reads via Edge R

Differential expression (DE) was evaluated via the Bioconductor tool EdgeR using the computed "expected counts" from RSEM. DE was calculated between 3 replicates of ISS exposed samples revived for 5 hours (ISS5h_R1 to_R3) and ISS exposed samples revived for 15 hours (ISS15h_R1 to _R3) and their respective ground controls (GC). EdgeR calculates the $\log_2$ value of counts per million (logCPM) by using TMM (Trimmed Mean of M) normalization of expression values. A normalization factor (norm.factor) is calculated and the initial library size (A) is adjusted for the subsequent DE evaluation. Norm.factors used for normalization are shown in (B). Thereby, initial logCPMs of genes (C) that were skewed due to differences in library sizes or due to composition bias, i.e. not because of biologically relevant causes bias unified (D), preventing the detection of false positive or false negative DE. For showing sample variation a Multi-dimensional scaling (MDS) plot was made measuring distances between samples and plotting the first 2 leading dimensions(dim) (E). Distance equals dissimilarity between samples (colored dots). In the ISS15h sample cluster one data point is covered by another because of high similarity.

revealed that especially GC15h samples probably contained several very highly upregulated genes which led to a lower median cpm compared to the most average sample. This resulted in a nom.factor below 1 (Fig.5 B), which led to downscaling of the library size by multiplication and an analogous upscaling of the gene counts, which were quantified as “counts per million” (Fig.5 C, D). ISS5h samples on the other hand had a slightly higher median cpm (Fig.5 C) than the other samples. In this case, the norm.factor was above 1, leading to the downscaling of the genes to unify the gene count across samples (Fig5. D).

To make an initial general analysis of sample relations and explore which factors had the biggest contribution on sample variances, a multi-dimensional scaling plot (MDS plot) was created (Fig.5, E). Plotting the log fold-change of the two dimensions that had the most impact on sample variance against each other revealed a clear separation between ground control samples on one hand and ISS exposed samples on the other. Overall, the plot showed a clustering of samples depending on the revival time and the exposure type, i.e., kept on earth or on the ISS. This indicates that both factors influenced the overall variance between the samples and equally contributed to the cluster separation in the MDS plot. Additionally, apparently RIN value did not have an impact on the observed variability, as the samples with the most similar RIN values, ISS 15h and GC 5h, clustered far apart from each other. This suggests that the observed variance is indeed caused by biologically relevant differences and not due to the level of RNA degradation. Furthermore, no batch effect was observed. Similarly, a heatmap showing the gene expression of all 3077 detected genes displayed hierarchical clustering of samples into a respective group depending on type of exposure and time of revival (Figure 6). As expected, the triplicates were most alike while the next hierarchical classification was depending on the type of treatment. Furthermore, the heatmap revealed that he initial naïve assumption that ISS samples revived for 15 hours will more closely resemble the transcriptome of early ground control cells did not hold true. The hypothesis was that they might be in a similar stage of recovery since the ground control cells experienced lower levels of stress. However, the gene expression of ISS15h samples were more similar to ISS

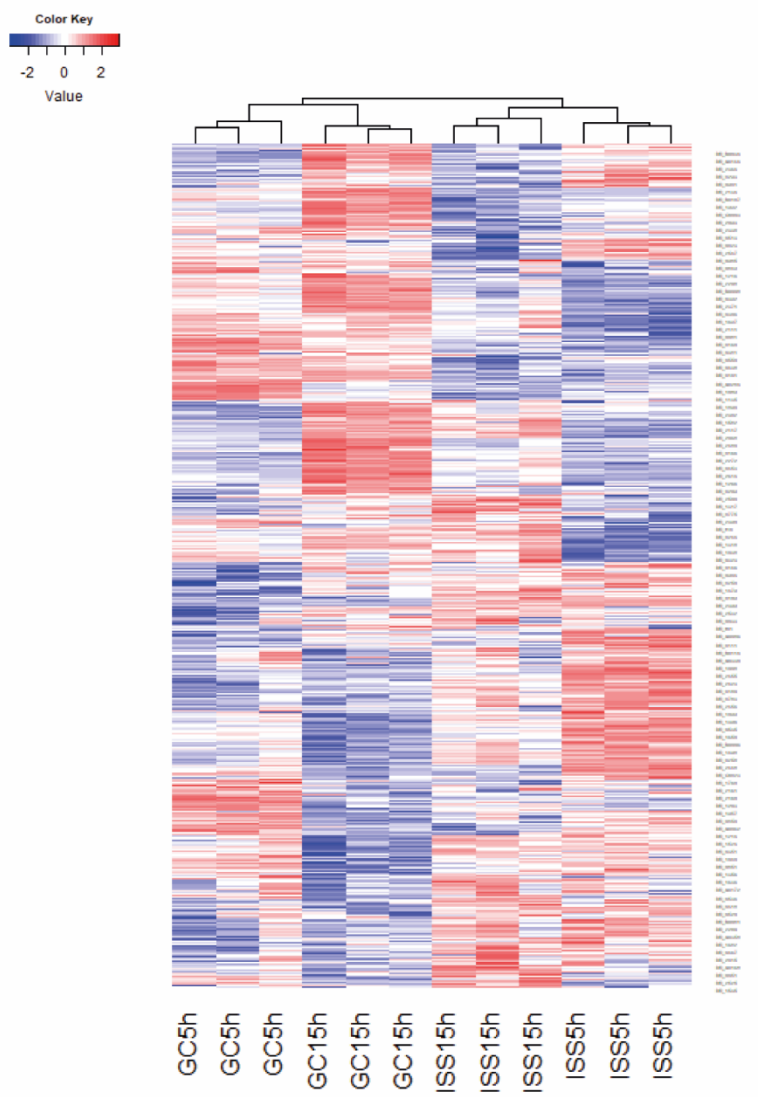


Figure 6. Investigating differential gene expression of all detected genes across samples

Heatmap was created by using counts per million (CPM) computed via EdgeR using all genes where expression could be detected, with a cutoff of 10 reads per gene. Color code denotes differential expression with red signalizing overexpression and blue lower expression (see color key). ISS5h and 15h stands for samples exposed to ISS outside condition and cell revival in TGB-Medium for 5 and 15 hours respectively and GC for ground control. Hierarchical clustering of gene expression was performed to investigate sample relations. Gene expression was measured in triplicates for all conditions.

exposed samples revived for 5 hours than to any of the ground control samples. Overall, many genes that were expressed in ISS5h cells were also expressed in ISS15h. ISS15h however, seemed to display a more toned-down expression of said genes. Additionally, ISS5h shared a small subset of relatively upregulated genes with GC15h cells. Nevertheless, most genes upregulated in GC15 were downregulated in ISS5h samples and vice versa.

2.4 Transcriptomic analysis of cells exposed to LEO conditions after 5 hours of revival.

Next, I investigated the early transcriptomic response in more detail by comparing DE gene expression of ISS5h samples with the control GC5h. To infer biologically meaningful statements from out investigations, only genes with a $\log_2$ fold-change

significantly greater than 1.2 were considered in the analysis. A quantification of differentially expressed genes revealed the presence of 701 downregulated and 660 upregulated genes in ISS samples revived for 5 hours compared to the control (Fig. 7 A). Additionally, a smear plot revealed that a small subset of genes was being overexpressed more than 30-fold in ISS5h samples (Fig. 7 A). A Heatmap showing the expression the top 200 DE genes in this comparison revealed a population of genes that were exclusively upregulated in ISS5h genes (Fig7. B, C). Some of these genes were still active after 15 hours of revival but the majority experienced downregulation. Regarding the function of the genes, at the time of the analysis only 9 of the 62 genes exclusively upregulated in ISS5h had a function annotated to them. Some were denoted as uncharacterized according to the Uniprot database but encoded known domains [112] which were inferred by homology. Others did not share at least 50% of sequence identity with any other species according to UniProt. Overall, this observation indicated that many potentially crucial key players for cell repair and stress coping pathways in *D. radiodurans* were poorly characterized, if at all.

2.4.1 Potential ROS scavenger pool enrichment by transcriptional upregulation of transporter proteins, Exopolyphosphatases and *SodC*

The first comparison to identify genes which were relevant for the regeneration upon outer space related stress was made between ISS5h and GC5h cells. Amongst the highly upregulated genes in ISS5h cells was the exopolyphosphatase (PPX) DR_A0185 (Fig.7 C). This enzyme catalyzes the hydrolysis of inorganic polyphosphates [113] to release orthophosphates. Compared to GC5h, PPX experienced a highly significant log2 fold change of more than 8 (Fig.8). Additionally, the outer space exposed cells experienced significant upregulation of the phosphate transporter encoding transcripts DR_0925 and DR_A0160 (pstB) with varying degrees of elevation (Fig.8). Due to their ubiquitous use in many different cellular functions poly and orthophosphates play an important role in the life cycle of bacteria and are involved in cell signaling, growth, stress response or as building blocks for essential molecules such as the phosphate backbone of DNA or ATP [114]. In *D. radiodurans* however, they may be also used as a source for the production of Mn-

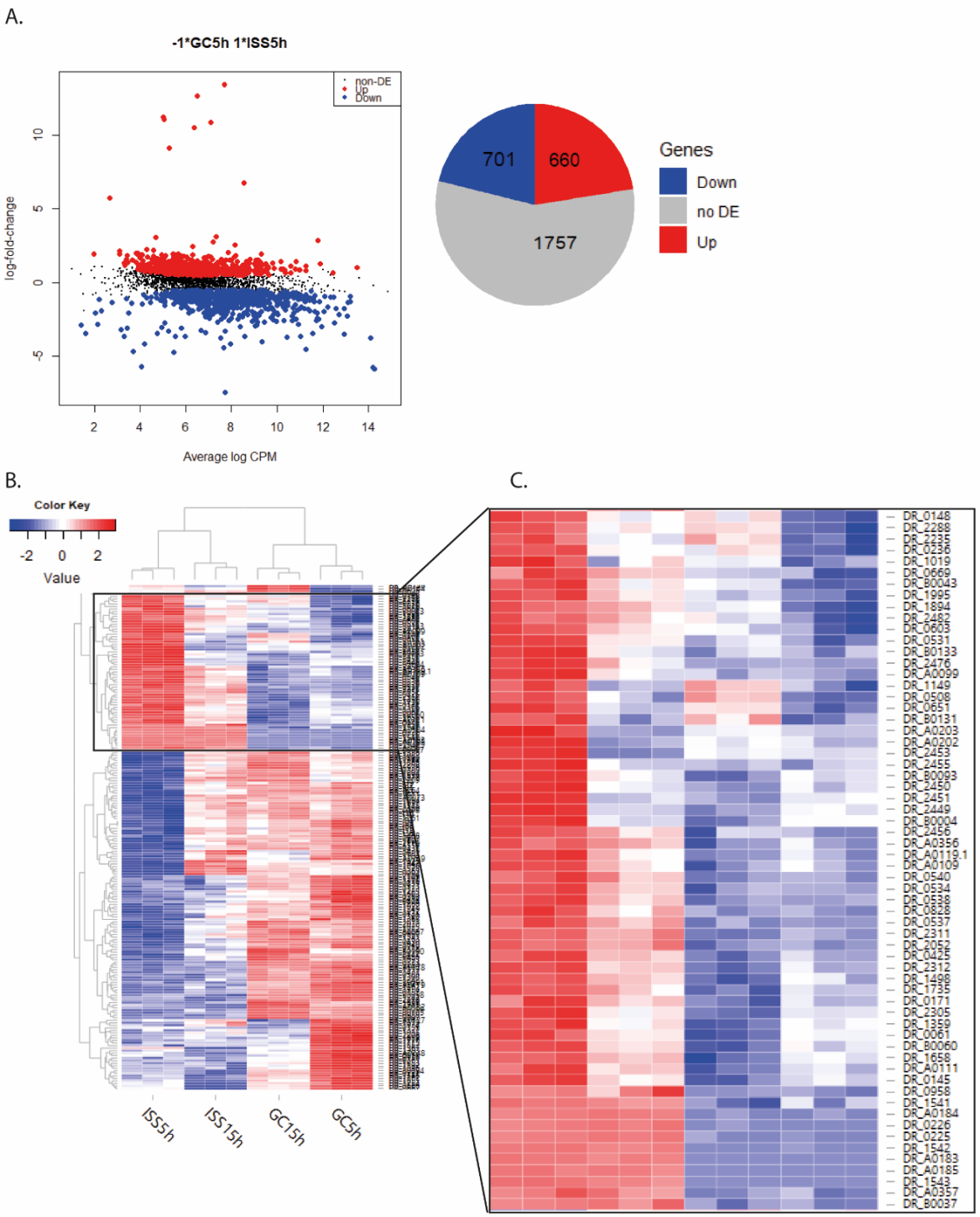


Figure 7. Top 2oo most significant DE genes of ISS5h samples compared to GC5h control to elucidate short term gene expression response upon LEO exposure.

EdgeR was used to identify differential gene expression (DE). Cells were revived in TGB for 5 hours (ISS5h) and subsequently compared to the control that was kept on earth and was also revived for 5hours (GC5h). RNA was extracted using Trizol and sequenced using Illumina sequencing. For the EdgeR analysis only genes with a $\log_2$ fold-change of more than 1.2 were considered as being overexpressed. In (A) the MDS plot shows the log-fold change and average abundance of all significant DE genes. Significantly up and down regulated genes ($p<0.05$, FDR cut off at 5%) are highlighted in red and blue respectively. Heatmap (B) was created based on the top 200 most significant DE genes in the comparison ISS5h vs. GC5h, showing the z-scoring of the log of cpm (counts per million). Level of gene expression is indicated by color coding (color key). A close-up of significantly upregulated genes found in ISS515h is depicted in (C). Triplicates of each condition are shown.

phosphate complexes which have been proposed to function as radical scavengers [45]. Multiple studies have revealed the formation of electron dense granules in the cells of the extremophile, which probably harbor manganese and phosphorous, and calcium moieties which might serve a source to synthesize protective ROS scavengers [115]. Furthermore, a model proposed for *E.coli* suggests that orthophosphates can alleviate cell stress by co-exporting Cu^{2+} ions which can generate ROS in the course of the Fenton reaction [116]. Moreover, a study on the resistance of *D. radiodurans* mutants revealed moderately decreased survivability of PPX-like protein mutants (DR_0826) against γ-rays and high sensitivity to UV-radiation and H_20_2 [117]. In our data, DR_0826 expression also experienced a 1.3-fold increase, albeit with a lower significance (p=0.08). This indicates that cells exposed to LEO conditions seem to have an increased demand for orthophosphate molecules in order to cope with the experienced stress. They could use the imported and processed phosphates either to form ROS scavengers using orthophosphates in complex with Mg (Fig.9) or they might employ phosphates to directly export ions that could lead to ROS build up.

Early transcriptional response to LEO led to the upregulation of multiple genes encoding transport-related proteins such as the Na+/H+ antiporter (DR_1149) and Na(+)-linked D-alanine-glycine permease (DR_B0133), a previously unrecognized peptide ABC-transporter protein DR_0958 and lastly also DR_2453 which encodes a Cation-transporting ATPase (Fig.7 C, Fig.8). According to the Pfam database the transcript of DR_2453 contains a predicted sequence for heavy metal-associated domain for copper ion-binding. Copper could be involved in enzymatic activities as a co-factor e.g., for the stress induced superoxide dismutase sodC. Furthermore, copper has been shown to be involved for the polymerization of the most abundant protein in the S-layer of *D. radiodurans* [118]. Therefore, the import of copper might help to repair cell in two ways: by alleviating cell wall damage by inducing polymerization and by functioning as co-factor during (Fig.9). Additionally, many more transporter genes which are not reported in the heatmap showing only the top 200 genes experienced significant elevation of expression such as DR_0205, DR_2118, DR_2283 and DR_1709 (Fig.8). The latter two transporter proteins are presumed to regulate Mn homeostasis [119]. Many of these transporters have already been observed to experience upregulation in response to acute radiation in *D. radiodurans* [43], [62]. It

was speculated that transporter upregulation could help with the export of damaged and potentially toxic molecules on

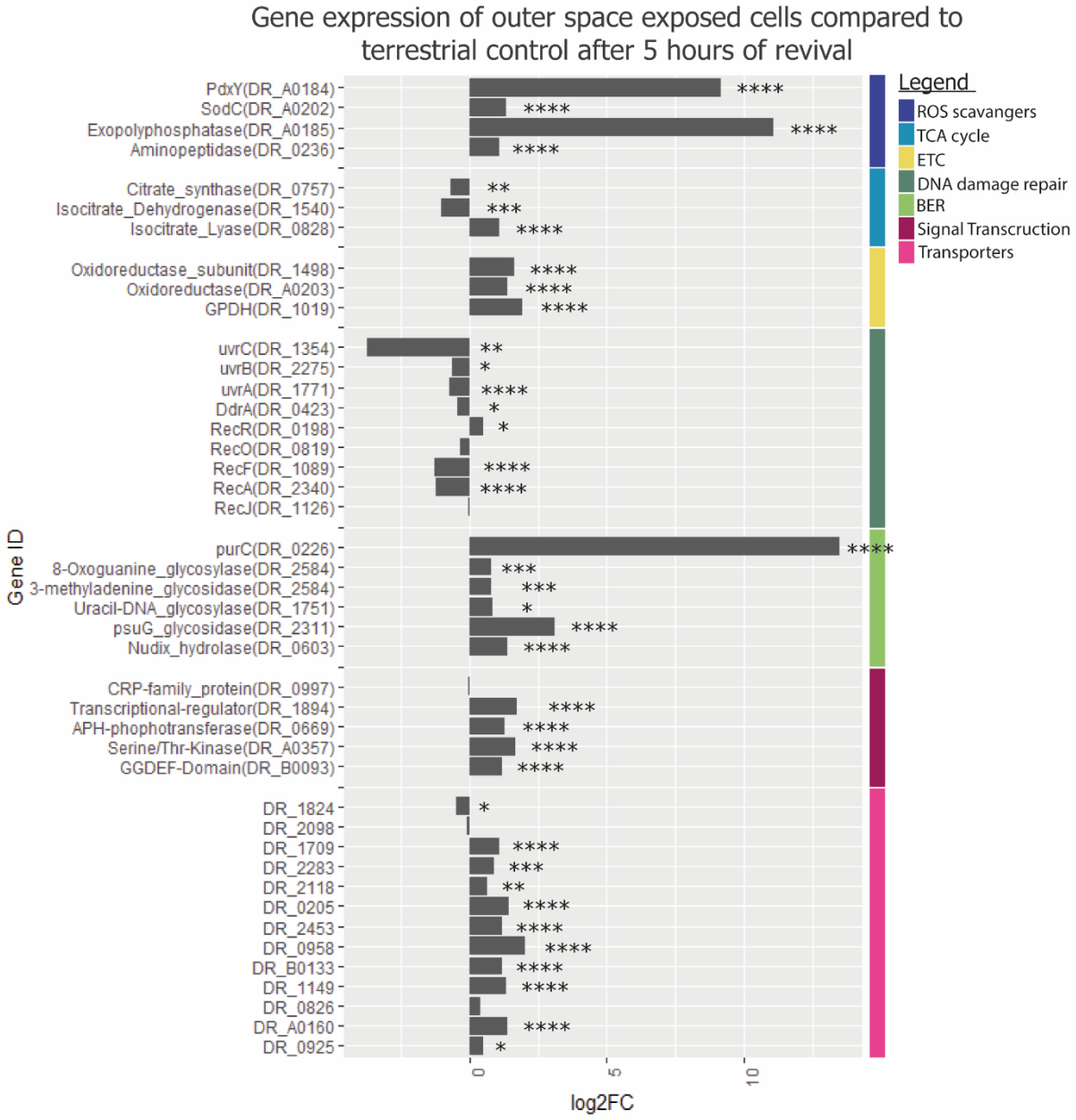


Figure 8. Upregulation of genes encoding transporters and repair enzymes in cells exposed to the ISS for 3 years during early response phase.

