The Ecology and Evolutionary History

of

the *Deinococcaceae* Family

A thesis

submitted in fulfilment

of the requirements for the degree

of

Doctor of Philosophy in Biological Sciences

at

The University of Waikato

By

Jaber Abbaszadeh

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2024

Thesis Abstract

Deinococcus radiodurans is the most studied member of the *Deinococcaceae* family for its outstanding capability to withstand extreme doses of gamma radiation and prolonged desiccation. *D. radiodurans* is also type species and the first described member of the *Deinococcus* genus. It demonstrates approximately 3000 times more radiation resistance capability than humans and 50 times more than *Escherichia coli*. The remarkable radiation resistance phenotype in *D. radiodurans* and certain members of this family has established the conventional belief that radiation resistance is the defining characteristic of *Deinococcaceae*. Consequently, microbiologists used gamma radiation as a pre-culture treatment to enrich radioresistant isolates in environmental samples before cultivation. However, a few studies did not use gamma radiation as an enrichment method and isolated sensitive *Deinococcus* species. These findings demonstrated that the radiation resistance capacity is highly diverse among *Deinococcus* species, with variations in the order of magnitude ranging from 1.5 kGy to 15 kGy. Because of this variability, the prevailing view is that genomic data cannot predict radiation resistance and that this phenotype remains without a genotype. Moreover, *Deinococcus* members have been isolated from diverse environments, including Antarctica, hot springs, deserts, atmospheric dust, animals, and garden soil. Nevertheless, the origins of this ubiquity have remained unexplored. Here, I discuss how the biased isolation methods and focus on highly resistant species created a gap in our understanding of the true ecology and evolution of *Deinococcaceae*. This thesis addresses this knowledge gap using comparative genomics methods by answering three questions. (I) What factors drive the ubiquity of the *Deinococcus* species? (ii) What drives the diversity of radiation resistance in the *Deinococcus* genus, and can we predict this phenotype using genomic data? (iii) How did the radiation resistance phenotype evolve, and what are the primary needs for the common ancestor of *Deinococcaceae* (proto-*Deinococcus*)? Our results indicate that the ubiquity of the *Deinococcus* genus arises from its large pangenome. A high abundance of transposase genes and efficient homologous recombination mechanisms enhanced their genetic flexibility and enabled *Deinococcus* to survive in diverse habitats. Conversely, this ecologically advantageous trait forms a dynamic accessory genome and impacts the radiation resistance phenotype. Phylogenomic and statistical analyses revealed a pattern in the radiation resistance level. The presence of 186 gene families, mainly related to metal and redox-related cofactors, can predict lower levels of resistance in two different clades of the *Deinococcus* genus. This observation

contradicts the conventional wisdom that radiation resistance is the defining characteristic of the *Deinococcaceae* family because the accessory genome controls this phenotype. Furthermore, we used the gene tree-species tree reconciliation method to infer the evolutionary events that gave rise to the proto-*Deinococcus*. The ancestral reconstruction analysis indicated that the proto-*Deinococcus* underwent significant genome expansion by gaining 1000 gene families, which is about 50% of the genome size of its ancestor, while *Thermaceae* retained the same genome size as their common ancestor. Overall, we conclude that the ubiquity of *Deinococcus* species is the result of acquiring diverse functions and surviving in various environments. While this is a beneficial evolution for the family, gaining specific genes can lower the radiation resistance capability. Therefore, radiation resistance is not the defining characteristic of *Deinococcaceae*. Moreover, ancestral reconstruction reveals that this proto-*Deinococcus* underwent genome expansion and likely gained genes that induced radiation resistance traits and genomic flexibility.

DEDICATION

I dedicate this PhD thesis to the lasting memory of Professor Craig Cary, who tragically passed away on my final day as his student.

His legacy is eternal.

More than an exemplary mentor, he was an exceptional friend

and a father figure in my life.

ACKNOWLEDGEMENTS

I started my PhD one year before the onset of the COVID-19 pandemic, a period that presented numerous challenges and uncertainties. Nonetheless, I made it to the finish line! I want to express my deepest gratitude to the exceptional individuals who supported and guided me throughout this transformative journey, enabling me to become a better thinker and human being.

First and foremost, I would like to thank my PhD supervisors, Associate Professor Charles Lee, Professor Craig Cary, and Dr Craig Herbold. Their constant support, guidance, and encouragement have been invaluable throughout the entire process. From the initial stages of refining my research proposal to the final submission of my thesis, their unwavering presence and wealth of wisdom have been instrumental in shaping my academic growth. I highly valued the weekly meetings, which not only served as crucial checkpoints to keep me on track academically but also provided me with plenty of encouragement. I am profoundly grateful for the immeasurable contributions they made to my development.

I am incredibly grateful for the opportunity to be a part of the Thermophile Research Unit (TRU), an exceptional group that provides deep insights into a broad range of marine and terrestrial environmental ecology. I want to thank my friends, lab mates, colleagues and research team – Anya Noble (my PhD buddy), Dr Shelly Brandt, Gamaliel Cabria, Dr Maria Monteiro, Jacinda Challis, Shir Livne, Dr Huw Richards, Dr Alexis Marshall, Dr Mafalda Baptista, and Dr Stephen Noell, Professor Ian McDonald, Roanna Richards-Babbage, and John Longmore for a cherished time spent together in the lab, and in social settings. I would also like to thank my Master's supervisor, Associate Professor Mojtaba Mohseni, for his support when I started my scientific journey.

Throughout my PhD journey, I have been fortunate to experience living in beautiful New Zealand and meeting people who have become my family. Saeed Nikghalb, Amin Rahdarian, Dalareen Fernandez, Lia Missena, Jeremy Toms, Dr Ali Vahabzadeh, Dr Jamal Abarashi, Karol Valeska, Dr Hossein Hosseini, Dr Renée Gilgen, and Rodney Arnold; I am grateful for your companionship and support from the very beginning of my PhD.

With profound gratitude, I extend my appreciation to the most precious thing in my life, my family. To Mom and Dad, your relentless dedication and unwavering commitment to providing for Faezeh and me, coupled with a consistent emphasis on the significance of education, have indelibly shaped my journey. Faeze and Majid, thank you for your enduring love and support. Leila, as a constant presence in my life over the past decade, through both the highs and lows, your encouragement has been an invaluable driving force for my accomplishment.

Table of contents

Table of contents for figures

Table of contents for tables

Description of supplementary data (separate file)

Supplementary Data 2 - Chapter II:

S2-1: Metadata for *Deinococcus* species used in the dataset of Chapter 2. All genomes belonging to the *Deinococcus* genus are presented.

S2-2: Rarefaction data for 10,000 samplings. Started with three genomes and added one genome till all genomes were added

S2-3: Pangenome composition dataset

Supplementary Data 3 - Chapter III:

S3-1: Metadata for the compiled dataset. This dataset includes *Deinococcus* species with a known D_{10} value and available genome in the public database, plus the two genomes sequenced in this study.

S3-2: outcome of the statistical tests of the accessory genes calculated for the N6 sensitive clades against species in moderate resistance and highly resistance groups

S3-3: outcome of the statistical tests for the accessory genes calculated for the N31 sensitive clades against species in moderate resistance and highly resistance groups

S3-4: List of orthogroups that are significantly correlated with radiation resistance levels and functional annotation results for clade N6

S3-5: List of orthogroups that are significantly correlated with radiation resistance levels and functional annotation results for clade N31

Supplementary Data 4 - data for Chapter IV:

S4-1: Metadata and list of genomes used in the main dataset for all *Deinococcota* phylum

S4-2: Metadata and list of subsample genomes used for ancestral reconstruction

S4-3: Functional annotation of orthogroups gained in N002 node (Common ancestor of *Deinococcaceae* and *Thermaceae*)

S4-4: Functional annotation of orthogroups gained in N006 node (Common ancestor of *Deinococcaceae* and *Trueperaceae*)

S4-5: Functional annotation of orthogroups gained in N006 node (proto-*Deinococcus*)

1 Chapter I – Background

1.1 Introduction to the thesis

1.1.1 *Deinococcus* genus

This PhD study attempts to demonstrate the bias around the *Deinococcus* genus and provides insight into the evolution and ecology of this enigmatic group of bacteria. The *Deinococcus* genus is best known for its type species, *Deinococcus radiodurans*, one of the most radiation-resistant organisms known to science (Slade & Radman, 2011). While only 5 to 10 grays (Gy) of gamma radiation kill humans within days, *D. radiodurans* shows no reduced viability after exposure to 5 kilograys (kGy) and has a D_{10} value (the dose required to eliminate 90 per cent of the population) of 12 kGy (Daly, 2009; Sharma et al., 2017). In comparison, *Thermus thermophilus,* a close relative of *D. radiodurans*, has a D_{10} value of 0.7 kGy, and this value for the model organism *Escherichia coli* is 0.8 kGy (Daly, 2009). *D. radiodurans* cells survived three years of exposure to the low orbit environment on the International Space Station (Kawaguchi et al., 2020). When dried and frozen, *D. radiodurans* cells survived a striking dose of 140 kGy and retained their metabolic activity when revived (Horne et al., 2022). *D. radiodurans* can grow under exposure to chronic irradiation of 60 Gy/h without affecting its growth rate in nutrient-rich conditions (Daly, 2000).

D. radiodurans was first isolated from irradiated meat can in 1956 and named *Micrococcus radiodurans* because of its morphological and physiological similarities with the *Micrococcus* genus (Anderson et al., 1956). Later, in 1981, Brooks and Murray used 16S rRNA gene classification and reclassified six radioresistant species in the *Micrococcus* genus to a new genus, *Deinococcus,* within a new family, *Deinococcaceae* (Brooks & Murray, 1981). *Deinococcus* was named after the Greek words *deinos*, meaning strange or unusual, and *coccus*, meaning grain or berry and became the type genus of the *Deinococcaceae* family (Brooks & Murray, 1981). This extremely radioresistant bacterium also exhibited high tolerance to desiccation and other forms of oxidative stress and became the model organism for radiation resistance studies (Daly, 2023; Slade & Radman, 2011).

Ionising radiation and UV rays are lethal for all organisms due to their direct and indirect impacts on macromolecules. The direct damage is through DNA modifications and double-strand breaks (DSB), and the indirect impact is the generation of reactive oxygen species (ROS) molecules, which causes DSBs. DSBs are the least repaired among all DNA damages, and their frequency is generally directly related to cell death (Slade & Radman, 2011).

Decades of research have answered important questions about the extreme radiation resistance capabilities observed in *D. radiodurans* and other members of the *Deinococcus* genus like *D. geothermalis*, *D. ficus*, and *D. deserti* (Makarova et al., 2007; Shashidhar et al., 2010). In the early years, multiple genome copy numbers (4-8 copies per cell in *D.radiodurans*), along with homologous recombination repair mechanisms, were proposed as a mechanism for radiation resistance in *Deinococcus* (Daly et al., 1994; Grimsley et al., 1991; Hansen, 1978). In 2009, Michael J. Daly and his colleagues proposed a new explanation for radiation resistance where a cell's proteome, rather than its genome, is the prime target for radiation-induced cell death, and the accumulation of Mn^{2+} antioxidants governs radiation resistance in *Deinococcus* and other vegetative radiation-resistant organisms by protecting their proteins (Daly, 2009, 2023; Daly et al., 2004; Horne et al., 2022; Sharma et al., 2017).

The latest theory proposed by Daly et al. (2009) has been tested multiple times and explains resistance to high doses of radiation. Our understanding of radiation resistance phenotype had two implications. First, it enabled gauging radiation resistance across all kingdoms of life using electron paramagnetic resonance (EPR) spectroscopy and by measuring concentrations of Mn^{2+} complexes (Sharma et al., 2017). Second, it facilitated the design of irradiated vaccines using a synthetic complex called Mn-decapeptide-phosphate complex (MnDpPi) inspired by cytosolic Mn^{2+} composition in *Deinococcus*. Supplementing MnDpPi resulted in substantial epitope preservation even at a very high dose of 50 kGy and safely inactivated viruses in the alphavirus vaccines (Gayen et al., 2017).