Cells were revived for 5hours in TGB (ISS5h) compared to the control that was kept on earth and was also revived for 5h (GC5h) to capture early responses. Barplot shows the log_2 fold change in expression in ISS5h cells compared to the control. Gene ID indicates the Locus Nname of the transcript as well as the Gene name if available or names of functional domains found in the transcript. Color coding indicates the known or predicted pathways of genes (legend in top right corner). Level of significance is indicated by asterisks: *= p<0.05, **= p>0.01, ***= p>0.001, ****= p<0.0001. P-values were FDR corrected using Benjamini-Hochberg correction. Abbreviations used in legend: BER= Base excision repair, ETC= electron transport chain, TCA=Tricarboxylic acid, ROS= radical oxygen species. Gene expression was measured in triplicates.

one hand, and with the import of proteins and ions for energy production, biosynthesis, and creation of ROS scavenging molecules on the other [43]. However, in this dataset transcripts previously described as encoding export proteins (DR_2098, DR_1825) were not upregulated in ISS5h cells relative to the control (Fig.8). Henceforth, high upregulation of various transporter proteins might signalize that cells which experienced LEO had an elevated demand for the import of proteins, cations, sugars and other molecules in order to promote biosynthesis, energy production and the generation of ROS scavenging molecules. For example, the Na+/H+ antiporter could provide Na+ to facilitate the uptake of D-alanine-glycine by the Na (+)-using permease DR_B0133. Both genes were represented in the top 200 list of DE genes (Fig.7 B). The amino acids D-alanine and glycine could be further used for the reconstruction peptidoglycan-containing cell wall layer of *D. radiodurans* which mainly consists of glutamic acid, alanine, glycine, and L-ornithine [40] and probably was one of the primary sites of damage (Fig.9). Furthermore, the upregulation of ABC-transporter protein DR_0958 and the other transporter genes were concurrent with DR_0236 encoding the putative aminopeptidase, a gene whose upregulation upon stress has not been recorded before (Fig.7 B, C). This aminopeptidase was significantly overexpressed by a factor of 2 in ISS5h cells compared to the control. According to Uniprot, the aminopeptidase harbors a metalloprotease domain belonging to the M29 family. In Listeria monocytogenes, M29 metalloproteases have been previously found to prefer the amino acid arginine for hydrolyzation [120]. Another significantly upregulated gene was DR_0651 encoding a probable arginase (Fig.7, A). Using Mn (II) in its active site, arginases are enzymes that catalyze the final step in the urea cycle by converting L-arginine to urea and L-ornithine [121]. Ornithine then can serve as a precursor for proline or polyamines like putrescine or spermidine which can stabilize DNA [122] and aid in the battle against ROS [123]. Henceforth, arginase could use L-arginine provided by the aminopeptidase in order to generate L-ornithine. Ornithine might then be used to produce polyamines which in turn could alleviate the damage that was inflicted on the cells by either being converted into a ROS scavenging spermidine or putrescine or by being incorporated into the cell wall to restore damage (Fig.9). As these enzymatic activities are highly dependent on the presence of Co-factors, the expression of genes encoding transportation proteins seemed to be elevated in cells that were exposed to LEO

conditions. Also involved in the synthesis of Co-factors is the highly DR_A0184 encoding pyridoxal kinase PdxY (Fig.7 C, Fig.8). In this study, the transcript experienced a highly significant log_2 fold change of 9 compared to control cells (Fig.8). This kinase is involved in the salvaging of pyridoxal 5'-phosphatase, better known as the catalytically active form of vitamin B6. Vitamin B6 is used as a co-factor by many enzymes [124] and is possibly also a quencher of singlet oxygen ROS [125]. Another well-known ROS scavenger that experienced transcriptional upregulation in LEO exposed cells relative to the control is the enzyme superoxide dismutase *SodC* (DR_A0202). This enzyme catalyzes the dismutation of the radical O_2^- into more harmless O_2 and H_2O_2. As it is dependent on copper and zinc it might again exemplify a reason of why the increased ion transportation is highly needed after the high stress the cells were exposed to. Overall, after 5 hours of regeneration cell which experienced outer space related stress were highly occupied with importing molecules for consumption. The defining criteria that set their gene expression apart from the control cells concerning cellular trafficking was the transportation of ions rather than the export of damaged cellular components (Fig.9). However, it has to be considered that ISS5h cells were compared to cells which too were exposed to desiccation for 3 years and therefore also experienced cell damage. Hence, it is likely that ISS5h cells did elevate expression of exporter proteins relative to unharmed cells, however not relative to the terrestrial control cells used in this study. Ion transporting proteins on the other hand did experience a pronounced gene overexpression in this comparison.

2.4.2 Identification of poorly characterized transcriptional regulators which elicit a putative live-saving response after 5 hours of revival.

Multiple genes encoding proteins involved in transcriptional regulation exhibited upregulation in the first 5 hours of revival relative to the control. Among them is a protein containing the GGDEF domain (DR_B0093), a domain named after the conserved sequence motif "Gly-Gly-Asp/Glu-Glu-Phe" (Fig.7 C, Fig.8). The transcript DR_B0093 also harbors transmembrane motifs indicating that the domain is in close proximity to the cell membrane. The GGDEF domain can act as a diguanylate cyclase and is associated with the synthesis of bacterial second messengers such as c-diGMP (cyclic diguanosine monophosphate), cGAMP (guanosine monophosphate–

adenosine

monophosphate) and cyclic di-AMP (adenosine monophosphate) after exposure to an external stimulus [126]. C-diGMP synthesis by GGDEF domain containing proteins can lead to subsequent signal transduction to elicit a transcriptional response. Overall, the GGDEF domain was shown to modulate many critical aspects of bacterial behavior. For example, signaling via GGDEF proteins can lead to changes in bacterial growth, biofilm formation and overall behavior by influencing the cell on a transcriptional, translational level and posttranslational level [127]. Previous studies on the behavior of organisms during space flights revealed that many organisms such as *Salmonella typhimurium* [128] *or Candida albicans* [129] exhibited changes in behavior by enhancing biofilm formation and cell aggregation. As for the nonpathogenic bacterium *D. radiodurans*, no peer reviewed study has found biofilm formation in this organism so far. However, expression of GGDEF domain could direct the transcriptional response after stress-inducing stimuli in a way, so that cell motility, cell-cell trafficking, nutrient uptake and other live-saving cell functions are enhanced. Furthermore, signal transduction of extracellular stimuli could be forwarded by the response regulator DR_A0357, the Serine/threonine kinase protein (DR_0534) and APH domain-containing phosphotransferase DR_0669. These genes were also highly upregulated (Fig.7 C, Fig.8) and represented in the top 200 DE (differentially expressed) gene list and could have been involved in the process of propagating stress signaling until it eventually leads to a transcriptional response and modulation of gene expression. Gene expression might have been regulated by the transcriptional regulators such as the transcriptional regulator from the Lrp/AsnC family (DR_1894) which was highly upregulated in ISS5h compared to the control (Fig.7 C). Hence, this chain of events triggered by GGDEF domains could have been in charge of eliciting a transcriptional response to ensure the survival of the cell.

Overall, the transcriptional response seemed to serve a multifactorial purpose. Not only did the cells experience a high upregulation of proteins in charge of transportation and were able to produce high amounts of ROS scavengers such as *SodC,* polyamines and Mg-complexes (Fig.7 C, Fig.9) as state above, they also seemed to lead the metabolism of the cells into a favorable direction. For example, a key feature of *D.*

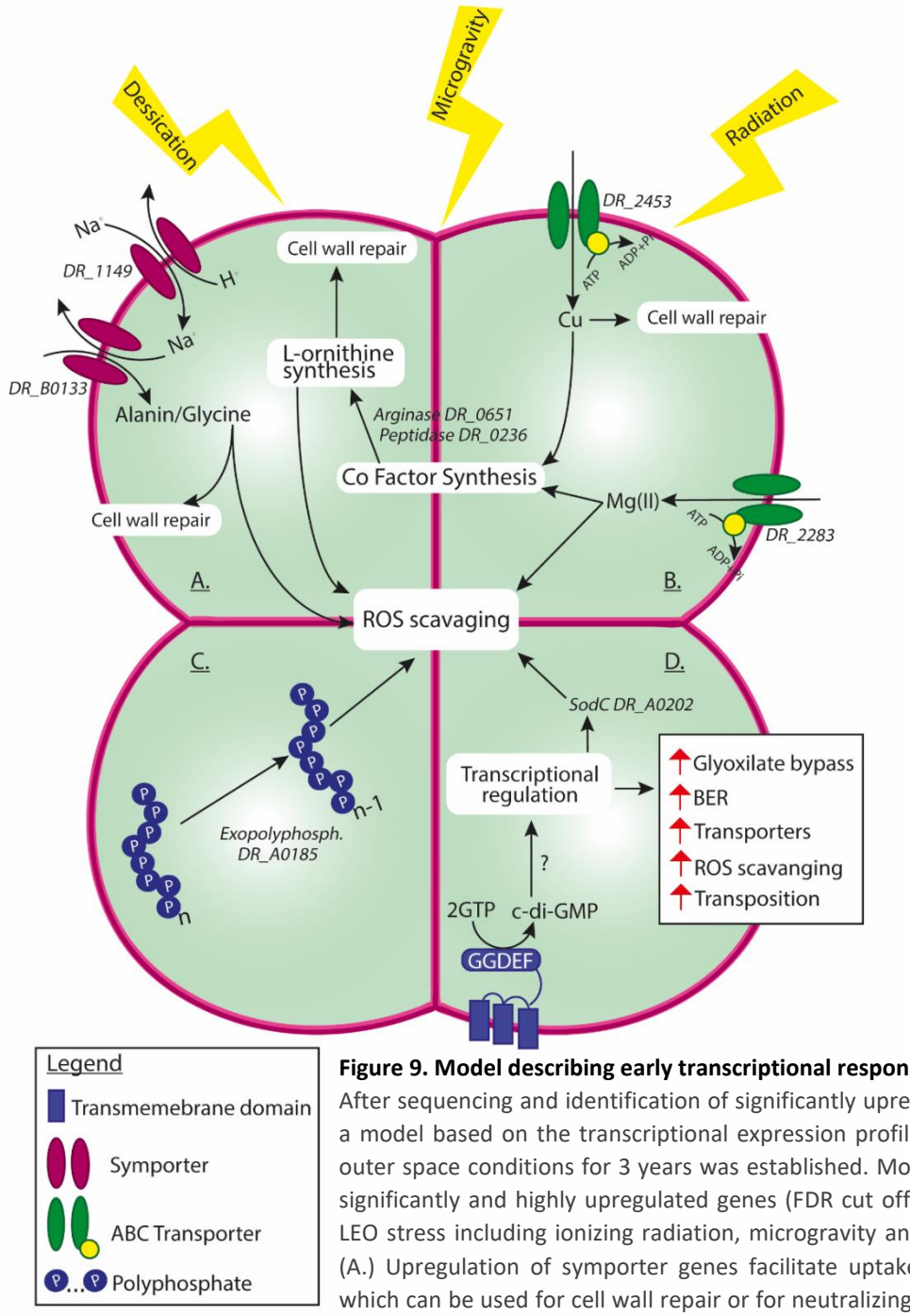


Figure 9. Model describing early transcriptional response upon LEO stress.
After sequencing and identification of significantly upregulated genes via EdgeR, a model based on the transcriptional expression profile of cells exposed to ISS outer space conditions for 3 years was established. Model is based on the most significantly and highly upregulated genes (FDR cut off at 5%, p>0.0005). Upon LEO stress including ionizing radiation, microgravity and prolonged desiccation. (A.) Upregulation of symporter genes facilitate uptake of alanine and glycine which can be used for cell wall repair or for neutralizing reactive oxygen species. Upregulation of ATP-transporter encoding genes facilitate the uptake of ions like Mg and Cu which can be used as co-factors (B.) during enzymatic activities of transcriptionally upregulated Arginases and Peptidases to generate L-ornithin. L-Ornithine can be used to mend cell wall damage or for the synthesis of ROS scavenging polyamides. Significant upregulation of exopolyphosphatases catalyzes the hydrolysis of inorganic polyphosphate which can be used to further feed the pool of ROS scavengers (C.). GGDEF domain containing proteins bound to membrane can transmit extracellular signals by producing the secondary messenger cyclic di-GMP which leads to transcriptional modulation (D.). Transcriptional regulations lead to elevated expression of ROS scavengers like superoxide dismutase like SodC, to induction of glyoxylate bypass and BER as well as transcriptional upregulation of transporter proteins and higher occurrence of transposition events.

radiodurans is its ability to employ the glyoxylate bypass. This is an alternative pathway to the TCA cycle in which CO_2 generating reactions of the TCA cycle are skipped so that the cells can suppress endogenous ROS production [42]. Indeed, we found significantly elevated amounts of the transcript DR_0828 which encodes isocitrate lyase (Fig.7 C, Fig.8), a key enzyme of the glyoxylate bypass. Conversely, the enzymes which catalyze the steps of the TCA cycle that are skipped during the glyoxylate bypass, namely isocitrate dehydrogenase (DR_1540) and citrate synthase (GltA) (DR_0757) displayed lower expression compared to the control (Fig.8). This indicates that cells which experienced LEO conditions indeed resorted to a downregulation of the TCA cycle whereas the glyoxylate pathway was upregulated to uphold energy generation and gluconeogenesis while simultaneously avoiding ROS generation (Fig.9).

Another catabolic enzyme which experienced very high transcriptional upregulation was Glycerol-3-phosphate dehydrogenase (GPDH) encoded by DR_1019 (Fig.7 C, Fig.8) which is involved in glycolysis, and phospholipid biosynthesis but also in the electron transport chain (ETC) for the generation of ATP by catalyzing the transfer of electrons to quinones [130]. Two other transcripts that encode putative proteins that have implications with ETC were found to be significantly upregulated, namely DR_A0203 encoding a putative oxidoreductase and DR_1498 encoding NuoH, the NADH-quinone oxidoreductase subunit (Fig.7 C, Fig.8). This observation indicated the transcriptional upregulation of proteins involved in ETC to cope with the increased demand of ATP.

Lastly, there was another set of genes which encoded probable signal transducers. Some of the highly upregulated genes such as DR_A0109 and DR_B0004 (Fig.7 C), harbored disordered polyampholyte domains, characteristic sequences often found in so-called intrinsically disordered proteins (IDP) [131]. This interesting yet cryptic class of proteins lacks hydrophobic, structure stabilizing amino acids and therefore do not retain a fixed 3D structure while retaining their function. This feature allows a high degree of flexibility and hence has been predicted to be involved in transporters, receptors, and catalytic activities in humans [132]. Overall, this type of protein region is more abundant in eukaryotes than in prokaryotes and are over-represented in regulatory and linker regions [133] which implies that they might be necessary to

modulate the more complex cellular activities. Another hypothesis regarding the function of these proteins could be that they belong to the class of late-embryogenesis-abundant (LEA) proteins. These proteins display a high hydrophilicity index and are often found to accumulate in cells upon desiccation or hyperosmotic stress [134]. The putative protective proteins have been found in broad variety of organisms including cotton seeds, the nematode *Caenorhabditis elegans*, rotifers and *D. radiodurans* but their exact function remains elusive. It has been proposed that they might act as chaperones, hydration buffers, membrane stabilizing proteins, ion sequesters or transcription factors [135] [136]. However, their true biological function remains largely unknown and is mostly a matter of speculation. Nonetheless, their properties and their significant and high upregulation in ISS5h cells suggest that this cryptic class of proteins might have a critical role in establishing the high radio-tolerance that is demonstrated by *D. radiodurans*. Hence, more advanced characterization of the putative proteins encoded by DR_A0109, DR_B0004 and other genes containing IDPs might be valuable to deepen the understanding of this organism.