1.1.2 Enigmatic nature of *Deinococcaceae*

"In 1986, Robert Murray sent a memorandum to microbiologists in which he accurately predicted that new information would continue to make a difference in our understanding of *Deinococcaceae* and its relatives" (Daly, 2023). Ever since, microbiologists have significantly progressed in decoding the basis for radiation resistance through the protection of proteome by a complex of Mn^{2+} antioxidants (Gaidamakova et al., 2022). However, many other aspects of this peculiar group of organisms have remained unknown, and the endeavour to describe those unanswered questions continues.

One mystery about the members of the *Deinococcus* genus is the inconsistency in their radiation resistance capabilities. *Deinococcus* species generally show diverse capacities in tolerating ionising radiation and desiccation (Battista, 2016; Shashidhar et al., 2010). For instance, *D. persicinus,* a species isolated from soil, is ten times more sensitive to gamma radiation than *D. radiodurans* (Jeon et al., 2016). Reportedly, *Deinococcus* species with different resistance levels did not show a significant difference in the concentration of Mn^{2+} when they were compared (Shashidhar et al., 2010). Also, some studies have reported that even strains of the same species showed variable D_{10} values (de Groot et al., 2005; Slade & Radman, 2011; Sun Joo et al., 2016). This variability in radiation resistance has remained a mysterious phenomenon.

Members of the *Deinococcus* genus express highly variable traits and occur over a wide geographic distribution within the natural and built environments. Hundreds of culture-dependent and independent studies have reported the presence of *Trueperaceae* and *Deinococcaceae* members (through isolation or amplicon sequencing) in diverse environments, including deserts, Antarctica, hot springs, animal guts and faeces, marine animals, and metal surfaces (Chen et al., 2011; Kaur et al., 2018; Kim et al., 2015; Makarova et al., 2007; Rainey et al., 2005; Rosenberg, 2014; Sleytr et al., 1976). This ubiquity and the fact that *Deinococcus* species can be found in diverse habitats complicated the understanding of their lifestyle. This geographical dispersal and phenotypic variability can be linked to the genetic flexibility of *Deinococcaceae*.

Horizontal Gene Transfer (HGT) in *D. radiodurans* is facilitated by transposase enzymes, its natural competence, and DSB repair enzymes such as RecA and RecFOR pathways (Ithurbide et al., 2020). The natural competence enables higher rates of HGT and provides a foundation for genomic diversity in *Deinococcus*. As a consequence, some *Deinococcus* species have acquired unique and rare traits. For instance, *D. multiflagellatus* isolated from a biofilm in a car air-conditioning system (Kim et al., 2018) is exceptional because it is the only known motile member of the *Deinococcota* phylum with peritrichous flagella. Flagella are complex structures and highly conserved among the flagellate bacteria (Chevance & Hughes, 2008). Typically, up to 50 genes are required to assemble, maintain and function these surface appendages (Gao et al., 2014). Flagella plays a crucial role in the ecological success of some phyla like enterobacteria through their motility and contributes to their pathogenicity in adherence, invasion and colonisation of host cells and tissues (De Maayer et al., 2020). Acquiring such a complex trait by a *Deinococcus* species highlights their genetic flexibility, enabling them to evolve in diverse habitats.

As a result of this genetic flexibility in *Deinococcaceae* and despite decades of research on this enigmatic family, our knowledge about their native habitat and their phenotypic and genomic diversity has remained very limited. The overarching goal of this thesis is to gain insight into the ecology and evolution of the *Deinococcus* genus through the lens of comparative genomics. In this thesis, I sequence two genomes which were previously isolated in a study by Callegan et, al and represent radiosensitive *Deinococcus* species to the currently available genomes in public databases (Callegan et al., 2008). Those genomes were added to the genomes available in the public database to answer three main questions. (i) What drives the ubiquity and diversity of the *Deinococcus* genus? (ii) Can radiation resistance be predicted by comparative genomics? (iii) What were the primary needs of a proto-*Deinococcus* in its natural habitat?

1.2 Research background

The *Deinococcaceae* family consists of two genera with validly published names, *Deinococcus* and *Deinobacterium* (Rosenberg, 2014). The *Deinobacterium* genus has only one known member, *Deinobacterium chartate*, isolated from a paper mill machine. It is thermophilic and highly resistant to gamma radiation, D_{10} value > 10 (Ekman et al., 2011) and diverged from the common ancestor of *Deinococcus* based on phylogenomic analysis. Because there is only one known isolate and genome from the *Deinobacterium* genus, this thesis will focus mainly on the *Deinococcus* genus rather than the *Deinococcaceae* family; therefore, we exclude *Deinobacterium* from our analysis and focus on the *Deinococcus* genus. I discuss the literature around the *Deinococcus* genus in three main themes. (i) Cultivation and isolation of *Deinococcus* species, (ii) Radiation resistance in *Deinococcus,* (iii) Resistance to desiccation

1.2.1 Radiation resistance in *D. radiodurans*

In biology, radiation typically refers to two types of high-energy electromagnetic waves: Ultraviolet (UV) and ionising radiation (IR) (Slade & Radman, 2011). These waves can cause severe damage to biomolecules, including DNA, proteins, and lipids in cells, through two primary mechanisms: direct impact on nucleotides by forming pyrimidine and purine photoproducts and indirect damage through the generation of free radicals in the form of Reactive Oxygen Species (ROS), subsequently leading to strand breaks, primarily single-strand breaks (SSBs) (Kciuk et al., 2020). In this thesis, we use the term 'radiation' interchangeably with 'ionising radiation'; the terms' resistant/resistance' usually refer to 'radiation resistant/resistance' unless stated otherwise.

DSBs caused by ROS and the direct impacts of IR are the most lethal form of DNA damage, preventing genome replication and leading to cell death (Cox & Battista, 2005). The numbers of DSBs per Gy per genome are similar between radiation-resistant and radiation-sensitive species, with about 0.002 to 0.006 DSBs/Gy/Mbp (Daly et al., 2010; Gérard et al., 2001; Gladyshev & Meselson, 2008). However, radiation-resistant organisms like *Deinococcus* benefit from diverse mechanisms enabling them to

maintain cellular function during and after exposure to IR and ROS. Preserving proteins, especially DNA repair machinery, allows them to reassemble the fragmented DNA molecules, replicate, and express their genome (Krisko & Radman, 2013). Studies of radiation resistance mechanisms in *D. radiodurans* can be discussed in three main areas: (i) efficient DNA repair mechanisms, (ii) proteome protection and cell cleansing and (iii) regulatory pathways.

1.2.1.1 DNA repair mechanisms

Studying DNA repair mechanisms has been vital to understanding the radiation resistance phenotype in *D. radiodurans* and other radioresistant organisms. Decades of research have discovered a broad range of DNA repair mechanisms in *D. radiodurans*, including base excision repair (BER), nucleotide excision repair (NER), mismatch repair (MMR), and most importantly, homologous recombination (HR) (Lim et al., 2018). Some examples of discovered proteins in *D. radiodurans* are (AlkA, MutM, MutY, Mpg.) for BER, (UvrABCD, Mfd) for NER, and (RecA, RadA, RecFOR, RuvABC, RecQ) for HR (Table 1-1. Proteins necessary for the four main DNA repair mechanisms:). In this section, I do not attempt to present an in-depth analysis of all mechanisms discovered in *D. radiodurans*. However, I discuss the most efficient DSB repair mechanism in *Deinococcus*, which is a two-stage RecA-dependent homologous recombination.

DSBs are the most challenging form of damage to repair, as both helices of the DNA are missing (Cox & Battista, 2005). High doses of ionising radiation and the presence of ROS cause hundreds of DSBs in the DNA molecule and are lethal if not repaired. Homologous recombination is the most efficient mechanism to repair DSBs accurately but requires intact DNA fragments as templates (Li & Heyer, 2008). It has been shown that *D. radiodurans* has multiple genome copies in all stages, which range from 2 to 8 copies per cell (Slade & Radman, 2011). Even though polyploidy is not causal for radiation resistance, it is a requirement because it provides the template for homologous recombination. Nonetheless, many polyploid microorganisms are not radioresistant (Horne et al., 2022; Omelchenko et al., 2005; Soppa, 2014).

Miroslav Radman and his team discovered an efficient double-stage RecA-dependent recombinational process where *D. radiodurans* uses "extended synthesis-dependent strand annealing" (ESDSA) followed by homologous recombination by crossovers to fix hundreds of DSBs and assemble the whole genome (Slade et al., 2009; Zahradka et al., 2006) (Figure 1-1). *D. radiodurans* lacks RecBC homologs in *E.coli* and alternatively uses the RecFOR pathway (Buljubašić et al., 2019; Lim et al., 2018). First, UvrD and RecJ helicase process the DSBs ends into 3' single-stranded DNA substrates where the RecFOR complex loads RecA recombinase (Bentchikou et al., 2010). Next, a complex of RecA and RadA, prime DNA repair synthesis on partially overlapping fragments to the templates. Then, DNA Pol III initiates DNA synthesis, with two possible scenarios: elongation proceeds by Pol III with Pol I filling up gaps arising from excision repair of damaged bases or (D2), and second, by Pol I alone (Slade et al., 2009; Zahradka et al., 2006). In the second stage, the overlapping, long, linear fragments, or linear intermediates longer than the chromosome, require crossovers by means of RecA-dependent homologous recombination (HR) to assemble the genome into its circular form (Figure 1-1)(Slade et al., 2009).

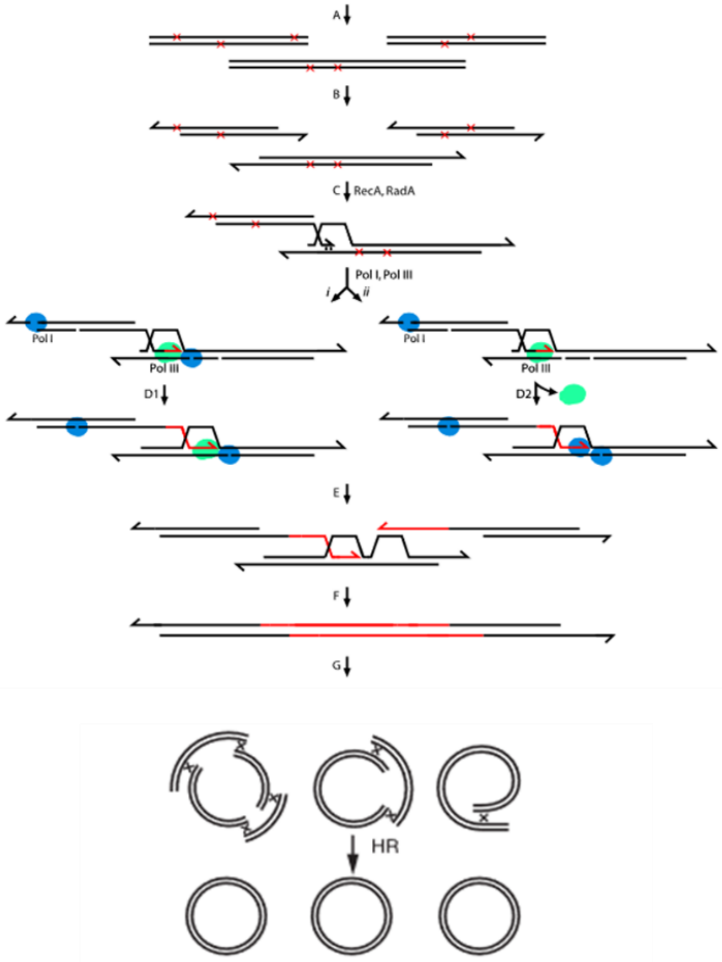

Figure 1-1. A Model for ESDSA DNA Repair in *D. radiodurans*

Following severe DNA damage (ionising radiation, desiccation) (A), the fragmented DNA is end recessed in 5′→3′ direction liberating single-stranded 3′ overhangs (B), which, through RecA- and RadA-mediated strand invasion, prime synthesis on overlapping fragments (C). DNA synthesis is initiated by Pol III and (D1) elongated by Pol III with Pol I filling up gaps arising from excision repair of damaged bases or (D2) by Pol I alone. (E) Two noncontiguous fragments are linked by convergent elongations on a third "bridging" fragment. (F) Newly synthesised single strands anneal to complementary single-stranded extensions forming dsDNA intermediates, which are (G) assembled into intact circular chromosomes by RecA-mediated homologous recombination. Previously published material was used without changes (Slade et al., 2009).