2.4.3 Upregulation of transcripts involved in BER but not of RecA-dependent pathways during early response period.

One of the major sites of damage upon radiation is DNA. Henceforth, cells usually respond by transcriptionally upregulating genes which are involved in the DNA damage response and the reconstruction of the shattered genome. One hallmark of the DNA repair system in *D. radiodurans* is the initial degradation of DNA and RNA before the cell resumes to restore the Genome (Dea Slade, 2011). This initial degradation period probably serves the purpose of cleansing the cell of potentially harmful mutagenized nucleotides and to create 3´-end overhangs which can be used to stitch the genome back together. One of the very first actions that organisms take after being exposed to a source of DNA damaging stress such as ionizing radiation is the removal of damaged nucleotides whereas rejoining double strand breaks takes more time [137]. In this data set, especially the putative “housecleaning” Nudix protein DR_0603 was significantly upregulated (Fig.7 C) by a factor of 2 (Fig.8). In general, Nudix proteins can detoxify and recycle modified nucleosides and are involved in

BER [138]. Interestingly, according to Uniprot this protein contains a methyltransferase domain in addition to the Nudix domain, suggesting that this protein might serve additional so-far unknown purposes besides hydrolyzing nucleotides for excisions. Another major contributor of BER are glycosylases and glycosidases which are in charge of removing alkylated, oxidized or deaminated bases upon radiation. In this study, we found high expression of pseudouridine-5'-phosphate glycosidase (psuG) encoded by DR_2311 (Fig.8) which catalyzes the degradation of the most common RNA pseudouridine modification in bacteria [139]. In fact, many of the glycosylases which have been observed to be upregulated in response to radiation (Dea Slade, 2011) were upregulated relative to the control after 5 hours of revival, however with varying degrees of upregulations and significances (Fig.8). This finding indicates that LEO exposure led to a high rate of base pair damage which resulted in the transcriptional upregulation of BER pathways during the early response period (Fig.9).

On the other side, biosynthesis for the reconstruction of the genome seemed to be initiated after five hours as well. I found extensive expression of DR_0226, a transcript that encodes purC (Fig.7 C, Fig.8). PurC is a phosphoribosylaminoimidazol e-succinocarboxamide synthase needed for purine biosynthesis. This could indicate that the cells are inducing the synthesis of new nucleotides. However, surprisingly the classical SOS DNA damage response was not upregulated in comparison to the 5-hour control. There was no relative upregulation of the recA recombinase nor of RecJ 5′-3′ exonuclease and the expression patters from the key players of the recFOR pathway were ambiguous (Fig.8). Moreover, the highly radiation- and desiccation-induced protein DdrA which is presumed to protect 3′ ssDNA overhangs from degradation to ensure RecA binding [69] was also not upregulated in ISS5h cells (Fig.8) compared to the control GC5h. Nevertheless, it is unlikely that the lack of expected upregulation of RecA was due to lack of DSB damage, considering the extensive upregulation of BER associated transcripts. The vast upregulation of glycosidases dealing with the neutralization of various mutagenized nucleotides and the shattered transcriptome displayed by the low RIN value indicated that extensive damage was inflicted on the cells during LEO exposure. Hence, it is more likely that control cells which were subjected to desiccation too also suffered from DNA breaks and therefore also had to upregulate the RecFOR and RecA-dependent pathways to mend DNA

breaks. It is possible that they already displayed the maximum of what is possible regarding DNA damage upregulation. Hence, upregulation in ISS5h cells was not visible in this comparison but would probably obvious when compared to cells which were not subjected to dehydration before. Alternatively, it is possible that the RecA-dependent pathways are upregulated at a later timepoint after BER is completed.

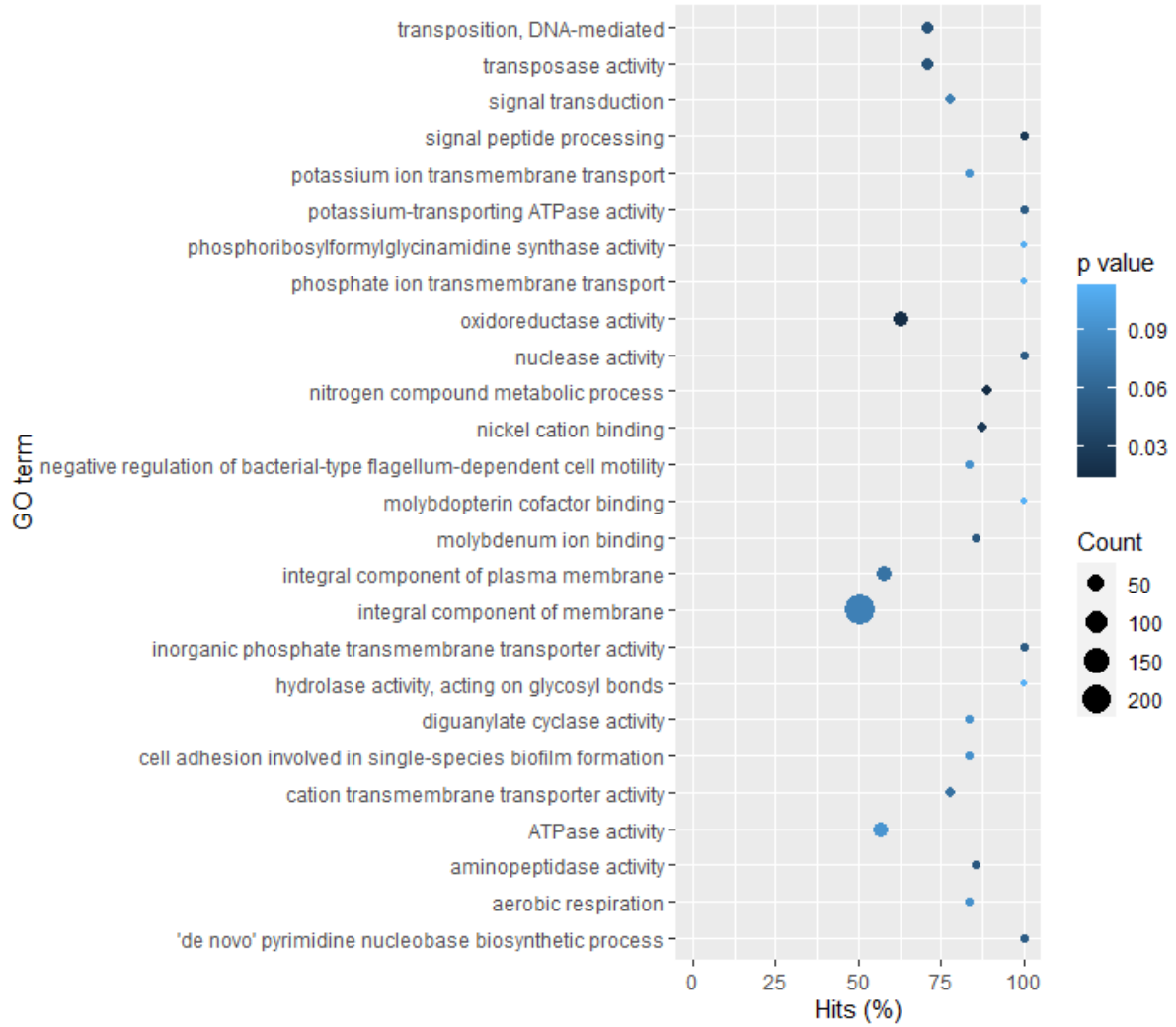


Figure 10. Gene ontology enrichment analysis in ISS exposed cells revived for 5 hours in TGB Compared to terrestrial control.

Go-term (gene ontology) enrichment analysis of the RNA seq data was performed using the Goseq Bioconductor tool. The analysis was conducted on RNA retrieved from samples after 5 hours of revival in TGB to investigate the early response and from cells revived for 15 hours to investigate the long-term effects on gene expression. The top 20 most significantly enriched Go-terms in the ISS5h samples are shown. Size of data points indicate the total number of genes that were assigned with respective Go-term. Go-terms represent upregulated terms belonging to the categories “biological process”, “molecular functions” and “cellular component”. Hits (%) indicate how many genes that are assigned with a specific Go-term experienced upregulation. Color of data points indicate *p*-value for over representation of the term in the differentially expressed genes. P-values of Go-term enrichment were calculated using the Wallenius method. P-values were corrected using the Benjamini-Hochberg procedure to avoid false positives.

Indeed, studies on the kinetics of DNA damage repair in 2.5-kGy-irradiated archaeon Halobacterium salinarum *revealed BER of oxidized bases in the first 2 hours whereas DSB were repaired within 8 hours* [137]*. However, even in the comparison ISS15h vs GC15h no upregulation of the recA-dependent repair pathways was found. This indicates that their upregulation could not be detected due to the nature of this experiment or because more extensive upregulation would take place after revival for more than 15 hours.*

2.4.4 Gene ontology enrichment analysis revealed elevated transposition activity after LEO exposure during early revival phase.

To reduce complexity of RNAseq data and reveal upregulation of the most prevalent biological pathways in cells, analyzing the over and underrepresentation of Go-terms is a popular approach [140]. Unfortunately, at the timepoint of the study more than 1000 Genes had no Go-terms assigned thus far. However, overall, the Go-term analysis of ISS5h versus GC5h cells reflected prior observations. The Go-term diguanylate cyclase activity was found to be enriched as well as the term for signal transduction as more than 75% of all genes assigned with this tag were over expressed (Fig.10). Puzzlingly, the analysis revealed enrichment of the tag “cell adhesion involved in biofilm formation” and “negative regulation of flagellum-dependent cell motility”. Again, more than 75% of genes with these tags displayed upregulation. However, so far there has been no peer-reviewed study published which showed biofilm formation or the expression of flagella in *D. radiodurans*. However, it could be possible that the bacterial cells investigated were representatives of a very close relative of *D. radiodurans* which developed means of motility and/or biofilm formation. Another relative for example, *Deinococcus geothermalis*, is able to produce biofilms when exposed to stressful conditions as demonstrated in a study where cells were exposed to simulated space and Mars conditions [141].

Closer inspections of the genes attributed with these tags revealed that all of them were proteins with GGDEF domains, a domain known to regulate biofilm formation and cell motility. The overrepresentation of this domain in the Go-tags was in

accordance with the high upregulation of the GGDEF domain containing protein DR_B0093 (Fig.8). The enrichment demonstrates that this form of signal transduction was apparently elevated in ISS5h cells compared to the control. This further suggests an important role of GGDEF-mediated responses in order to cope with extreme stress and modulate cell behavior in a favorable way.

Overall, the exact function of this interesting group of genes during stress response remains elusive. The transcriptional analysis suggests that they are important early response genes upon LEO exposure. However, further studies are needed to investigate their potential implication for either signal transduction, cell motility or another type of unknown function.

Moreover, multiple Go-terms indicated an enriched presence of membrane associated proteins. This was demonstrated by the of the Go-tag “integral component of the membrane” which was enriched in 50% of the 200 genes that were assigned to this tag. The finding suggests that the membrane probably was one of the most damaged units of the cells and that membrane reconstruction was therefore prioritized during the repair process leading to the upregulation of relevant genes. Additionally, upregulation of proteins associated with transportation of molecules such as inorganic-phosphates, potassium or, more general, cations was found (Fig.10). This observation mirrors the extensive upregulation of transporter-related genes that was described previously. Furthermore, the Goseq analysis also revealed enrichment of upregulated genes associated with the synthesis of pyrimidine nucleobases. All known genes involved in this process and henceforth assigned with the term “de-novo pyrimidine nucleobase biosynthetic process” were upregulated, indicating that cells are in the process of rebuilding the genome as well.

Interestingly, the analysis uncovered the upregulation of genes involved in transposition (Fig.10). Around 60% of genes which fall under the category of transposition were upregulated in ISS15h. This occurrence implied an elevated occurrence of transposition events in cells subjected to LEO conditions. In short, transposition describes the excision of a DNA sequence and a subsequence insertion of the sequence into a different locus. Higher occurrences of transposition events upon high radiation in *D. radiodurans* have been previously reported [142], [143]. Studies on transposition events in bacteria propose various consequences of elevated

transposition upon stress. On one hand, insertion of a non-native sequence into a gene body can result in downregulation or disruption of a gene. On the other hand, modulation of neighboring genes can occur by providing a new promotor sequence or by disrupting negative regulators [144]. Furthermore, the activation of transposable elements after stress have also been observed in eukaryotes, however their function remains elusive [96]. It is hypothesized that induction of transposition activity could facilitate the adaption to adverse environments by creating gene diversity and subsequently improving cell fitness. Alternatively, it is possible that the activation of transposons is the consequence of upregulation of nearby genes or the result of the disruption of regulatory pathways upon stress exposure. Ultimately, this finding poses the question whether cells actively induce transposition upon stress to improve cell fitness or whether transposition is more or less a passive consequence caused by other events like the disruption of regulatory elements.

Finally, the Go-term enrichment analysis also revealed a relative elevation of aerobic respiration in cells that experienced outer-space related stress during the first 5 hours. This was implied by the enrichment of terms like “aerobic respiration” in combination with “oxidoreductase activity” (Fig.10) where around 60% and 75% of the genes associated with this Go-term respectively experienced upregulation. The upregulation of genes involved in aerobic respiration could indicate that cells were trying to cope with an increased demand for energy generation.

2.4.5 Organization of upregulated genes in operons

Some of these significantly upregulated genes in ISS5h samples were found to colocalize at the same locus, such as DR_1541 to DR_1543 as well as DR_2449 to DR_2455 (Fig.7 C, Fig.11 A, C). More colocalizing ORFs that were upregulated together were DR_0225, DR_0226 and DR_A0356, DR_A0357 as well as DR_A0202 and DR_A0203 and DR_A0356 and DR_A0357 (Fig.7, Fig. 11 B, D, E). This suggests that these genes are organized as operons, a cluster of ORFs that are regulated by a shared promotor and transcribed into a single mRNA molecule [145]. Usually, operons code for genes from the same pathway, however there are also operons containing genes which do not share a common pathway but get expressed after the same stimulus [146]. In the first two clusters, there were two already characterized transcripts.

DR_1542 encodes a putative beta subunit of propionyl-CoA carboxylase and DR_2453 encodes a putative cation-transporting ATPase (Fig.11 A, C). Furthermore, the latter locus encoded a short transcript containing a predicted metal-binding transcriptional regulator that is possibly in charge of regulating genes depending on the presence or absence of metal (Fig.11 C). The locus DR_0226 harbours the protein purC which is involved in the de novo IMP (inosine 5'-monophosphate) biosynthesis for purine production which was upregulated in combination with DR_0225 that harbors helix-domains (Fig.11 D). DR_A0357 is predicted to code for a response regulator protein whereas DR_A0357 harbors a so-called domain of

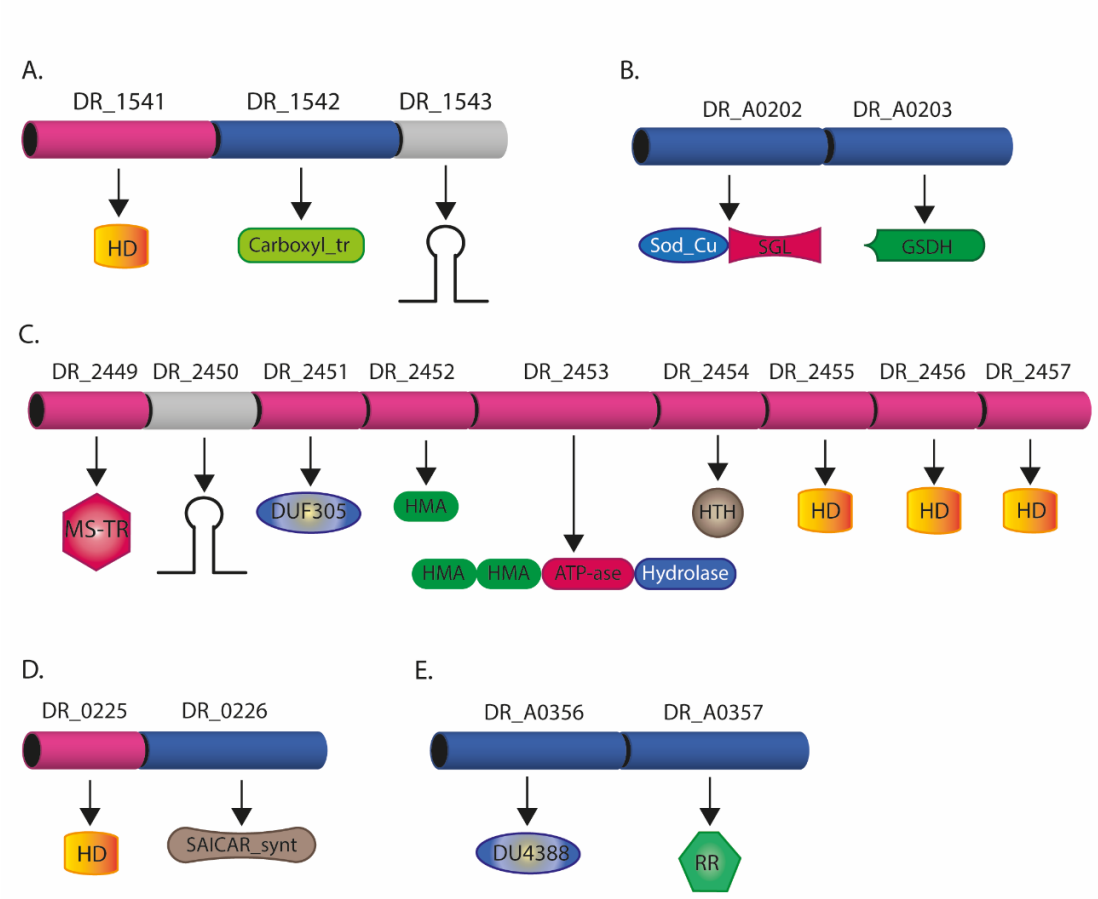


Figure 11. Upregulated genes located in operons in ISS exposed cells after 5 hours of revival.
After analyzing differential gene expression in cells that were exposed to outer space conditions on the ISS for 3 years, some significantly upregulated genes that were listed in the top 200 results were found to be located in clusters of highly upregulated genes (A-E). Colors represent the status of characterization of the transcripts. Blue indicates characterized genes where protein identity is known and is usually inferred from homology. Pink indicates uncharacterized genes where known domains (Symbols) are present in the sequence. Gray indicates uncharacterized transcripts without domains assigned to them. Smaller sized tubes represent transcripts with less than 200 base pairs in size. Hairpin symbol denotes ability of transcript to form stable hairpin loops. Data was taken from Pfam database. Abbreviations used: HD= helix domain, Carboxyl_tr= carboxyltransferase, HMA=heavy metal associated domain, DUF= domain of unknown function, Sod_Cu= superoxide dismutase, HTH=helix turn helix, RR=response regulator, SAICAR_synt= phosphoribosylaminoimidazolesuccinocarboxamide synthase,

unknown function (DUF). The loci DR_A0202 and DR_A0203 are interesting as they are translated into the superoxide dismutase from the Cu-Zn family (SodC) and into a putative oxidoreductase respectively. Both genes are only highly upregulated in ISS5h compared to the other samples (Fig.7 C, Fig.11 B).

Overall, most upregulated genes in these operons are not yet characterized. Some of them such as DR_2455, DR_2456, DR_B0131, DR_A0183 or DR_1995 harbor one or multiple trans-helical domains (Fig.7 C, Fig.11). Hence, it is possible that they might encode membrane associated proteins, however these trans-helical domains could be also found by chance by the Pfam algorithm as they are in general fairly short. Additionally, the last two contain permease domain for the transport of xanthine and uracil and a mechanosensitive ion channel domain respectively. Lastly, there are some upregulated ORFs without similarities to other proteins which are so short (<200bp) they might not be translated at all. Instead, investigating the folding properties of the short transcripts like DR_1543, DR_0225, DR_2449, DR_2456 revealed that many are capable of forming stable hairpin loops so they might act as short no-coding RNAs regulating translation in *D. radiodurans* after LEO stress exposure [147].

2.5 Transcriptomic analysis of cells exposed to LEO conditions after 15 hours of revival.

Similar to cells that were revived for 5 hours, cells after 15 hours displayed a small number of genes that experienced a significant $\log_2$ fold-change of over 10 in the 635 genes which were found to be overexpressed relative to GC15h (Fig. 12 A). Creating a heatmap based in the top 200 most differentially expressed genes in ISS15h compared to GC15h revealed that a small subset of 11 genes were only expressed in ISS15h cells, whereas the majority of highly upregulated genes after 15h were also expressed during the early transcriptional response phase after 5 hours.

2.5.1 Transcriptional overexpression of peptide transporters, pilin-associated proteins and peptidases after 15 hours of regeneration

Similar to the early response, a lot of upregulated genes found in ISS15h encoded transporter proteins such as DR_0280, DR_2469, DR_2122, DR_1324, DR_0958 and DR_2052 (Fig.12 C). However, in contrast to upregulated transporters ISS5h, the majority of these putative ABC transporters (namely DR_0958, DR_0280, DR_2469, DR_2122, DR_0958) were in charge of transporting peptides and possibly other macro molecules rather than ions or cations. Almost all of them experienced more than 4-fold increase of gene expression compared to the 15-hour control (Fig.13 A). On the other hand, ion transporters like the cation transporter DR_2453 or the Na+/H+ antiporter DR_1149 that were upregulated during the earlier response period (Fig.8) displayed. slightly lower expression levels after 15 hours (Fig.13 A). This observation suggested that during this later stage of regeneration cells which were subjected to LEO were more engaged in the uptake of proteins compared GC15h and early ISS5h cells. Moreover, there was high upregulation of a transcript denoted as DR_1964 encoding a protein annotated as general secretion pathway protein E according to the Uniprot database. This protein contained domains prevalent in Type II and type IV secretion system protein (Fig.12 C). In bacteria, the Type II secretion system (T2SS) is needed for protein secretion and is made up of a large multiprotein machinery consisting out of multiple subunits which are encoded in a single operon [148]. The assembled complex contains pseudo-pili which are structurally similar to the pili found in the secretion system type IV (T4SS), a secretion system that is also able to export DNA in addition to proteins [149]. Because of its ability to transport DNA, T4SS is used for bacterial conjugation, the exchange of DNA with the extracellular space [149]. The pseudo-pili T2SS pathway on the other hand is prevalent in plant and animal pathogens and is dedicated to the secretion of exoproteins such as virulence factors, cytochromes and macromolecule hydrolyzing enzymes [150]. Its main function, however, seems to be the acquisition of nutrients. Curiously, even though there is no evidence for pili production in *D. radiodurans*, the extremophile does harbor genes associated with pili-formation [42]. Hence, it is probable that genes associated with pili-generation are used in the secretion machinery rather than for conjugation or

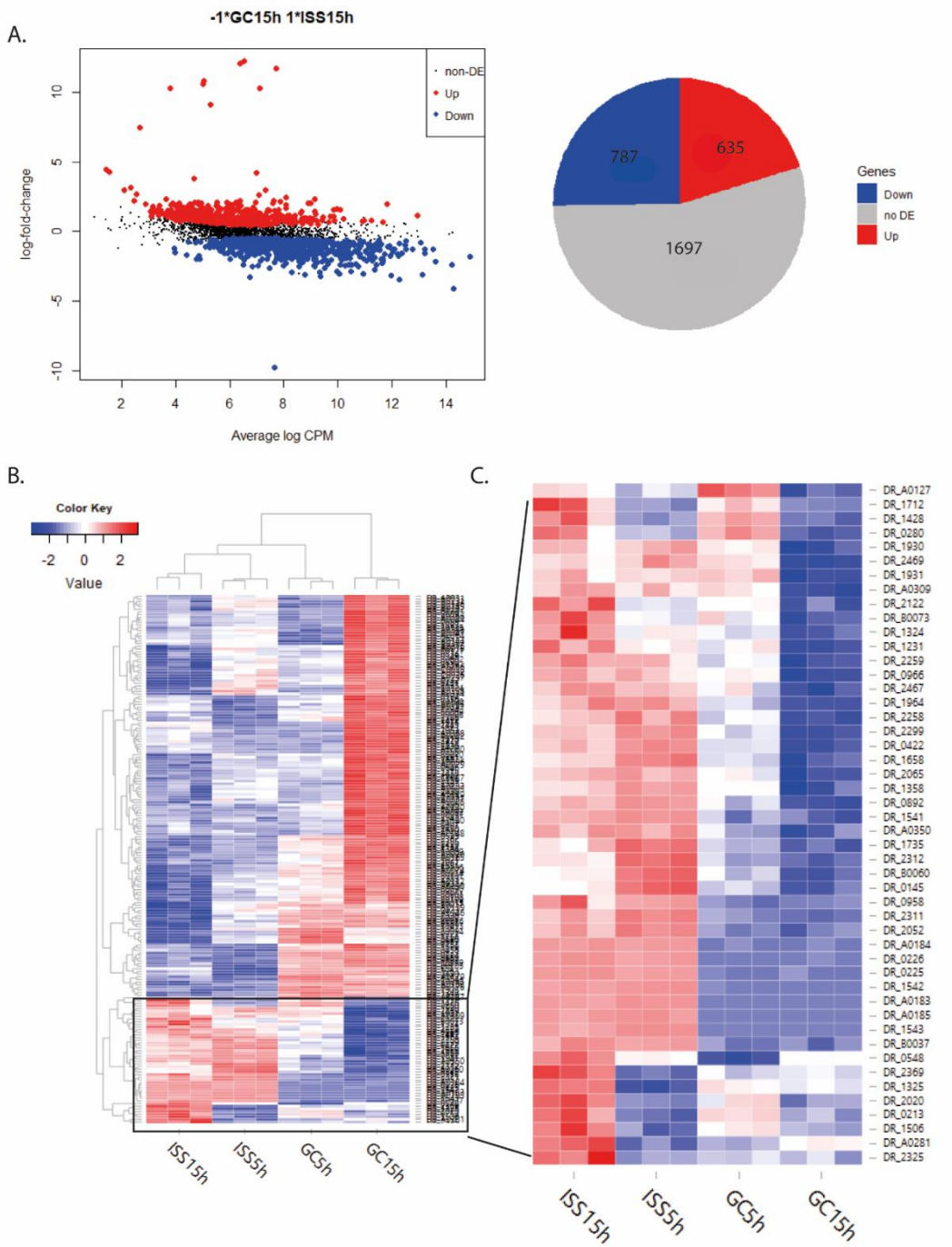


Figure 12. Investigation of the 2oo most significant DE genes of ISS15h samples compared to GC15h to elucidate long term gene expression response upon LEO exposure.

To investigate the transcriptomic response of *D.radiodurans* at later stages of recovery upon a 3 year exposure to Low Earth orbit(LEO) conditions outside of the ISS in a dehydrated state, EdgeR was used to identify differential gene expression (DE). Cells were revived for 15 hours in TGB (ISS15h) and subsequently compared to the control that was kept on earth and was also revived for 15hours (GC15h). RNA was extracted using Trizol and sequenced using Illumina sequencing. For the EdgeR analysis only genes with a $\log_2$ fold-change of more than 1.2 were considered as being overexpressed. In (A) the MDS plot shows the log-fold change and average abundance of all significant DE genes. Significantly up and down regulated genes ($p<0.05$, FDR cut off at 5%) are highlighted in red and blue respectively. Heatmap (B) was created based on the top 200 most significant DE genes in the comparison ISS15h vs. GC15h, showing the z-scoring of the log of cpm (counts per million). Level of gene expression is indicated by color coding (color key). A close-up of significantly upregulated genes found in ISS515h is depicted in (C). Triplicates of each condition are shown.

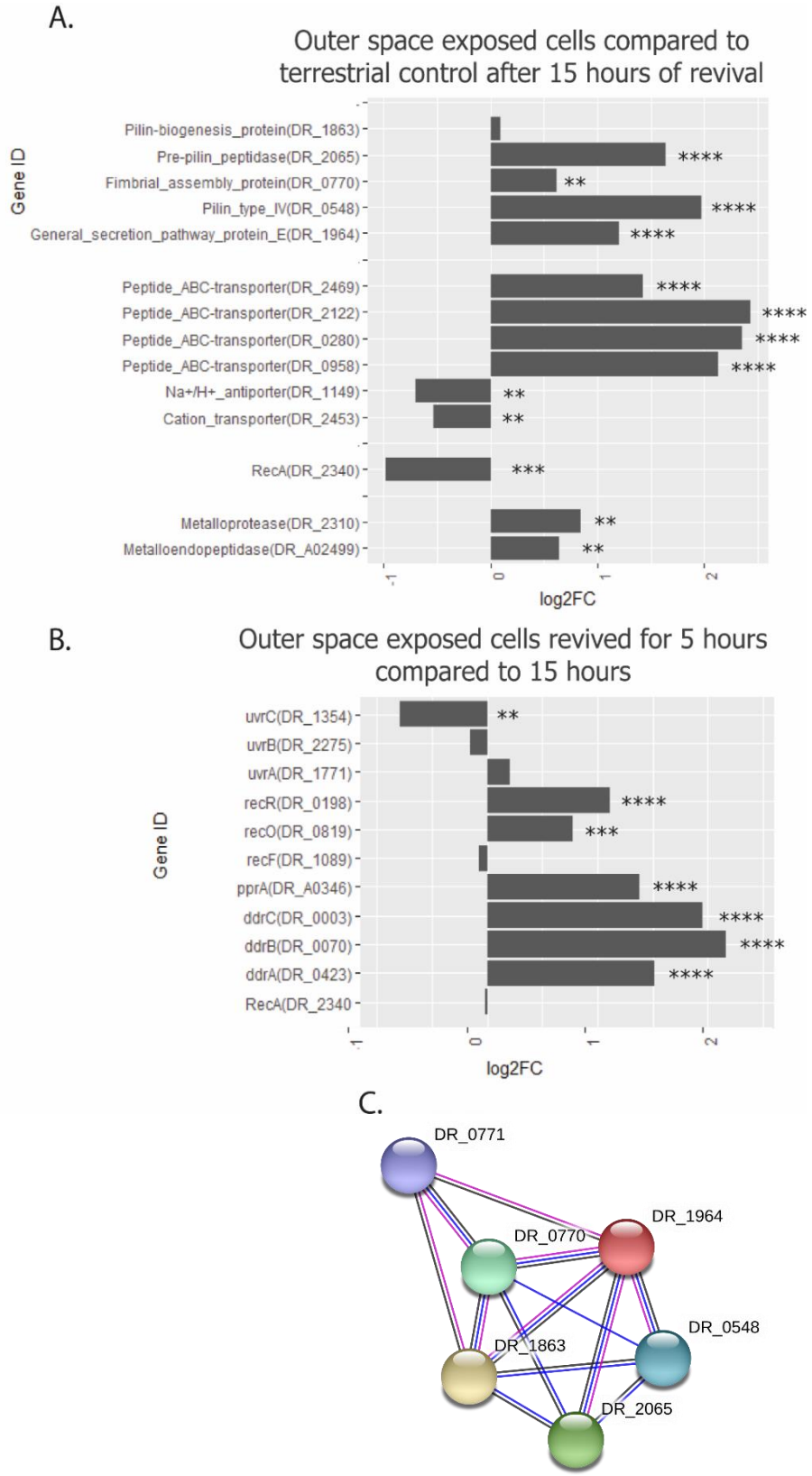


Figure 13. Gene-expression of transport and damage repair associated genes and STRING protein-protein interaction of general secretion pathway protein E.

Expression of genes associated with DNA damage repair and cell trafficking were investigated in ISS exposed cells and compared to ground control cells after 15 hours of recovery (A). Furthermore, ISS exposed cells revived for 5 hours were compared to ISS exposed cells after 15 hours of recovery (B). For the EdgeR analysis used to identify differentially expressed genes, only genes with a $\log_2$ fold-change of more than 1.2 were considered as being overexpressed. X-axis depicts the $\log_2$ of the fold change of respective genes, whose gene name is given on the y-axis. Locus name is depicted in brackets. Level of significance is indicated by asterisks: *= $p<0.05$, **= $p>0.01$, ***= $p>0.001$, ****= $p<0.0001$. P-values were adjusted using the Benjamini-Hochberg method. (C) The protein interaction database was used to identify association of genes with the query input gene DR_1964 (red), showing co expression evidence as black string, co-occurring genes as blue string, and experimental evidence of association as pink string.

adherence like in other bacteria. In fact, according to the STRING database, in

D. radiodurans upregulation of the secretion system protein encoding locus DR_1964 is linked to co-expression of Pilin, type IV encoded by DR_0548, of a Fimbrial assembly protein (DR_0770) and a pre-pilin peptidase type IV (DR_2065) which helps with assembly and secretion [151] and a pilin biogenesis protein (DR_1863) (Fig.13 C). Indeed, all of these were also upregulated in ISS15h (Fig.13 A). The Type IV Pilin locus DR_0548 was even found in the top 200 most significantly upregulated genes in ISS15h datasets (Fig.12 C). Hence, this finding suggested that this cluster of genes was pivotal for the

regeneration of cells after LEO stress. They might act in secretion system homologues like T2SS or T4SS which are dependent on the assembly of pilin or pseudo-pili proteins or they could be encoding pores or other substructures of the cellular membrane. Similar to the better characterized secretin homologue DR_0774 (a pore molecule) which is an important component of the secretory pathway and one of the most abundant components of the cell wall in *D. radiodurans* [152], it can be assumed that the final location of these transcriptionally upregulated pili-associated proteins resides in the cell wall of *D. radiodurans*. Overall, transcriptional upregulation of these genes might be caused by extensive cell wall damage the samples were subjected to. Moreover, the final assembled complex consisting out of these putative pili-associated proteins might aid in the assembly of T2SS and/or T4SS secretion systems which subsequently help the cells during regeneration by excreting damaged and therefore toxic proteins and nucleotides or by inducing the uptake of nutrients upon release of hydrolyzing enzymes. Indeed, the top 200 DE genes revealed a series of significantly upregulated transcripts that code for proteins involved in the degradation of proteins. Amongst them were transcripts encoding a putative serine protease (DR_2325), an extracellular solute-binding protein (DR_1712), and a protein containing a hydrolase_4 domain which is often found in serine aminopeptidases as well as DR_1325 which was annotated as endopeptidase-related protein (Fig.12 C). Moreover, a closer inspection of the expression pattern of aforementioned genes revealed, that most of these protein hydrolyzing enzymes experienced higher expression after 15 hours of revival compared to 5 hours of revival. Overall, this observation allows the assumption that after a longer period of regeneration cells