BER fixes small lesions in the DNA double helix structure caused by chemical DNA alterations. In BER, glycosylases like AlkA,.MutM, MutY, and Mpg cleave the bonds between deoxyribose and excision and repair a large number of base lesions, each of which is recognised by one or more overlapping specific DNA glycosylases. The DNA glycosylases recognise and remove damaged bases, leaving a gap which is processed by short or long-patch repair, which involves different proteins (Krokan & Bjørås, 2013).

NER in *D. radiodurans* involves two main pathways: UvrABC and UvsE. UvsE, is Mn^{2+} dependent endonuclease with specificity for pyrimidine dimers and is induced by UV radiation, and the UV damage endonuclease (UvsE)-dependent excision repair (UVER) pathway can effectively remove cyclobutane pyrimidine dimers (CPDs) and pyrimidine (6-4) pyrimidine photoproducts (6-4PPs) from genomic DNA (Tanaka et al., 2004) (Tanaka et al. 2005).

Table 1-1. Proteins necessary for the four main DNA repair mechanisms:

Base excision repair BER		
AlkA	COG0122	3-methyladenine DNA glycosylase/8-oxoguanine DNA glycosylase
Mpg	COG2094	3-methyladenine DNA glycosylase
Ung	COG0692	Uracil DNA glycosylase
Udg4	COG1573	Uracil-DNA glycosylase
Mug	COG3663	G:T/U-mismatch repair DNA glycosylase
MutY	COG1194	Adenine-specific DNA glycosylase, acts on AG and A-oxoG pairs
MutM (Fpg)	COG0266	Formamidopyrimidine-DNA glycosylase
Nth	COG0177	Endonuclease III
XthA	COG0708	Exonuclease III
Mismatch repair MMR		
MutL	COG0323	DNA mismatch repair ATPase
MutS	COG0249	DNA mismatch repair ATPase
MutS2	COG1193	dsDNA-specific endonuclease/ATPase
Nucleotide excision repair NER		
UvrA-A2	COG0178	Excinuclease UvrABC, ATPase subunit
UvrB	COG0556	Excinuclease UvrABC, helicase subunit
UvrC	COG0322	Excinuclease UvrABC, nuclease subunit
UvrD	COG0210	Superfamily I DNA or RNA helicase
SSL2 (Rad25)	COG1061	Superfamily II DNA or RNA helicase (multi-form)
Mfd	COG1197	Transcription-repair coupling factor (superfamily II helicase)
Uve (UvsE)	COG4294	UV DNA damage repair endonuclease
Atl1	COG3695	Alkylated DNA nucleotide flippase Atl1, Ada-like DNA-binding domain
Homologous Recombinational repair HR. ...		
RecA	COG0468	Protein RecA; Recombinase A
RecF	COG1195	Recombinational DNA repair ATPase RecF
RecO	COG1381	DNA repair protein RecO; Recombination protein O
RecR	COG0353	Recombinational DNA repair protein RecR
RecJ	COG0608	Single-stranded DNA-specific exonuclease
RecN	COG0497	DNA repair ATPase RecN
RecQ	COG0514	Superfamily II DNA helicase RecQ (including HRDC domains)
SbcC	COG0419	DNA repair exonuclease SbcCD, ATPase subunit
SbcD	COG0420	DNA repair exonuclease SbcCD, nuclease subunit
RuvA	COG0632	Holliday junction resolvasome RuvABC, DNA-binding subunit
RuvB	COG2255	Holliday junction resolvasome RuvABC, DNA helicase subunit
RuvC	COG0817	Holliday junction resolvasome RuvABC, endonuclease subunit
RecG	COG1200	RecG-like helicase
RecX	COG2137	SOS response regulatory protein OraA/RecX, interacts with RecA
RadA (Sms)	COG1066	DNA repair protein RadA

*Genes in this table were vetted from review articles (Lim et al., 2018; Slade & Radman, 2011)

1.2.1.2 Proteome protection and metal homeostasis

Extensive studies on DNA repair mechanisms and whole genome sequencing of *D. radiodurans* could not answer many questions about radiation resistance mechanisms (Makarova et al., 2007; White et al., 1999). Early genomic studies revealed that most repair enzymes in *D. radiodurans* are homologs of known enzymes in radiation-sensitive bacteria like *E. coli* and that *D. radiodurans* does not have any remarkable enzyme involved in radiation resistance (Makarova et al., 2001; Omelchenko et al., 2005). Later studies noted that the number of genes considered "essential" in radiation resistance was further reduced since genes thought necessary for radiation resistance in *D. radiodurans* were absent in *D. geothermalis* (Makarova et al., 2007). However, it was suggested that if the proteome is protected, the repair machinery can maintain the cell function before the following cell cycle and duplication, which requires an intact genome (Daly, 2009)

In the early years of radiobiology, the classical view was that DNA damage is responsible for the cytotoxic impacts of ionising radiation (Hutchinson, 1966). In 2004, Michael J Daly showed that radioresistant organisms accumulate high concentrations of Mn^{2+}, which protects proteins rather than DNA, and challenged the conventional wisdom (Daly et al., 2004). Later, his team proposed a new theory that the cytotoxic impacts of ionising radiation are mainly due to protein oxidation rather than DNA damage. They showed that Mn^{2+}-metabolite complexes accumulated in extremely resistant cells can protect enzymes needed to repair DNA and allow survival (Daly, 2009; Daly et al., 2010; Gaidamakova et al., 2022). They showed that protein-free cell extracts of *D. radiodurans* protected proteins from high doses of IR and increased the radiation resistance of *E. coli* and human Jurkat T cells. However, the chemical basis of resistance to these oxidising conditions in Mn^{2+}-accumulating bacteria remained unknown (Daly et al., 2010). Other studies investigated radiation resistance in different kingdoms of life, including archaea, bacteria, fungi, and simple animals and confirmed that intracellular accumulation of Mn^{2+} antioxidants complex peptides and

orthophosphate (Pi) is universal in radioresistant organisms (Barnese et al., 2008; Horne et al., 2022; Robinson et al., 2011).