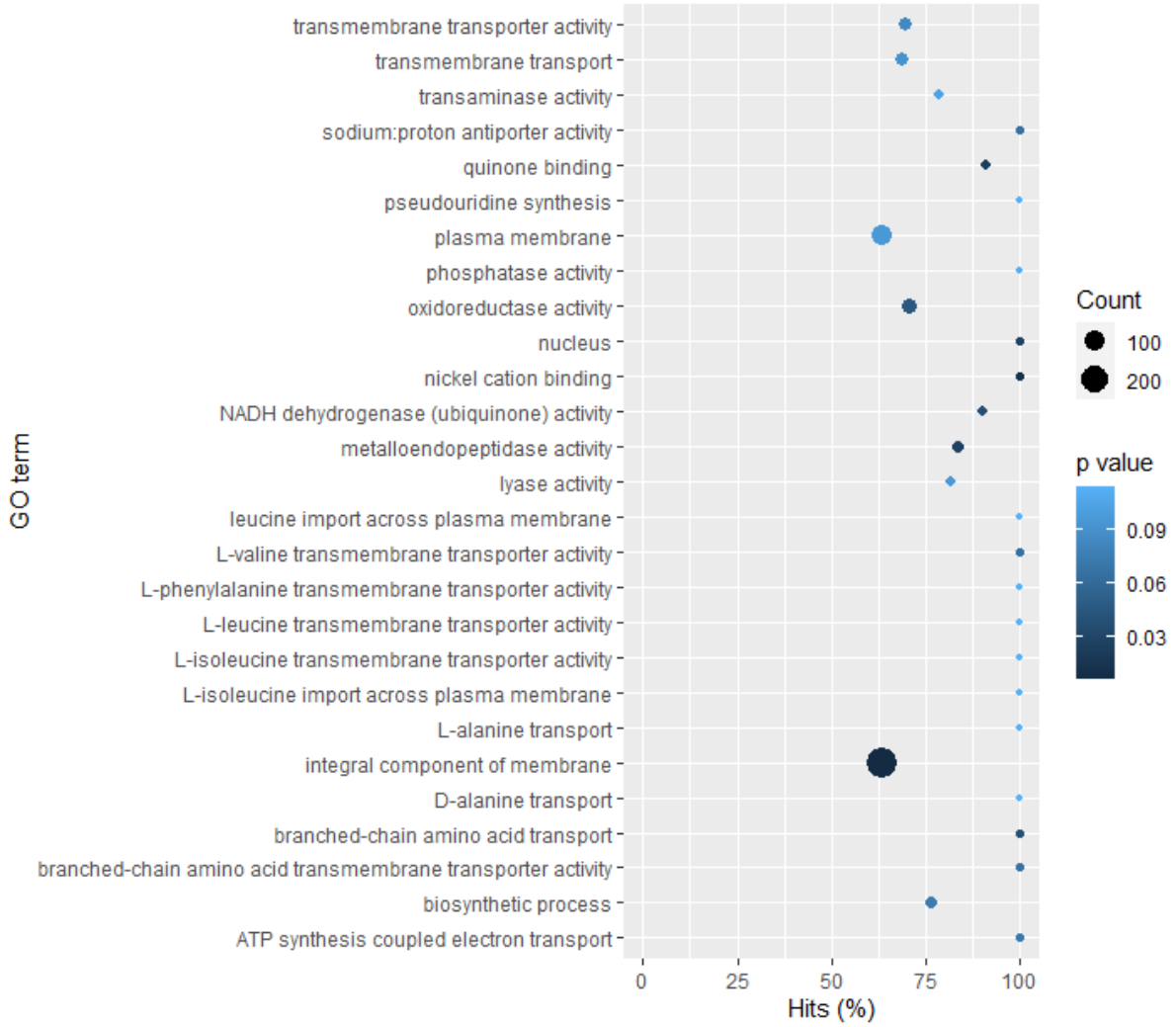


Figure 14. Top 26 results of Gene ontology enrichment analysis in ISS exposed cells revived for 15 hours.
To explore the transcriptional response of *D.radiodurans* to outer space conditions after exposure on the ISS for 3 years, Go-term (gene ontology) enrichment analysis of the RNA seq data was performed using the Goseq Bioconductor tool. The analysis was conducted on RNA retrieved from samples after 5 hours of revival in TGB to investigate the early response and from cells revived for 15 hours to investigate the long-term effects on gene expression. The top 20 results for ISS15h are shown. Size of data points indicate the total number of genes that were assigned with respective Go-term. Go-terms represent upregulated terms belonging to the categories biological process, molecular functions, and cellular component. Hits (%) indicate how many genes that are assigned with a specific Go-term experienced upregulation. Color of data points indicate *p*-value for over representation of the term in the differentially expressed genes. P-values of Go-term enrichment were calculated using the Wallenius method. P-values were corrected using the Benjamini-Hochberg procedure to avoid false positives.

were occupied with overexpressing genes that are important for the degradation of proteins more so than during the early period of regeneration. Hence, some of the proteases could have been used to degrade damaged toxic proteins and using excretion systems to subsequently eliminate the degraded products. Alternatively, after a longer period of regeneration the enzymes could have been employed for hunting proteins. The cells could have preyed upon proteins by excreting the

hydrolyzing enzymes into the extracellular space using T2SS or T4SS in order to salvage the preferred nutrient source of *D. radiodurans*.

2.5.2 Similarities between early and late transcriptional response

A plethora of significant DE transcripts experienced almost equal overexpression during both early and late response such as the exopolyphosphatase DR_A0185, pyridoxal kinase PdxY (DR_A0184), DR_A0183 encoding an uncharacterized protein, purC (DR_0226) and a putative propionyl-CoA carboxylase (DR_1542) (Fig.12 C). Again, some of these genes were possibly organized in operons which were upregulated in response to the treatment, for example the locus ranging from DR_A0183 to DR_A0185.

Another enigmatic locus that experienced overexpression relative to the ground control in both early and later transcriptional response was DR_1735 (Fig.12 C). The locus which was found to be listed in the top 200 overexpressed genes in both ISS5h and ISS15h samples displaying more pronounced overexpression in ISS5h cells, harbored an N-terminus domain of a flagellar assembly protein T (FlgT_N) that makes up the basal body of a flagellum according to the Pfam database. In more detail, the protein FlgT was first characterized in *Vibrio cholerae* [153]. In this organism it is located in the periplasm and is in charge of stabilizing the basal ring-like structure of flagella which consist out of the T and the H ring, hereby anchoring the flagellum in the membrane. However, according to current knowledge, *D. radiodurans* does not produce flagella and is unlikely to perform chemotactic movements [42]. Nevertheless, this locus appeared to play an important role during the response period after LEO stress exposure, considering its apparent upregulation that persisted for at least 15 hours. Hence, this protein presumably serves a different purpose in *D. radiodurans* rather than enabling chemotactic mobility. For instance, it could be possible that the transcribed protein was a part of the cell wall in *D. radiodurans* for example as part of a pore and had to be reproduced in the course of regeneration. Uniprot showed a 20 bp long signal peptide sequence, which indicated that the transcribed protein could have been involved in signal transduction. However, it is also possible that the emergence of this signal peptide does not contribute to the function of the proteins and is rather an artifact of the search engine. Overall, in the then-current state of knowledge it was not feasible to assign any definitive function to the putative FlgT

homologue protein.

2.5.3 Go term analysis of ISS15 in comparison to GC15h cells reveals significant upregulation of endopeptidases in ISS exposed cells.

As with the samples harvested after 5 hours, a Go-term enrichment analysis was performed for cells revived for 15 hours to elucidate the most abundant differentially expressed pathways (Fig.14) and hence paint a more comprehensive picture of the gene expression patterns and late transcriptional response to LEO stress. The most significantly enriched categories included genes for transportation, particularly genes involved in the transportation of branched-chain amino acids. About 70% of the genes assigned with the terms “transmembrane transport” and “transporter activity” were upregulated according to the Goseq software. As for genes that were involved in the transportation of various amino acids in general, several terms were significantly enriched in ISS15h samples compared to GC15h. For example, 100% of the genes decorated with the term “branched-chain amino acid transport”, or genes for the transport of specific amino acids such as L-valine, L-phenylalanine or L-isoleucine were found to be significantly upregulated in ISS15h cells (Fig.14). Additionally, Goseq analysis revealed an enrichment of more than 75% for metalloendopeptidase encoding genes. Indeed, for example a Zn-dependent endopeptidase (DR_A02499) that previously noted as overexpressed in ionized *D. radiodurans* cells [154] or the excreted serralysin metalloprotease DR_2310 [155] experienced a small but significant upregulation (Fig. 13 A). Overall, this observation reiterated the notion stated before that after 15 hours of revival, cells which were exposed to LEO stress avidly produced proteases to degrade proteins and transcriptionally upregulated transporter proteins to facilitate their uptake.

Furthermore, the analysis revealed enrichment of genes linked to biosynthesis (Fig.14). This was indicated by the term biosynthetic process where respective genes were upregulated in 75% of the cases and by the more specific term pseudouridine synthesis where upregulation was found in 100% of the genes (Fig. 14). Hence, ISS exposed cells appeared to be more involved in biosynthetic processes than their terrestrial counterparts, apparently especially in the synthesis of pseudo-uridine as

proposed by the enriched term “pseudouridine synthesis”. Pseudo-uridine is the most common post-transcriptional modification found in all branches of life. Overall, the in-cooperation of pseudo-uridine is linked to stabilization of the phosphate-backbone and improved base-pairing [156]. Additionally, a previous study has shown that this modification can be induced by stress in budding yeast [157]. Hence, multiple scenarios arise that could explain the relative upregulation of pseudo-uridine associated genes. It is possible that the relative upregulation of genes was a byproduct of the overall enhanced biosynthetic process and was upregulated to maintain de status quo of the modification status. On the other hand, it is possible that outer-space exposed cells upregulated pseudo-uridine modification related genes de novo after stress-exposure in order to enhance cell fitness and ensure survival. However, as the biological function of this modification is still largely a mystery, the exact mechanisms of how this protective process might take place is not known and could be a topic of future studies. For example, it might be interesting to investigate whether stress stimuli do indeed lead to an accumulation of pseudo-uridine in bacteria and if so, whether this occurs globally or at specific RNA sites. Furthermore, the modified RNA sites could be assessed in terms of stability or genes expression to infer possible functions of this modification.

Another pathway that seemed to experience upregulation relative to the terrestrial control was the electron transport chain (ETC), illustrated for example by the term “ATP synthesis coupled electron transport”. 100% of the genes assigned with this Go-term were upregulated in the ISS15h. Other enriched Go-terms implying elevated ETC activity were “NADH dehydrogenase ubiquinone activity” and “oxidoreductase activity” (Fig.14), Go-terms that were already overrepresented in the ISS5h and GC5h comparison (Fig.8). Although the Go-category mentions “ubiquinone”, the major respiratory quinone in *Deinococcus sp*. is menaquinone [158]. This demonstrates that GO-categories assigned to genes are oftentimes lacking organism-specific names. In general, NADH-quinone or menaquinone dehydrogenases, respectively, facilitate the electron transfer from NADH to the respective electron acceptor. Hence, they are an important part of the ETC for energy generation across all trees of life [159]. The increased representations of significantly overexpressed genes assigned with this GO-term suggest an elevated demand for the expression of genes that encode respiratory chain dehydrogenases. It can be hypothesized that, like in the ISS5h cells,

there appeared to be an increased need for ATP in ISS15h cells. Hence, it is possible that bacteria tried to meet this demand by increasing electron transfer activity in the cells for which respiratory chain dehydrogenases could have been upregulated.

2.6 Comparing the gene expression patterns between early and late response cells.

Finally, I was interested in how the gene expression of cells harvested after 5 hours was different from the gene expression of cells revived for 15 hours in order to infer gene expression kinetics of the transcriptional response in *D. radiodurans* after outer-space related stress exposure. Overall, 626 genes were significantly upregulated and 613 were downregulated in ISS5h compared to ISS15h cells (Fig.15 A). This meant that even though cells were not growing in numbers as shown by the OD_{600} measurements (Fig.2 C), there was a considerable change in the transcriptional landscape of the cells during the regeneration period. To investigate the changes in the expression profile of the cells in more detail, again a Go-term enrichment analysis was performed with the differentially expressed genes between the comparison of ISS5h with ISS15h (Fig.15 B). The analysis revealed that more than 50% of genes that were assigned a Go-term related to transposition were significantly overexpressed in ISS5h cells, meaning that transposition events were high at the beginning and decreased over time. Interestingly, genes associated with stress sources such as gamma radiation, desiccation, heat, and iron starvation, i.e. stress that is commonly found in an outer space environment, were highly enriched. In fact, 100% of these Go-terms were overexpressed in ISS5h relative to ISS15h. Contrary to the initial assumption, categories of these stress-types did not experience enrichment when comparing ISS5h with GC5h samples (Fig.10). This was surprising since the initial assumption was that starvation, radiation and heat might be one of the primary sources that would influence gene expression during the early phase of recovery after being exposed to outer- space like conditions. Hence, apparently ISS5h and GC5h alike upregulated genes associated with starvation, radiation and heat, leading to no relative upregulation of these categories in ISS5h cells. This included well characterized genes like ddrA, ddrB and ddrC which translate into ssDNA binding proteins involved in DNA damage response after radiation or pprA

which binds dsDNA. All of them were more highly expressed in ISS5h compared to ISS15h (Fig.13 B). The relative upregulation of these stress associated genes in ISS5h compared to ISS15h cells hence can be seen as evidence that these genes truly did experience upregulation during the early phase of regeneration, but it could not be identified when compared to GC5h because the genes were even more upregulated in terrestrial control cells. Furthermore, most genes encoding proteins of *D. radiodurans´* SOS response displayed ambiguous expression patterns during the early transcriptional response compared to 15 hours of regeneration. For example, recF and recA were downregulated in ISS exposed cells at the beginning of the regeneration phase compared to later time points, however not significantly. RecO and recR on the other hand experienced slight but significant elevation (Fig.13 B). Moreover, genes encoding the genes homologues of the UvrABC endonuclease complex which is upregulated as a response to mend UV-radiation induced lesions did not display big changes in the expression of the three key players in ISS15h cells compared to ISS5h (Fig.13 B). During the early phase of regeneration, all three of them experienced lower expression in ISS5h cells compared to GC5h to varying degrees (Fig.8). This observation indicates that the RecFOR and UvrABC pathways did not experience major changes in expression during the 10 hours of additional growth time.

Overall, it seemed that classical stress-response genes like RecA which had Go terms like “cellular response to desiccation” and “to gamma radiation” as well as “DNA repair” attached to it, were not the defining feature of the transcriptomic response to outer-space related stress, at least not when cells were compared to dehydrated ground control cells (Fig 8). Nevertheless, as some of the DNA damage response genes apparently experienced downregulation after 15 hours compared to cells after 5 hours it is possible that DNA repair and SOS response were fully activated in a period between 5 and 15 hours. However, but this active period might have been missed during the analysis. Compared to terrestrial cells, the initial early phase of the transcriptional response in ISS-exposed cells was more focused on energy generation via the glyoxylate bypass, synthesis of ROS scavengers like polyamines and manganese-complexes to protect the proteome from the toxic byproducts of radiation and the uptake of nutrients and ions. These measures might have been taken in order to brace the cells for the upcoming repair process during the next few

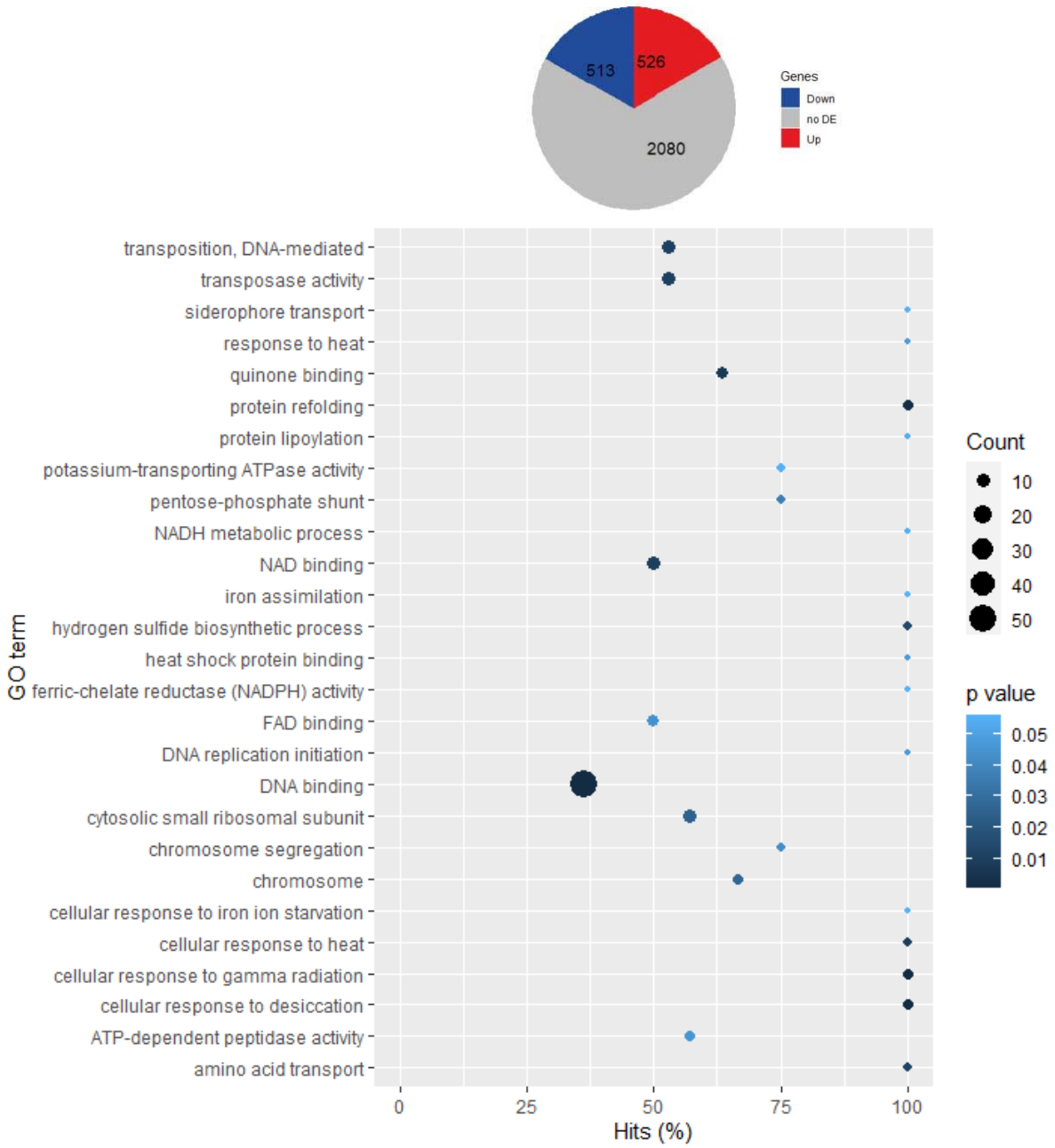


Figure 15. Comparison of the transcriptomes of ISS exposed cells after 5 hours of revival with ISS exposed cells after 15 hours of revival

Differentially expressed (DE) genes were determined using EdgeR and Go-term (gene ontology) enrichment analysis of the RNA seq data was performed using the Goseq Bioconductor tool. The analysis was conducted on RNA retrieved from samples after 5 hours of revival in TGB to investigate the early response and from cells revived for 15 hours to investigate the long-term effects on gene expression (ISS5h and ISS15h respectively). Genes with a $\log_2$ fold-change of more than 1.2 were considered as being overexpressed and quantification of significant DE genes is shown in the pie chart. The top 26 results for the Go-term enrichment in ISS5h cells compared to ISS15h cells are shown below. Size of data points indicate the total number of genes that were assigned with respective Go-term. Go-terms include the categories biological process, molecular functions, and cellular component. Hits (%) indicate how many genes that are assigned with a specific Go-term experienced upregulation compared to the total amount of genes of this category. Color of data points indicate *p*-value for over representation of the term in the differentially expressed genes. P-values of Go-term enrichment were calculated using the Wallenius method. P-values were corrected using the Benjamini-Hochberg procedure to avoid false positives.

hours of

regeneration (Fig.9).

Finally, an enrichment of genes tagged with terms like “amino acid transport” and “ATP-dependent peptidase” was found in ISS5h cells compared to ISS15h. This demonstrated that even though cells were avidly transporting single amino acids like valine, alanine or leucine after 15 hours (Fig. 14), the uptake and processing of peptides were most probably the highest during the first 5 hours and decreased over time. However, overall, it is hard to distinguish between export and import of proteins based on Go-term enrichment analysis, two processes which probably serve very different purposes. It could be argued that the initial early transport might be more geared towards exporting oxidatively damaged and dysfunctional proteins whereas the transportation after 15 hours was geared towards nutrient uptake. A study showed that upon gamma-irradiation, chaperons like DnaK or GroEL are one of the first damaged proteins subjected to proteolytic degradation in the processes of cellular sanitation. However, they are also one of the first to be resynthesized [92] [47]. Indeed, according to EdgeR all genes encoding known chaperons which are typically assigned with the Go-term “proteins folding” were overexpressed in ISS5h cells compared to ISS15h indicating the production of chaperons (Fig.15). Hence, it is likely that early on during regeneration transporters were primarily used to export degraded proteins to prevent the build-up of toxic protein accumulations. In ISS15h cell on the other hand, enriched terms like “Leucine import across plasma membrane” and “metalloendopeptidase acitivity” (Fig.14). Additionally, the elevated expression exhibited by ISS15h cells hint at an elevated proteolytic activity for the sake of nutrient consumption.

2.7 Conclusion and outlook

Overall, the exposure of cells to outer space like environment on the outside of the ISS led to vastly different expression patterns across the transcriptomic landscape in cells which were revived for 5 and 15 hours. Almost half of the expressed genes experienced down or upregulation during the designated early and late transcriptomic response phases (Fig.7 A). However, surprisingly the treatment did not lead to

differences of expression of the initial suspects like for example genes that are involved in DNA damage response. Instead, the analysis revealed upregulation of many formerly unrecognized transcripts, many of which were poorly characterized (Fig.7 C). One of the main contributors to signal transduction and hence transcriptional regulation during the first five hours seemed to be proteins containing a GGDEF domain. The transcript “DR_B0093” encoding a protein harboring this domain did not only appear in the top 200 of most significantly expressed genes, but upregulated transcripts also encoding GGDEF proteins were enriched in ISS5h according to the Go-seq analysis (Fig.10). Hence, according to this study GGDEF domain containing proteins played a major role during signal transduction after LEO-stress and seemed to modulate gene expression to induce cellular recovery. Furthermore, the data displayed the occurrence of IDP in many upregulated transcripts, for example in DR_A0109 and DR_B0004 (Fig.7 C). The function of this enigmatic domain is largely unknown. However, prior studies mention that they could be involved in signal transduction.

Notably, many of the most significantly elevated genes were organized in operons (Fig.13). However, at least for the genes which were better characterized it appeared that they did not necessarily belong to the same pathway of action indicating that they served different purposes and probably acted in different pathways but responded to the same stress stimulus. One of the most important steps during the first 5 hours of cell regeneration seemed to be the upregulation of energy generating pathways. Analysis of the most upregulated genes revealed that in comparison to the terrestrial control, ISS exposed cells displayed elevated expression of the isocitrate lyase, a key enzyme of the glyoxylate bypass and genes involved in the ETC like putative oxidoreductase DR_A0203 and DR_1498 encoding NuoH (Fig.7). Furthermore, the differential gene expression patterns suggested that the cell wall of cells subjected to outer space-conditions endured more damage compared to terrestrial cells. Both early and late response ISS cells exhibited high upregulation of transcripts whose products were “integral part of the membrane” according to the Go-tag analysis compared to their respective control counterparts (Fig.10). This suggests that compared to the terrestrial control, ISS5h cells endured more damage to their cell wall. Hence, cells actively induced the expression of genes encoding components of the cell wall in order to facilitate the reconstruction during the first hours of

regeneration (Fig.9). In addition, the same Go-term was enriched in ISS15h (Fig.14) samples. This indicated that the cells continued to upregulate membrane associated gene even after 15 hours. Moreover, another prominently upregulated set of genes encoded proteins in charge of transportation, such as ABC transporters (Fig.8, Fig.13). Hence, transcriptional elevation of transportation proved to be one of the most prominent differences between ISS exposed and control cells during both late and early regeneration period. However, early and late cells displayed differences in their preferences for what they would traffic. According to the transcriptomic analysis, ISS5 cells would preferably transport ions, for example potassium, copper, or phosphate ions. These could have been used for example as cofactors, to restore osmotic equilibrium or as ROS scavengers. Furthermore, chaperon encoding genes were significantly overexpressed in ISS5h samples compared to ISS15h cells but not compared to GC5h. This was indicated by upregulation of all genes assigned with the Go-tag “protein refolding” (Fig.15). Generally, oxidized chaperons are prone to be degraded and exported first to sanitize the cells. Nevertheless, their expression is quickly upregulated according to prior studies. This co-expression of chaperons and various ABC protein transporters in ISS5h samples could indicate, that during the early period of regeneration cells might primarily use transporters to export broken and toxic proteins. On the other hand, in ISS15h samples, secreted protein-hydrolyzing enzymes were amongst the most highly upregulated genes compared to all other samples (Fig.12, Fig.13 A). This observation suggests that protein trafficking could have had slightly different roles at the beginning of the regeneration compared to later time points. The co-expression of protein hydrolyzing genes and protein transporter genes in these samples could mean that after 15 hours in TGB cells were mainly focused on the uptake of proteins to support their preferred proteolytic lifestyle. Surprising was the very high transcriptional upregulation of the flagellar assembly protein FlgT_N in ISS15h and to a slightly lesser degree in ISS5h samples. Thus far, flagella production has not been found in the *D. radiodurans*, however there are examples of flagella usage in the *Deinococcus* genus such as *Deinococcus multiflagellatum* [160]*.* It is possible that FlgT_N and other co-expressed flagellin associated transcripts were parts of the membrane or might even form T2SS or T4SS protein complexes, however it cannot be excluded that cells exposed to outer space conditions could have used flagella if they indeed belong to a different species than

D. radiodurans or experienced gain-of function mutations.

Another big aspect of regeneration from outer-space related stress seemed to be the synthesis of ROS scavengers. During the early phase of regeneration, upregulation of SodC, PdxY and several other transcripts with implication in the production of ROS scavenging complexes were found to be one of the most highly upregulated genes in the transcriptome of ISS5h cells compared to control cells.

Overall, what probably set apart the early transcriptomic response apart from all other conditions was the high upregulation of transposon associated genes in ISS5h cells (Fig.7). The role of transposition upon stress is a widely discussed topic in current literature. Some suggest that transposons are activated upon stress to create more genetic diversity in order to overcome an adverse environment. On the other hand, transposition could be induced passively by the disruption of regulatory elements or by the upregulation of nearby stress-associated genes. In the ISS exposed cells too, genes associated with transposition seemed to be elevate during the first five hours and might have beneficial effects on cell survival during adverse conditions. Nevertheless, to elucidate the purpose of transposition in this extremophile more in depth analysis must be conducted.

Moreover, the study revealed that in ISS5h samples SOS damage response proteins, DNA damage proteins like ddrD or ddrA, as well as recA and genes of the RecFOR pathway were not upregulated compared to GC5h cells (Fig.8). Further, after 15 hours of regeneration no upregulation was found (Fig.13 A). What was upregulated indeed were genes involved in BER like the putative Nudix hydrolase DR_0603 and several glycosidases as well as PurC facilitating the synthesis of purine bases. Nevertheless, the comparison of ISS5h to ISS15h showed that DNA damage repair genes like, ddrA, ddrB and ddrC and the RecFOR pathway were upregulated in ISS5h samples in contrast to cells that had more time to regenerate (Fig.13 B). These comparisons reveal that surprisingly the classical pathways upregulated in cells during desiccation or ionizing radiation and other DNA damage inducing agents apparently are not the defining feature of the expression pattern in cells that were exposed to outer space-related stress. The very low RIN value of ISS5h leaves little doubt that the cells suffered great damage. However, ground control cells presumably experienced similar amounts of double and single strand breaks, possibly because

of the dehydrated state they were subjected to for 3 years. This seemingly led to overexpression of DNA damage related genes which can mend double and single strand breaks. Apparently however, the amount of base pair lesions was greater in cells that additionally to desiccation were subjected to outer space-related stress factors, which is suggested by the upregulation of BER associated genes like NUDIX hydrolases and various glycosidases (Fig.8).

Finally, some caveats as well as future perspectives of the study are discussed. One of the main shortages of this analysis was the fact that the majority of the genes that make up the transcriptome of *D. radiodurans* are largely unknown. Most upregulated genes were uncharacterized according to Uniprot or, in the best case, were assigned a putative function based on homologies. Very few however, displayed actual known functions based on experimental evidence. This fact has to be considered during analyses like for example Go-term enrichment where many genes were not assigned with Go-terms yet at the time the analysis was conducted. Lack of characterization can influence the outcome of such analysis and other approaches to interpret a transcriptome. Furthermore, it might be possible that the very high degradation level of the RNA in ISS5h cells could have been a potential source of non-biological variation in terms of gene expression. The MDS plot indicated that the normalization conducted via EdgeR was probably sufficient to infer biologically relevant differential gene expression, however the possibility to detect false positives cannot be totally ruled out. Nevertheless, the expression of the proposed candidate genes can be further assessed in future studies using qRT-PCR. The explanatory power of such studies can be further enhanced by testing genes in cells which were subjected to the various stress sources of outer space separately, i.e., microgravity or ionizing radiation similar to outer-space conditions and so on.

Furthermore, due to the limited number of samples, only a few time points were investigated. Considering the RNA degradation level and the growth, cells were not yet finished with their regeneration. Hence it is likely that important parts of the kinetics of the transcriptomic reaction between and after the inspected time points have been missed. Additional, more advanced studies applying methods like SLAMseq [161] which can detect nascent RNA and RNA degradation might be an effective approach to further explore the kinetics of the transcriptional response and

enhance the temporal resolution of interesting candidate genes.

Lastly, it is important to note that the transcriptome of a cell is not necessarily congruent with its proteome. The transcription of mRNA does not necessary result in the translation of it, as a plethora of modulation can influence not only its translation but also its function and its turnover rate. It is therefore advised to be careful before implying function solely based on the rate of transcription. Hence, a proteomic and metabolic approach could alleviate the drawback of this analysis and further substantiate or invalidate the findings of this study. Furthermore, it might be interesting to dissect the multiple stress sources the cells are subjected to onboard of the ISS, i.e. microgravity, radiation, vacuum and desiccation and replicate these conditions on earth regarding their effect on the transcriptome and proteome. Nevertheless, this study managed to identify potential new candidate genes that might aid *D. radiodurans* in its remarkable protection against the multi-factorial stress source “outer-space”. These candidate genes could be further used to generate mutants in order to evaluate if and to what degree they contribute to the resistance of the extremophile. Overall, advancing the knowledge of the stress-related biology of *Deinococcus spp.* and other polyextremophile organisms not only helps to understand how extraterrestrial life might survive under high environmental pressure. It also opens new prospects for future biotechnological applications like the generation of sturdy bioleaching microorganisms or genetic modulation of crops to enhance their perseverance against drought and other stress factors. Lastly, elucidating the molecular mechanisms of survival under extreme conditions like outer space might help us to answer one of humankinds most intriguing questions: does life exist outside of the “spaceship earth”?

3. MATERIAL AND METHODS

3.1 Cultivation

For the Tanpopo exposure experiment *Deinococcus radiodurans* type strain R1 ATCC 13939 and *Deinococcus aerius* strain TR0125 JCM 11750 were used [162]. For the experiment, bacterial pellets were dehydrated using a desiccator (SANYO, SPD−WVGS300) as described in Kawaguchi et al. [163] and stored in sterilized aluminum plates referred to as exposure panels (EP). Eps were attached to the Exposure Handrail Attachment Mechanism (ExHAM) and placed on the Exposure Facility of the Japanese Experiment Module of the ISS were they exposed to the Low earth orbit environment for 3 years.

After receiving samples, dehydrated cell pellets were removed from aluminum plate sample holder and dissolved in 1 ml od cooled 1% PBS (Phosphate-buffered saline, pH7.0) by shaking at 1400rpm for 1 hour at 12 C°. Afterwards, cells were resuspended briefly to evenly dissolve the cells. Afterwards, 60ml of liquid TGB medium [1 vol.-% Tryptone, 0.6 vol.-% Beef Extract, 0.2 vol.-% Glucose] [164] was inoculated with the 1% PBS cell suspension. The medium was then split into 3 equal amounts to generate replicates. Cells were revived in liquid TGB for 5 and 15 hours respectively by shaking at 170 rpm at 30°C. For RNA extraction, cells were then harvested and washed with1%PBS 3 times by centrifugation (3000g for 5 minutes on 4°C). The first two washing steps were performed with 1%PBS, the last one with water that has been purified using the Milli-Q Water system.

For the growth curve, 30ml liquid TGB was inoculated with cells which were stored inside of the ISS (because of a very small number of cells stored outside of the ISS) for 3 years and dissolved in 1% of PBS using aforementioned method. OD_{600} measurements were taken by a spectrophotometer every three hours for 48 hours using TGB medium as blank.

3.2 RNA extraction

After washing, cell pellets were subjected to RNA extraction using TRIzol™ reagent (Invitrogen™) employing the standard protocol [165] if not described otherwise. First TRIzol™ reagent was administered to the pellets, then cells were homogenized using glass bead for shredding. Samples were shredded for 6,5 m/s for a total of 2 minutes with a bead beater homogenizer. In between the shredding procedure after 1 minute, cells were cooled down for 30 seconds by placing them on ice. After homogenization, the phase separation was performed using 200µl chloroform. Samples were incubated in chloroform for 3 minutes on room temperature. After formation of the phases, the aqueous phase was used for RNA precipitation. Precipitation was performed using 2.5 to 3 vols of 100% ice cold ethanol which was added to the aqueous phase and incubated for 10 minutes at room temperature before centrifugation at 4°C for 2 minutes with 12.000 g. Subsequently, the RNA was washed using ice cold 90% Ethanol and then 75% ethanol. For these steps, EconoSpin® RNA Spin Columns were used. Finally, RNA was eluted with RNAse free water and incubated for 10 minutes before being centrifuged for 2 minutes with 6000g.

RNA quality was first assessed using the NanoDrop (Thermofisher scientific) for the initial the spike in experiment using samples stored inside of the ISS (i.e., experiment to investigate RNA extraction efficiency, sRNA depletion and sequency quality). Quality was assessed by measuring the 260/280 and 260/230 absorbance ratio. For the final experiment using the ground control and ISS outside samples, the Agilent RNA 6000 Pico Kit for Bioanalyzer was used to assess RNA quality.

3.3 Sequencing and transcriptomic analysis

Prior to Illumina sequencing, 16S rRNA was depleted using the Ribo-Zero rRNA removal kit (Illumina). Illumina Sequencing was performed using HiSeqV4 SR50. The 16S rRNA depletion, library preparation and sequencing were performed by the Next Generation Sequencing Facility at Vienna BioCenter Core Facilities (VBCF), member of the Vienna BioCenter (VBC), Austria.

Quality of raw reads was evaluated using the tool FastQC. Adapters of raw reads were trimmed using the Trimmomatic tool.

Initially, reads were used for a de-novo transcriptome assembly using the software tool Trinity [99]. Furthermore, comparative genome analysis of the de- novo transcriptome and the published reference genomes of *D. radiodurans* (accession number GCA_000008565.1) and *D. aerius* (accession number BFAG00000000.1) were performed using the Mauve alignment tool [100]. Due to higher similarity to the genome of *D. radiodurans*, the published reference genome was used for further downstream analysis and not the de-novo transcriptome. Investigation of 16S rRNA similarity to the reference gene was performed using the IGV gene viewer, sequence similarity comparisons of aligned 16S rRNA reads were performed manually using BLAST [166].

To quantify the read count of each detected gene, trimmed reads were aligned to the complete reference genome assembly of *D. radiodurans* R1 by calling the command "rsem-calculate-expression -- bowtie2" in the software tool rsem [167]. The resulting expected counts were used to calculate differential expression using the Bioconductor tool EdgeR [168]. For this, an EdgeR specific object named DGEList was created using a matrix of read counts obtained from rsem. Furthermore, reads were filtered using the filterByExpr(DGEList, min.total.count=10) command which resulted in calling only genes with more than 10 respective reads "expressed". Then, using EdgeR, the command calcNormFactors() was used for TMM normalization and quasi-likelihood tests were performed by applying the functions glmQLFit() and qlmQLFit(). The function glmTreat() was used to calculate differential expression between samples with the default minimum fold-change threshold. Heatmaps depicting differential expression were created using counts per million (cpm) which were calculated by calling the command cpm(). Multi-dimensional scaling plots to depict distances between gene expression profiles of samples were created by calling the EdgeR command plotMDS() on the DGEList object. Gene ontology enrichment analysis for RNA sequencing data was performed using the Bioconductor tool GOseq [140]. False discovery rate was calculated using Benjamini-Hochberg correction [140].