Unlike Mn^{2+}, which is the most efficient ROS scavenger, Fe^{2+} ions are highly toxic to proteins in the presence of ROS agents. Iron usually exists as a cofactor in proteins, and when it is exposed to radical Oxygen, the high energy $O^{\bullet}$ liberates the Fe^{2+}, which can react with H_2O_2 in the Fenton reaction to produce $OH^{\bullet}$ (Imlay, 2006). It has been shown that *Deinococcus* overcame this problem by reducing its dependence on iron-containing proteins. Ghosal showed that *D. radiodurans* has fewer proteins with iron-sulfur clusters ([2Fe-2S] and [4Fe-4S]) than the radiation-sensitive microorganism *Shewanella oneidensis* (Ghosal et al., 2005). They also showed that *D. radiodurans* lack most of the Fe-chelating and transport systems found in radiation-sensitive bacteria. Instead, the necessary iron is sequestered outside the cytosol in the septum between dividing cells. As a result, a high ratio of Mn/Fe is usually another indicator of radiation and desiccation-resistant organisms (Fredrickson et al., 2008). Santos and colleagues noted that in addition to the high Mn/Fe ratio, cellular localisation and trafficking of the metals are also essential for radiation resistance (Santos et al., 2019). They showed that *D. radiodurans* do not encode ferritin-like proteins. Alternatively, two proteins, Dps1 and Dps2, regulate Mn and Fe homeostasis and are associated with the tolerance against ROS in *D. radiodurans.*(Santos et al., 2017).

Deinococcus species also use enzymatic systems for oxidative stress defence. However, those enzymes are widespread in other organisms and do not seem to be sufficient for the extreme radiation resistance in *Deinococcus* species (Shashidhar et al., 2010). Superoxide dismutases, especially MnSOD, peroxidases, thioredoxins, and glutaredoxin-like proteins, are enzymes involved in ROS cleaning in *D. radiodurans* and other species (Figure 1-2) (Lim et al., 2018). Other antioxidant molecules are carotenoids, which are widespread natural pigments and act as ROS scavengers in non-phototrophic bacteria for cellular protection. *D. radiodurans* and many other *Deinococcus* species produce a red-pigmented carotenoid called deinoxanthin, which gives *Deinococcus* its orange-red to pink-coloured colonies (Lemee et al., 1997; Lim et

al., 2018; Zhou et al., 2015). Gene knockout studies have shown that disruption of genes involved in carotenoid biosynthesis makes *D. radiodurans* more sensitive to UV radiation and oxidative stress than the wild-type species *D. radiodurans* R1, indicating that they are important for antioxidant activity (Zhou et al., 2015).

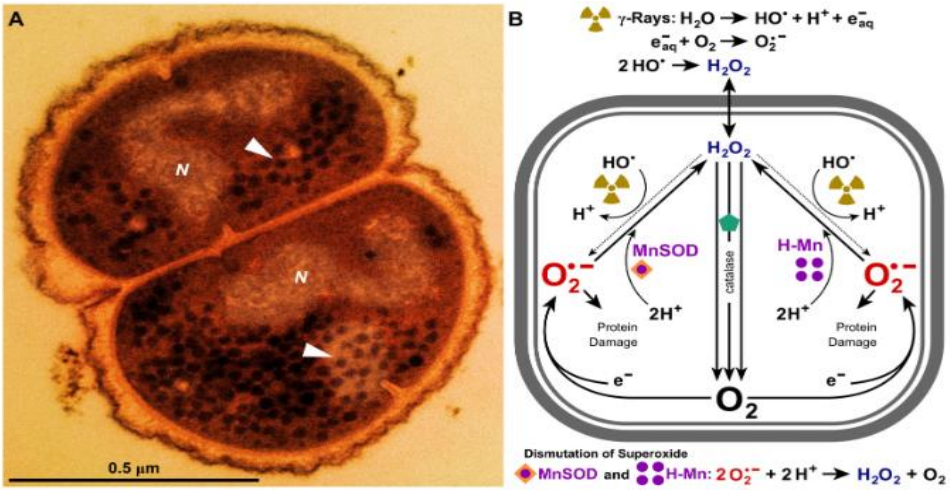

Figure 1-2. *D. radiodurans* antioxidant defence system.

"(A) Electron-dense granules are highly enriched in Mn2+ and phosphate precursors of Mn antioxidants, visualised by transmission electron microscopy (TEM), identified by arrows inside a dividing diplococcus (Daly et al. 2007; Santos et al. 2019). (B) Ionising radiation-driven reactions that generate reactive oxygen species (ROS) and cellular mechanisms related to the dismutation of superoxide (O2•−) by Mn-dependent superoxide dismutase (MnSOD) and Mn antioxidants, illustrating the two complementary catalysts of superoxide defence, MnSOD and Mn antioxidants (H-Mn). Importantly, O2•− damages proteins, but not the DNA " material used (Daly et al., 2004; Daly, 2009).

1.2.1.3 Regulatory pathways

Genome sequencing of *D. radiodurans* (White et al., 1999) provided a platform for gene knockout and transcriptomics experiments. In 2004, Tanaka and colleagues showed that exposure of *D. radiodurans* to ionising radiation and desiccation induces various novel genes, including some with unknown functions in irradiated cells (Tanaka et al., 2004). Most of the upregulated genes encode proteins involved in DNA repair (RecA, RuvB, UvrA, UvrB, UvrD), DNA supercoiling (GyrA and GyrB), and several *Deinococcus*-specific proteins (PprA, DdrA, DdrB, DdrC, DdrD) which are highly expressed in response to DNA damage Table 1-1 (Lim et al., 2018). Later studies showed a

palindromic 17 bp sequence named RDRM (Radiation Desiccation Response Motif) is conserved in the promoter of the regulon that controls these genes called Radiation Desiccation Response (RDR). This palindromic sequence was found in *D.radiodurans* and other *Deinococcus* species, *D. geothermalis* and *D. deserti* (De Groot et al., 2009; Ludanyi et al., 2014, 2014; Makarova et al., 2007).