5. Bibliography

1. Schein, M., Jesse, W. P. & Wollan, E. O. The nature of the primary cosmic radiation and the origin of the mesotron [4]. *Physical Review* (1941). doi:10.1103/PhysRev.59.615

2. Pilmanis, A. A. & Sears, W. J. Physiological hazards of flight at high altitude. *Lancet* (2003).

3. Richman, M. Survival in Space: Medical Problems of Manned Spaceflight. *JAMA J. Am. Med. Assoc.* (1990). doi:10.1001/jama.1990.03440150134042

4. Chambers, A. *Modern vacuum physics*. *Modern Vacuum Physics* (2004).

5. Jönsson, K. I., Rabbow, E., Schill, R. O., Harms-Ringdahl, M. & Rettberg, P. Tardigrades survive exposure to space in low Earth orbit. *Current Biology* (2008). doi:10.1016/j.cub.2008.06.048

6. Horneck, G. *et al.* Resistance of Bacterial Endospores to Outer Space for Planetary Protection Purposes—Experiment PROTECT of the EXPOSE-E Mission. *Astrobiology* (2012). doi:10.1089/ast.2011.0737

7. de Vera, J. P. Lichens as survivors in space and on Mars. *Fungal Ecol.* (2012). doi:10.1016/j.funeco.2012.01.008

8. Mattimore, V. & Battista, J. R. Radioresistance of Deinococcus radiodurans: Functions necessary to survive ionizing radiation are also necessary to survive prolonged desiccation. *J. Bacteriol.* (1996). doi:10.1128/jb.178.3.633-637.1996

9. Setlow, P. Mechanisms which contribute to the long-term survival of spores of Bacillus species. *J. Appl. Bacteriol.* (1994). doi:10.1111/j.1365-2672.1994.tb04357.x

10. Crowe, J. H. & Madin, K. A. Anhydrobiosis in Tardigrades and Nematodes. *Trans. Am. Microsc. Soc.* (1974). doi:10.2307/3225155

11. Boothby, T. C. *et al.* Tardigrades Use Intrinsically Disordered Proteins to Survive Desiccation. *Mol. Cell* (2017). doi:10.1016/j.molcel.2017.02.018

12. Slade, D., Lindner, A. B., Paul, G. & Radman, M. Recombination and Replication in DNA Repair of Heavily Irradiated Deinococcus radiodurans. *Cell* (2009). doi:10.1016/j.cell.2009.01.018

13. Zea, L. *et al.* Phenotypic changes exhibited by E. coli cultured in space. *Front. Microbiol.* (2017). doi:10.3389/fmicb.2017.01598

14. Ricco, A. J. *et al.* Pharmasat: Drug dose dependence results from an autonomous microsystem-based small satellite in low earth orbit. in *Technical Digest - Solid-State Sensors, Actuators, and Microsystems Workshop* (2010).

15. Kitts, C. *et al.* Flight Results from the GeneSat-1 Biological Microsatellite Mission. *2007 AIAA Small Satell. Conf.* (2007). doi:SSC08-II-4

16. Kim, W. *et al.* Spaceflight Promotes Biofilm Formation by Pseudomonas aeruginosa. *PLoS One* (2013). doi:10.1371/journal.pone.0062437

17. Crabbé, A. *et al.* Transcriptional and proteomic responses of Pseudomonas aeruginosa PAO1 to spaceflight conditions involve Hfq regulation and reveal a role for oxygen. *Appl. Environ. Microbiol.* (2011). doi:10.1128/AEM.01582-10

18. Brim, H. *et al.* Engineering Deinococcus radiodurans for metal remediation in radioactive mixed waste environments. *Nat. Biotechnol.* (2000). doi:10.1038/71986

19. Gupta, P. *et al.* MDP: A Deinococcus Mn2+-decapeptide complex protects mice from ionizing radiation. *PLoS One* (2016). doi:10.1371/journal.pone.0160575

20. Lin, S. M. *et al.* Antioxidant Activities of an Exopolysaccharide (DeinoPol) Produced by the Extreme Radiation-Resistant Bacterium Deinococcus radiodurans. *Sci. Rep.* (2020). doi:10.1038/s41598-019-56141-3

21. Maqbool, I. *et al.* Crude Cell-Free Extract From Deinococcus radiodurans Exhibit Anticancer Activity by Inducing Apoptosis in Triple-Negative Breast Cancer Cells. *Front. Cell Dev. Biol.* (2020). doi:10.3389/fcell.2020.00707

22. Baglioni, P., Sabbatini, M. & Horneck, G. Astrobiology Experiments in Low Earth Orbit: Facilities,Instrumentation, and Results. in *Complete Course in Astrobiology* (2008). doi:10.1002/9783527618996.ch11

23. Wallace, J. M. & Hobbs, P. V. Atmospheric science - an introductory survey. (1977). doi:10.5860/choice.44-3925

24. Cucinotta, F. A., Wilson, J. W., Williams, J. R. & Dicello, J. F. Analysis of MIR-18 results for physical and biological dosimetry: Radiation shielding effectiveness in LEO. *Radiat. Meas.* (2000). doi:10.1016/S1350-4487(99)00273-5

25. Horneck, G. HZE particle effects in space. *Acta Astronaut.* (1994). doi:10.1016/0094-5765(94)90170-8

26. Zimmermann, H., Schäfer, M., Schmitz, C. & Bücker, H. Effects of heavy ions on inactivation and DNA double strand breaks in Deinococcus radiodurans R1. *Adv. Sp. Res.* (1994). doi:10.1016/0273-1177(94)90470-7

27. Horneck, G., Klaus, D. M. & Mancinelli, R. L. Space Microbiology. *Microbiol. Mol. Biol. Rev.* (2010). doi:10.1128/mmbr.00016-09

28. Badhwar, G. D. Radiation measurements in low Earth orbit: U.S. and Russian results. *Health Phys.* (2000). doi:10.1097/00004032-200011000-00007

29. Reisz, J. A., Bansal, N., Qian, J., Zhao, W. & Furdui, C. M. Effects of ionizing radiation on biological molecules - mechanisms of damage and emerging methods of detection. *Antioxidants and Redox Signaling* (2014). doi:10.1089/ars.2013.5489

30. Benderitter, M., Vincent-Genod, L., Pouget, J. P. & Voisin, P. The cell membrane as a biosensor of oxidative stress induced by radiation exposure: A multiparameter investigation. *Radiat. Res.* (2003). doi:10.1667/0033-7587(2003)159[0471:TCMAAB]2.0.CO;2

31. Madian, A. G. & Regnier, F. E. Proteomic identification of carbonylated proteins and their oxidation sites. *J. Proteome Res.* (2010). doi:10.1021/pr1002609

32. Slade, D. & Radman, M. Oxidative Stress Resistance in Deinococcus radiodurans. *Microbiol. Mol. Biol. Rev.* (2011). doi:10.1128/MMBR.00015-10

33. Ito, H., Watanabe, H., Takehisa, M. & Iizuka, H. Isolation and identification of radiation-resistant cocci belonging to the genus deinococcus from sewage sludges and animal feeds. *Agric. Biol. Chem.* (1983). doi:10.1080/00021369.1983.10866087

34. Azzam, E. I., Jay-Gerin, J. P. & Pain, D. Ionizing radiation-induced metabolic oxidative stress and prolonged cell injury. *Cancer Letters* (2012). doi:10.1016/j.canlet.2011.12.012

35. Desouky, O., Ding, N. & Zhou, G. Targeted and non-targeted effects of ionizing radiation. *J. Radiat. Res. Appl. Sci.* (2015). doi:10.1016/j.jrras.2015.03.003

36. Valko, M. *et al.* Free radicals and antioxidants in normal physiological functions and human disease. *International Journal of Biochemistry and Cell Biology* (2007). doi:10.1016/j.biocel.2006.07.001

37. Dias, V., Junn, E. & Mouradian, M. M. The role of oxidative stress in parkinson's disease. *Journal of Parkinson's Disease* (2013). doi:10.3233/JPD-130230

38. James, S. J. *et al.* Metabolic biomarkers of increased oxidative stress and impaired methylation capacity in children with autism. *Am. J. Clin. Nutr.* (2004).

39. Halliwell, B. Oxidative stress and cancer: Have we moved forward? *Biochemical Journal* (2007). doi:10.1042/BJ20061131

40. Work, E. & Griffiths, H. Morphology and chemistry of cell walls of Micrococcus radiodurans. *J. Bacteriol.* (1968).

41. Lancy, P. & Murray, R. G. E. The envelope of Micrococcus radiodurans: isolation, purification, and preliminary analysis of the wall layers. *Can. J. Microbiol.* (1978).

42. Makarova, K. S. *et al.* Genome of the Extremely Radiation-Resistant Bacterium Deinococcus radiodurans Viewed from the Perspective of Comparative Genomics. *Microbiol. Mol. Biol. Rev.* (2001). doi:10.1128/mmbr.65.1.44-79.2001

43. Liu, Y. *et al.* Transcriptome dynamics of deinococcus radiodurans recovering from ionizing radiation. *Proc. Natl. Acad. Sci. U. S. A.* (2003). doi:10.1073/pnas.0630387100

44. GHOSAL, D. *et al.* How radiation kills cells: Survival of and under oxidative stress. *FEMS Microbiol. Rev.* (2005). doi:10.1016/j.femsre.2004.12.007

45. Daly, M. J. *et al.* Small-molecule antioxidant proteome-shields in Deinococcus radiodurans. *PLoS One* (2010). doi:10.1371/journal.pone.0012570

46. Zhang, Y. M., Wong, T. Y., Chen, L. Y., Lin, C. S. & Liu, J. K. Induction of a futile Embden-Meyerhof-Parnas pathway in Deinococcus radiodurans by Mn: Possible role of the pentose phosphate pathway in cell survival. *Appl. Environ. Microbiol.* (2000). doi:10.1128/AEM.66.1.105-112.2000

47. Zhang, C. *et al.* Proteomic analysis of Deinococcus radiodurans recovering from γ-irradiation. *Proteomics* (2005). doi:10.1002/pmic.200300875

48. Thornley, M. J., Horne, R. W. & Glauert, A. M. The fine structure of micrococcus radiodurans. *Arch. Mikrobiol.* (1965). doi:10.1007/BF00408143

49. Beblo, K. *et al.* Survival of thermophilic and hyperthermophilic microorganisms after exposure to UV-C, ionizing radiation and desiccation. *Arch. Microbiol.* (2011). doi:10.1007/s00203-011-0718-5

50. Zahradka, K. *et al.* Reassembly of shattered chromosomes in Deinococcus radiodurans. *Nature* (2006). doi:10.1038/nature05160

51. Bentchikou, E., Servant, P., Coste, G. & Sommer, S. A major role of the RecFOR pathway in DNA double-strand-break repair through ESDSA in Deinococcus radiodurans. *PLoS Genet.* (2010). doi:10.1371/journal.pgen.1000774

52. Driedger, A. A., James, A. P. & Grayston, M. J. Cell Survival and X-Ray-Induced DNA Degradation in Micrococcus radiodurans. *Radiat. Res.* (1970). doi:10.2307/3573160

53. Lett, J. T., Feldschreiber, P., Little, J. G., Steele, K. & Dean, C. J. The repair of x-ray damage to the deoxyribonucleic acid in Micrococcus radiodurans: a study of the excision process. *Proc. R. Soc. London. Ser. B. Biol. Sci.* (1967). doi:10.1098/rspb.1967.0023

54. White, O. *et al.* Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1. *Science (80-.).* (1999). doi:10.1126/science.286.5444.1571

55. Bernstein, D. A. *et al.* Crystal structure of the Deinococcus radiodurans single-stranded DNA-binding protein suggests a mechanism for coping with DNA damage. *Proc. Natl. Acad. Sci. U. S. A.* (2004). doi:10.1073/pnas.0401331101

56. Uranga, L. A., Reyes, E. D., Patidar, P. L., Redman, L. N. & Lusetti, S. L. The cohesin-like RecN protein stimulates RecA-mediated recombinational repair of DNA double-strand breaks.

Nat. Commun. (2017). doi:10.1038/ncomms15282

57. George, N. P. *et al.* Structure and cellular dynamics of deinococcus radiodurans single-stranded DNA (ssDNA)-binding protein (SSB)-DNA complexes. *J. Biol. Chem.* (2012). doi:10.1074/jbc.M112.367573

58. Daly, M. J. & Minton, K. W. An alternative pathway of recombination of chromosomal fragments precedes recA-dependent recombination in the radioresistant bacterium Deinococcus radiodurans. *J. Bacteriol.* (1996). doi:10.1128/jb.178.15.4461-4471.1996

59. Gutman, P. D., David Carroll, J., lan Masters, C. & Minton, K. W. Sequencing, targeted mutagenesis and expression of a recA gene required for the extreme radioresistance of Deinococcus radiodurans. *Gene* (1994). doi:10.1016/0378-1119(94)90124-4

60. Moseley, B. E. B. & Copland, H. J. R. Isolation and properties of a recombination deficient mutant of Micrococcus radiodurans. *J. Bacteriol.* (1975).

61. Narumi, I. *et al.* Molecular analysis of the Deinococcus radiodurans recA locus and identification of a mutation site in a DNA repair-deficient mutant, rec30. *Mutat. Res. - DNA Repair* (1999). doi:10.1016/S0921-8777(99)00048-8

62. Lipton, M. S. *et al.* Global analysis of the Deinococcus radiodurans proteome by using accurate mass tags. *Proc. Natl. Acad. Sci. U. S. A.* (2002). doi:10.1073/pnas.172170199

63. Carroll, J. D., Daly, M. J. & Minton, K. W. Expression of recA in Deinococcus radiodurans. *J. Bacteriol.* (1996). doi:10.1128/jb.178.1.130-135.1996

64. Lecointe, F., Shevelev, I. V., Bailone, A., Sommer, S. & Hübscher, U. Involvement of an X family DNA polymerase in double-stranded break repair in the radioresistant organism Deinococcus radiodurans. *Mol. Microbiol.* (2004). doi:10.1111/j.1365-2958.2004.04233.x

65. Khairnar, N. P. & Misra, H. S. DNA polymerase X from Deinococcus radiodurans implicated in bacterial tolerance to DNA damage is characterized as a short patch base excision repair polymerase. *Microbiology* (2009). doi:10.1099/mic.0.029223-0

66. Michel, B. Replication fork arrest and DNA recombination. *Trends in Biochemical Sciences* (2000). doi:10.1016/S0968-0004(00)01560-7

67. Omelchenko, M. V. *et al.* Comparative genomics of Thermus thermophilus and Deinococcus radiodurans: Divergent routes of adaptation to thermophily and radiation resistance. *BMC Evol. Biol.* (2005). doi:10.1186/1471-2148-5-57

68. Harris, D. R. *et al.* Preserving genome integrity: The DdrA protein of Deinococcus radiodurans R1. *PLoS Biol.* (2004). doi:10.1371/journal.pbio.0020304

69. Jolivet, E. *et al.* Limited concentration of RecA delays DNA double-strand break repair in Deinococcus radiodurans R1. *Mol. Microbiol.* (2006). doi:10.1111/j.1365-2958.2005.04946.x

70. Norais, C. A., Chitteni-Pattu, S., Wood, E. A., Inman, R. B. & Cox, M. M. DdrB protein, an alternative Deinococcus radiodurans SSB induced by ionizing radiation. *J. Biol. Chem.* (2009). doi:10.1074/jbc.M109.010454

71. Sugiman-Marangos, S. N., Weiss, Y. M. & Junop, M. S. Mechanism for accurate, protein-assisted DNA annealing by deinococcus radiodurans DdrB. *Proc. Natl. Acad. Sci. U. S. A.* (2016). doi:10.1073/pnas.1520847113

72. Narumi, I. *et al.* PprA: A novel protein from Deinococcus radiodurans that stimulates DNA ligation. *Mol. Microbiol.* (2004). doi:10.1111/j.1365-2958.2004.04272.x

73. Tanaka, M. *et al.* Analysis of Deinococcus radiodurans's transcriptional response to ionizing radiation and desiccation reveals novel proteins that contribute to extreme radioresistance. *Genetics* (2004). doi:10.1534/genetics.104.029249

74. Battista, J. R. AGAINST ALL ODDS:The Survival Strategies of Deinococcus radiodurans . *Annu. Rev. Microbiol.* (1997). doi:10.1146/annurev.micro.51.1.203

75. Kitayama, S. Adaptive repair of cross-links in DNA of Micrococcus radiodurans. *BBA - Gene Struct. Expr.* (1982). doi:10.1016/0167-4781(82)90103-8

76. Daly, M. J. A new perspective on radiation resistance based on Deinococcus radiodurans. *Nat. Rev. Microbiol.* (2009). doi:10.1038/nrmicro2073

77. Daly, M. J. *et al.* Accumulation of Mn(II) in Deinicoccus radiodurans facilitates gamma-radiation resistance. *Science (80-.).* (2004). doi:10.1126/science.1103185

78. Gutman, P. D., Fuchs, P. & Minton, K. W. Restoration of the DNA damage resistance of Deinococcus radiodurans DNA polymerase mutants by Escherichia coli DNA polymerase I and Klenow fragment. *Mutat. Res. Repair* (1994). doi:10.1016/0921-8777(94)90064-7

79. Schlesinger, D. J. Role of RecA in DNA damage repair in Deinococcus radiodurans. *FEMS Microbiol. Lett.* (2007). doi:10.1111/j.1574-6968.2007.00862.x

80. Xu, G. *et al.* RecO is essential for DNA damage repair in Deinococcus radiodurans. *J. Bacteriol.* (2008). doi:10.1128/JB.01851-07

81. Daly, M. J. *et al.* Protein oxidation implicated as the primary determinant of bacterial radioresistance. *PLoS Biol.* (2007). doi:10.1371/journal.pbio.0050092

82. Fredrickson, J. K. *et al.* Protein oxidation: Key to bacterial desiccation resistance? *ISME J.* (2008). doi:10.1038/ismej.2007.116

83. Von Frijtag Drabbe Künzel, J. K., Van Der Zee, J. & Ijzerman, A. P. Radical scavenging properties of adenosine and derivatives in vitro. *Drug Dev. Res.* (1996). doi:10.1002/(SICI)1098-2299(199601)37:1<48::AID-DDR3>3.0.CO;2-M

84. Driedger, A. A. & Grayston, M. J. The enhancement of X-ray-induced DNA degradation in Micrococcus radiodurans by phenethyl alcohol. *Can. J. Microbiol.* (1971). doi:10.1139/m71-081

85. Hariharan, P. V. & Cerutti, P. A. Formation and repair of γ-ray induced thymine damage in Micrococcus radiodurans. *J. Mol. Biol.* (1972). doi:10.1016/S0022-2836(72)80006-8

86. Kota, S., Kumar, C. V. & Misra, H. S. Characterization of an ATP-regulated DNA-processing enzyme and thermotolerant phosphoesterase in the radioresistant bacterium Deinococcus radiodurans. *Biochem. J.* (2010). doi:10.1042/BJ20100446

87. Xu, W., Shen, J., Dunn, C. A., Desai, S. & Bessman, M. J. The Nudix hydrolases of Deinococcus radiodurans. *Mol. Microbiol.* (2001). doi:10.1046/j.1365-2958.2001.02267.x

88. Doolittle, R. F. *et al.* Domainal evolution of a prokaryotic DNA repair protein and its relationship to active-transport proteins. *Nature* (1986). doi:10.1038/323451a0

89. Servant, P. *et al.* The ClpPX protease is required for radioresistance and regulates cell division after γ-irradiation in Deinococcus radiodurans. *Mol. Microbiol.* (2007). doi:10.1111/j.1365-2958.2007.06003.x

90. Makarova, K. S., Aravind, L., Daly, M. J. & Koonin, E. V. Specific expansion of protein families in the radioresistant bacterium Deinococcus radiodurans. *Genetica* (2000). doi:10.1023/A:1004035424657

91. Dalle-Donne, I. *et al.* Protein carbonylation, cellular dysfunction, and disease progression. *Journal of Cellular and Molecular Medicine* (2006). doi:10.1111/j.1582-4934.2006.tb00407.x

92. Joshi, B., Schmid, R., Altendorf, K. & Apte, S. K. Protein recycling is a major component of post-irradiation recovery in Deinococcus radiodurans strain R1. *Biochem. Biophys. Res. Commun.* (2004). doi:10.1016/j.bbrc.2004.06.062

93. Rouault, T. A. & Klausner, R. D. Iron-sulfur clusters as biosensors of oxidants and iron. *Trends in Biochemical Sciences* (1996). doi:10.1016/S0968-0004(96)10024-4

94. Arrhenius, S. & Borns, H. Worlds in the Making. The Evolution of the Universe. *Bull. Am. Geogr. Soc.* (1909). doi:10.2307/200804

95. Kawaguchi, Y. *et al.* Investigation of the Interplanetary Transfer of Microbes in the Tanpopo Mission at the Exposed Facility of the International Space Station. *Astrobiology* (2016). doi:10.1089/ast.2015.1415

96. Horváth, V., Merenciano, M. & González, J. Revisiting the Relationship between Transposable Elements and the Eukaryotic Stress Response. *Trends in Genetics* (2017). doi:10.1016/j.tig.2017.08.007

97. Gallego Romero, I., Pai, A. A., Tung, J. & Gilad, Y. RNA-seq: Impact of RNA degradation on

transcript quantification. *BMC Biol.* (2014). doi:10.1186/1741-7007-12-42

98. Satoh, K. *et al.* Draft genome sequence of the radioresistant bacterium Deinococcus aerius TR0125, isolated from the high atmosphere above Japan. *Genome Announc.* (2018). doi:10.1128/genomeA.00080-18

99. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* (2011). doi:10.1038/nbt.1883

100. Darling, A. E., Tritt, A., Eisen, J. A. & Facciotti, M. T. Mauve assembly metrics. *Bioinformatics* (2011). doi:10.1093/bioinformatics/btr451

101. Varghese, N. J. *et al.* Microbial species delineation using whole genome sequences. *Nucleic Acids Res.* (2015). doi:10.1093/nar/gkv657

102. Konstantinidis, K. T. & Tiedje, J. M. Genomic insights that advance the species definition for prokaryotes. *Proc. Natl. Acad. Sci. U. S. A.* (2005). doi:10.1073/pnas.0409727102

103. Rainey, F. A. *et al.* 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.* (2005). doi:10.1128/AEM.71.9.5225-5235.2005

104. Zhang, Z., Schwartz, S., Wagner, L. & Miller, W. A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology* (2000). doi:10.1089/10665270050081478

105. Kim, M. & Chun, J. 16S rRNA gene-based identification of bacteria and archaea using the EzTaxon server. in *Methods in Microbiology* (2014). doi:10.1016/bs.mim.2014.08.001

106. Robinson, J. T. *et al.* Integrative genomics viewer. *Nature Biotechnology* (2011). doi:10.1038/nbt.1754

107. Morgulis, A. *et al.* Database indexing for production MegaBLAST searches. in *Bioinformatics* (2008). doi:10.1093/bioinformatics/btn322

108. Robinson, J. T., Thorvaldsdóttir, H., Wenger, A. M., Zehir, A. & Mesirov, J. P. Variant review with the integrative genomics viewer. *Cancer Research* (2017). doi:10.1158/0008-5472.CAN-17-0337

109. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* (2010). doi:10.1186/gb-2010-11-3-r25

110. Dillies, M. A. *et al.* A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Brief. Bioinform.* (2013). doi:10.1093/bib/bbs046

111. Lin, Y. *et al.* Comparison of normalization and differential expression analyses using RNA-Seq data from 726 individual Drosophila melanogaster. *BMC Genomics* (2016).

doi:10.1186/s12864-015-2353-z

112. Bateman, A. *et al.* UniProt: The universal protein knowledgebase. *Nucleic Acids Res.* (2017). doi:10.1093/nar/gkw1099

113. Akiyama, M., Crooke, E. & Kornberg, A. An exopolyphosphatase of Escherichia coli. The enzyme and its ppx gene in a polyphosphate operon. *J. Biol. Chem.* (1993).

114. Rao, N. N., Gómez-García, M. R. & Kornberg, A. Inorganic Polyphosphate: Essential for Growth and Survival. *Annu. Rev. Biochem.* (2009). doi:10.1146/annurev.biochem.77.083007.093039

115. Santos, S. P. *et al.* The interplay between Mn and Fe in Deinococcus radiodurans triggers cellular protection during paraquat-induced oxidative stress. *Sci. Rep.* (2019). doi:10.1038/s41598-019-53140-2

116. Grillo-Puertas, M., Schurig-Briccio, L. A., Rodríguez-Montelongo, L., Rintoul, M. R. & Rapisarda, V. A. Copper tolerance mediated by polyphosphate degradation and low-affinity inorganic phosphate transport system in Escherichia coli. *BMC Microbiol.* (2014). doi:10.1186/1471-2180-14-72

117. Dulermo, R. *et al.* Identification of new genes contributing to the extreme radioresistance of Deinococcus radiodurans using a Tn5-based transposon mutant library. *PLoS One* (2015). doi:10.1371/journal.pone.0124358

118. Farci, D., Guadalupi, G., Bierla, K., Lobinski, R. & Piano, D. The role of iron and copper on the oligomerization dynamics of DR_2577, the main S-layer protein of deinococcus radiodurans. *Front. Microbiol.* (2019). doi:10.3389/fmicb.2019.01450

119. Makarova, K. S. *et al.* Deinococcus geothermalis: The pool of extreme radiation resistance genes shrinks. *PLoS One* (2007). doi:10.1371/journal.pone.0000955

120. Cheng, C. *et al.* Aminopeptidase T of M29 Family Acts as A Novel Intracellular Virulence Factor for Listeria monocytogenes Infection. *Sci. Rep.* (2015). doi:10.1038/srep17370

121. Lu, C. D. Pathways and regulation of bacterial arginine metabolism and perspectives for obtaining arginine overproducing strains. *Applied Microbiology and Biotechnology* (2006). doi:10.1007/s00253-005-0308-z

122. Katz, A. M. *et al.* Spermine Condenses DNA, but Not RNA Duplexes. *Biophys. J.* (2017). doi:10.1016/j.bpj.2016.11.018

123. Valdés-Santiago, L. & Ruiz-Herrera, J. Stress and polyamine metabolism in fungi. *Front. Chem.* (2014). doi:10.3389/fchem.2013.00042

124. Di Salvo, M. L., Contestabile, R. & Safo, M. K. Vitamin B 6 salvage enzymes: Mechanism, structure and regulation. *Biochimica et Biophysica Acta - Proteins and Proteomics* (2011).

doi:10.1016/j.bbapap.2010.12.006

125. Bilski, P., Li, M. Y., Ehrenshaft, M., Daub, M. E. & Chignell, C. F. Vitamin B6 (Pyridoxine) and Its Derivatives Are Efficient Singlet Oxygen Quenchers and Potential Fungal Antioxidants. *Photochem. Photobiol.* (2007). doi:10.1562/0031-8655(2000)0710129sipvbp2.0.co2

126. Hallberg, Z. F. *et al.* Structure and mechanism of a hypr GGDEF enzyme that activates cGAMP signaling to control extracellular metal respiration. *Elife* (2019). doi:10.7554/eLife.43959

127. Jenal, U., Reinders, A. & Lori, C. Cyclic di-GMP: Second messenger extraordinaire. *Nature Reviews Microbiology* (2017). doi:10.1038/nrmicro.2016.190

128. Wilson, J. W. *et al.* Space flight alters bacterial gene expression and virulence and reveals a role for global regulator Hfq. *Proc. Natl. Acad. Sci. U. S. A.* (2007). doi:10.1073/pnas.0707155104

129. Crabbé, A. *et al.* Spaceflight enhances cell aggregation and random budding in Candida albicans. *PLoS One* (2013). doi:10.1371/journal.pone.0080677

130. Yeh, J. I., Chinte, U. & Du, S. Structure of glycerol-3-phosphate dehydrogenase, an essential monotopic membrane enzyme involved in respiration and metabolism. *Proc. Natl. Acad. Sci. U. S. A.* (2008). doi:10.1073/pnas.0712331105

131. Uversky, V. N. Intrinsically disordered proteins and their 'Mysterious' (meta)physics. *Frontiers in Physics* (2019). doi:10.3389/fphy.2019.00010

132. Deiana, A., Forcelloni, S., Porrello, A. & Giansanti, A. Intrinsically disordered proteins and structured proteins with intrinsically disordered regions have different functional roles in the cell. *PLoS One* (2019). doi:10.1371/journal.pone.0217889

133. Basile, W., Salvatore, M., Bassot, C. & Elofsson, A. Why do eukaryotic proteins contain more intrinsically disordered regions? *PLoS Comput. Biol.* (2019). doi:10.1371/journal.pcbi.1007186

134. Garay-Arroyo, A., Colmenero-Flores, J. M., Garciarrubio, A. & Covarrubias, A. A. Highly hydrophilic proteins in prokaryotes and eukaryotes are common during conditions of water deficit. *J. Biol. Chem.* (2000). doi:10.1074/jbc.275.8.5668

135. Caramelo, J. J. & Iusem, N. D. When cells lose water: Lessons from biophysics and molecular biology. *Progress in Biophysics and Molecular Biology* (2009). doi:10.1016/j.pbiomolbio.2008.10.001

136. Crowe, J. H., Hoekstra, F. A. & Crowe, L. M. Anhydrobiosis. *Annu. Rev. Physiol.* (1992). doi:10.1146/annurev.ph.54.030192.003051

137. Kish, A. *et al.* Salt shield: Intracellular salts provide cellular protection against ionizing

radiation in the halophilic archaeon, Halobacterium salinarum NRC-1. *Environ. Microbiol.* (2009). doi:10.1111/j.1462-2920.2008.01828.x

138. Bessman, M. J., Frick, D. N. & O'Handley, S. F. The MutT proteins or 'Nudix' hydrolases, a family of versatile, widely distributed, 'housecleaning' enzymes. *Journal of Biological Chemistry* (1996). doi:10.1074/jbc.271.41.25059

139. Preumont, A., Snoussi, K., Stroobant, V., Collet, J. F. & Van Schaftingen, E. Molecular identification of pseudouridine-metabolizing enzymes. *J. Biol. Chem.* (2008). doi:10.1074/jbc.M804122200

140. Young, M. D., Wakefield, M. J., Smyth, G. K. & Oshlack, A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* (2010). doi:10.1186/gb-2010-11-2-r14

141. Frösler, J., Panitz, C., Wingender, J., Flemming, H. C. & Rettberg, P. Survival of Deinococcus geothermalis in Biofilms under Desiccation and Simulated Space and Martian Conditions. *Astrobiology* (2017). doi:10.1089/ast.2015.1431

142. Mennecier, S., Servant, P., Coste, G., Bailone, A. & Sommer, S. Mutagenesis via IS transposition in Deinococcus radiodurans. *Mol. Microbiol.* (2006). doi:10.1111/j.1365-2958.2005.04936.x

143. Pasternak, C. *et al.* Irradiation-induced Deinococcus radiodurans genome fragmentation triggers transposition of a single resident insertion sequence. *PLoS Genet.* (2010). doi:10.1371/journal.pgen.1000799

144. Vandecraen, J., Chandler, M., Aertsen, A. & Van Houdt, R. The impact of insertion sequences on bacterial genome plasticity and adaptability. *Critical Reviews in Microbiology* (2017). doi:10.1080/1040841X.2017.1303661

145. Jacob, F. & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology* (1961). doi:10.1016/S0022-2836(61)80072-7

146. Price, M. N., Arkin, A. P. & Alm, E. J. The life-cycle of operons. *PLoS Genet.* (2006). doi:10.1371/journal.pgen.0020096

147. Stav, S. *et al.* Genome-wide discovery of structured noncoding RNAs in bacteria. *BMC Microbiol.* (2019). doi:10.1186/s12866-019-1433-7

148. Filloux, A. The underlying mechanisms of type II protein secretion. *Biochimica et Biophysica Acta - Molecular Cell Research* (2004). doi:10.1016/j.bbamcr.2004.05.003

149. Craig, L., Pique, M. E. & Tainer, J. A. Type IV pilus structure and bacterial pathogenicity. *Nature Reviews Microbiology* (2004). doi:10.1038/nrmicro885

150. Nivaskumar, M. & Francetic, O. Type II secretion system: A magic beanstalk or a protein escalator. *Biochimica et Biophysica Acta - Molecular Cell Research* (2014).

doi:10.1016/j.bbamcr.2013.12.020

151. Liles, M. R., Edelstein, P. H. & Cianciotto, N. P. The prepilin peptidase is required for protein secretion by and the virulence of the intracellular pathogen Legionella pneumophila. *Mol. Microbiol.* (1999). doi:10.1046/j.1365-2958.1999.01239.x

152. Farci, D. *et al.* New features of the cell wall of the radio-resistant bacterium Deinococcus radiodurans. *Biochim. Biophys. Acta - Biomembr.* (2014). doi:10.1016/j.bbamem.2014.02.014

153. Terashima, H. *et al.* Insight into the assembly mechanism in the supramolecular rings of the sodium-driven Vibrio flagellar motor from the structure of FlgT. *Proc. Natl. Acad. Sci. U. S. A.* (2013). doi:10.1073/pnas.1222655110

154. Karlin, S. & Mrázek, J. Predicted highly expressed and putative alien genes of Deinococcus radiodurans and implications for resistance to ionizing radiation damage. *Proc. Natl. Acad. Sci. U. S. A.* (2001). doi:10.1073/pnas.081077598

155. Basu, B. & Apte, S. K. A novel serralysin metalloprotease from Deinococcus radiodurans. *Biochim. Biophys. Acta - Proteins Proteomics* (2008). doi:10.1016/j.bbapap.2008.05.009

156. Ge, J. & Yu, Y. T. RNA pseudouridylation: New insights into an old modification. *Trends in Biochemical Sciences* (2013). doi:10.1016/j.tibs.2013.01.002

157. Wu, G., Xiao, M., Yang, C. & Yu, Y. T. U2 snRNA is inducibly pseudouridylated at novel sites by Pus7p and snR81 RNP. *EMBO J.* (2011). doi:10.1038/emboj.2010.316

158. Yamada, Y., Takinami, H., Tahara, Y. & Kondo, K. The menaquinone system in the classification of radiation-resistant micrococci. *J. Gen. Appl. Microbiol.* (1977). doi:10.2323/jgam.23.105

159. Lennarz, W. J. & Lane, M. D. *Encyclopedia of Biological Chemistry: Second Edition*. *Encyclopedia of Biological Chemistry: Second Edition* (2013).

160. Kim, D. U. *et al.* Deinococcus multiflagellatus sp. nov., isolated from a car air-conditioning system. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* (2018). doi:10.1007/s10482-017-0982-8

161. Herzog, V. A. *et al.* Thiol-linked alkylation of RNA to assess expression dynamics. *Nat. Methods* (2017). doi:10.1038/nmeth.4435

162. Kawaguchi, Y. *et al.* The Possible Interplanetary Transfer of Microbes: Assessing the Viability of Deinococcus spp. Under the ISS Environmental Conditions for Performing Exposure Experiments of Microbes in the Tanpopo Mission. *Orig. Life Evol. Biosph.* (2013). doi:10.1007/s11084-013-9346-1

163. Kawaguchi, Y. *et al.* DNA Damage and Survival Time Course of Deinococcal Cell Pellets During 3 Years of Exposure to Outer Space. *Front. Microbiol.* (2020).

doi:10.3389/fmicb.2020.02050

164. *Compendium of Methods for the Microbiological Examination of Foods*. *Compendium of Methods for the Microbiological Examination of Foods* (2015). doi:10.2105/mbef.0222

165. TRI reagent. *Cold Spring Harb. Protoc.* (2008). doi:10.1101/pdb.caut2701

166. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* (1990). doi:10.1016/S0022-2836(05)80360-2

167. Li, B. & Dewey, C. N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* (2011). doi:10.1186/1471-2105-12-323

168. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* (2009). doi:10.1093/bioinformatics/btp616