Two regulatory proteins, DdrO and IrrE metalloprotease, control this promoter. DdrO is the repressor of the RDR regulon by binding to the RDRM sequence on the DNA. After exposure to ionising radiation or prolonged desiccation, IrrE cleaves the DdrO repressor and enables the expression of the genes within the RDR regulon and, ultimately, the repair of damaged DNA (Devigne et al., 2015). Recent studies have shown that the RDR regulon in *D. radiodurans* is more complex than previously thought and is composed of at least 35 genes, including genes encoding DNA and RNA metabolism proteins, such as RecG and HelD helicases, and the prokaryotic replicative DNA ligase LigA, and new genes with role in diverse metabolic pathways such as translation or encoding of proteins of unknown function (Narasimha & Basu, 2021).

1.2.2 Cultivation of *Deinococcus*

Deinococcus species have been isolated from diverse environments, including hot deserts, Antarctica, high atmosphere, hot springs, freshwater, radiation-polluted sites, air, animal guts, and many metal surfaces (Table 3-1) (Chen et al., 2011; de Groot et al., 2005; Ferreira et al., 1997; Hirsch et al., 2004; Kim et al., 2015; Lee et al., 2016; Yang et al., 2009). When writing this thesis in December 2023, 92 *Deinococcus* species have been named, from which 89 are validly published according to the List of Prokaryotic Names with Standing in Nomenclature (LPSN) (https://www.bacterio.net/genus/*Deinococcus*). Additionally, over 300 16S rRNA sequences are deposited in NCBI taxonomy designated as *Deinococcus* sp. as putative *Deinococcus* species that requires further identification and genome sequencing https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=1298). However, according to the Genome Taxonomy Database (GTDB), the *Deinococcaceae* family is categorised into seven genera, including *Deinobacterium*, *Deinococcus*,

Deinococcus_A, Deinococcus_B, and *Deinococcus_C* represented by isolates and two uncultivated genera JACMOA01 and JAJZIR01 (Parks et al., 2018).

Many of the *Deinococcus* species have been isolated from deserts and arid environments. A single study by Rainey and colleagues reported cultivating more than 130 species from the arid soil of the Sonoran Desert and non-arid forest soil, resulting in the isolation of 60 *Deinococcus* strains, nine of which were novel species with D_{10} values higher than 10 kGy. They used various doses of ionising radiation as pre-culture treatment and reported that some isolates survived 30 kGy from the Sonoran Desert soil, but no isolates were recovered from the non-arid forest soil after exposure to doses greater than 13 kGy (Rainey et al., 2005).

In another study, Mohseni and Abbaszadeh isolated two *Deinococcus* species from the Lut desert in Iran, a hyper-arid desert with a five-year record as the hottest place on Earth (Mohseni et al., 2014). Like the previous study, they used 10 kGy of gamma radiation as a pre-culture treatment to eradicate sensitive microorganisms and enrich radioresistant isolates. They reported a D_{10} value of 7.15 and 5.94 kGy for two *Deinococcus* isolates (Mohseni et al., 2014). Many *Deinococcus* species are isolated from deserts using the same method with gamma radiation. Highly radioresistant *D. deserti,* isolated from the Sahara desert, was enriched by 15 kGy of gamma radiation to soil pre-cultivation (de Groot et al., 2005). *D. peraridilitoris* (Rainey et al., 2007), *D. xinjiangensis* (Peng et al., 2009), and *D. taklimakanensis* (Liu et al., 2017) are some other *Deinococcus* species isolated from deserts and mostly show high levels of radiation resistance.

Conversely, most *Deinococcus* species isolated from cold environments are sensitive to gamma radiation and form a phylogenetic cluster with each other (Table 3-1, Figure 3-1). Those species have been isolated from Antarctica, the Arctic, high-altitude and alpine environments, and freshwater sediments. *D. persicinus* was isolated from soil near a water stream in Seoul and has a D_{10} value of 1.5 kGy, which is ten times more sensitive than *D. radiodurans* (Jeon et al., 2016). *D. arcticus* was isolated from the rhizosphere soil of the Arctic tundra and showed a D_{10} value of 3 kGy (X. P. Wang et

al., 2019). *D. detaillensis*, with a D_{10} value of 4.8, was isolated from Detaille Island in Antarctica (Zhang et al., 2020a). These *Deinococcus* species show a pattern of declined resistance when evolving to cold environments.

In a screening study of culturable bacterial diversity in alpine regions, Callegan et al. described four novel psychrophilic and radiation-sensitive *Deinococcus* species from Pico de Orizaba, Mexico, and Mount Evans, CO, USA. *D. radiomollis, D. claudionis, D. altitudinis,* and *D. alpinitundrae* are reported species and show similar phenotype despite the geographical distance (Callegan et al., 2008). *D. rubrus* is another radiation-sensitive species isolated from Antarctica with a D_{10} value of 3kGy. *D. sedimenti, D.seoulensis, D. persicinus,* and *D. aquaticus* are other examples of sensitive *Deinococcus* species mostly isolated from sediments and soil samples collected near or in freshwater (Im et al., 2008; Jeon et al., 2016; Lee et al., 2016). Another study reported three novel species from Antarctica: *D. frigens, D. saxicola*, and *D. marmoris*, but unfortunately, did not report gamma radiation resistance and the dose of UV radiation resistance (Hirsch et al., 2004). Several other studies lack data on radiation resistance levels, complicating studying the ecology and evolution of radiation resistance in *Deinococcus* as a genus.

1.2.3 The origin of adaptation to radiation

On the planet Earth, natural ionising radiation only occurs in very small doses and is referred to as background radiation (Asgarani et al., 2012). The average worldwide background radiation dose is 2.4 millisievert (mSv) per year (1 Siv = 1 Gy), but it can be more in some areas. For instance, Ramsar, a city in northern Iran, recorded the highest measured background radiation on Earth, at 260 mSv per year (Ghiassi-nejad et al., 2002). However, no significant difference has been reported between people in the high background areas compared to people in the normal background areas (Monfared et al., 2005). The epicentre of those elevated radioactive environments is a region with few hot springs which exert Radon gas from underneath minerals. One study reported the isolation of a radiation-resistant isolate, *Kocuria* sp. ASB 107 with a D_{10} value of 2 kGy (Asgarani et al., 2012).

Geothermal pools are also known to be a source of DNA damage, and multiple *Deinococcus* species have been isolated from different hot springs (Battista, 2016). Moreover, prokaryotes are known to be much more resilient than eukaryotes due to their smaller genome size and less complicated cellular structure (Daly et al., 2010). For instance, the D_{10} value of *E. coli* is about 700 Gy, and *Shewanella oneidensis*, which is one of the most sensitive microorganisms, has a D_{10} of 70 Gy (Ghosal et al., 2005; Shashidhar et al., 2010). In comparison, human Jurkat T cells can tolerate only very low doses, 10-15 Gy of Gamma radiation (Daly et al., 2010). Therefore, naturally-elevated background radiation per se cannot be accounted as an evolutionary pressure for the evolution of radiation resistance phenotype as a strategy for survival.

On the other hand, the Earth's surface has never been exposed to high fluxes of ionising radiation. Therefore, it was suggested that the radiation-resistance phenotype could not be an evolutionary adaptation process since there is no selective advantage to this trait in the natural world (Mattimore & Battista, 1996). Several theories have been proposed to solve the problem of evolution to high doses of ionising radiation. Some of those theories, like panspermia, suggesting that life has extraterrestrial origins and that *Deinococcus* is a candidate that travelled to Earth on meteorites and originated life on Earth (Pavlov et al., 2006), seem unrealistic.

It is believed that the radiation resistance trait is the byproduct of adapting to desiccation or heat-induced cellular damage (Battista, 2015). This theory is in agreement with the observation in the majority of cultivation studies where ionising radiation-resistant species are also resistant to desiccation and/or elevated temperatures (Albuquerque et al., 2005; Ekman et al., 2011; Horne et al., 2022, 2022; Mattimore & Battista, 1996; Peng et al., 2009). Desiccation and heat do not introduce the same types of damage caused by ionising radiation-induced ROS but can cause DNA double-strand breaks (DSBs), the lesions most commonly associated with ionising radiation-induced lethality (Battista, 2015).

1.2.4 Comparative Genomics

The whole genome sequence of *D. radiodurans* (White et al., 1999) revolutionised our understanding of radiation resistance. The comparative genomic race started when *D. radiodurans* was sequenced to find an answer for the radiation resistance phenomenon. The computational assembly of the four genome partitions of *D. radiodurans* R1 was stalled due to the large number of repeated sequences dispersed across its chromosomes and plasmids (Daly, 2023; Makarova et al., 1999).

One year later, Makarova and colleagues conducted another comparative analysis on *D. radiodurans* against *E. coli* (Makarova et al., 2001) and later transcriptome analysis of strain R1 recovering from 17.5 kGy (Liu et al. 2003). The comparative analyses revealed that DNA repair genes (e.g., recA) as well as antioxidant genes like sodA and katA in *D. radiodurans* are unremarkable, and the *D. radiodurans* genome encodes approximately the same number and types of genes as sensitive bacteria (Makarova et al., 2001). In 2005, Omelchenko conducted the first comparative genomic analysis on *D. radiodurans* and its close relative, *Thermus thermophilus.* This study also could not find a decisive answer to the radiation resistance conundrum (Omelchenko et al., 2005)

The failure of comparative genomics studies to find a universal mechanism for radiation and desiccation resistance by comparing *D. radiodurans* and distant species prompted further comparative genomic studies between members of the *Deinococcaceae* family in an attempt to identify genes shared between *Deinococcus* species but absent in other bacteria. In 2007, Makarova and colleagues compared two highly resistant *Deinococcus* species, the mesophilic *D. radiodurans* R1 and the thermophilic *D. geothermalis* and concluded that most of the *Deinococcus*-specific genes initially considered to have a role in the extreme radiation resistance phenotype of *D. radiodurans* were absent in *D. geothermalis* (Makarova et al., 2007). This study also could not find a definitive answer and concluded that comparative genomics could not find a specific set of genes that define radiation resistance phenotype in all *Deinococcus* species.

Later comparative genomics studies were more review studies that compared radiation resistance mechanisms known from *D. radiodurans* to other *Deinococcus* genomes (de

la Tour et al., 2013; Lim et al., 2018; Makarova & Daly, 2014). One of those studies included 11 radiation-resistant *Deinococcus* species and compared mechanisms involved in radiation resistance, oxidative stress defence, DNA repair, and their regulation (Lim et al., 2018), but there was not any novel finding about already known mechanisms. Recently, Jiao and colleagues conducted a comprehensive comparative genomic study on *Thermaceae* and found a genomic potential for incomplete denitrification in *Thermus.* They demonstrated the evolutionary history of the denitrification pathway in the *Thermaceae* and concluded that *Thermus* is an important heterotrophic denitrifier in geothermal environments (Jiao et al., 2022). Such studies have never been conducted on the *Deinococcus* genus to explore its metabolic capabilities.

1.3 Challenges in our understanding of *Deinococcaceae*

1.3.1 Ubiquity of *Deinococcus* species

Deinococcus species are usually ubiquitous but have a lower abundance than other bacteria occupying the same ecological niche. It has been postulated that other bacteria usually overgrow *Deinococcus* members in nature because they grow faster in the standard laboratory media than *Deinococcus* (Krisko & Radman, 2013). In other words, *Deinococcus* made a metabolic investment in the efficiency of survival, whereas other bacteria are devoted to the efficiency of growth. The chemoorganotrophic lifestyle of *D. radiodurans* suggests that this bacterium is a scavenger of organic substrates from other organisms. Besides, its mosaic genome composition and remarkable genomic variability indicate a pre-adaptation to diverse environments among the *Deinococcus* species, hence, the ubiquitous presence of similar variants in diverse habitats but with low abundance (Krisko & Radman, 2013; Makarova et al., 2007; Slade & Radman, 2011; White et al., 1999). This chemoorganotrophic nature of *Deinococcus* is another challenge to its members' cultivation and isolation because other fast-growing organisms usually overgrow them.

1.3.2 Biased cultivation of *Deinococcus* species

The focus on the extreme radiation resistance phenotype in some *Deinococcus* species led to the prevailing view that radiation resistance is the defining characteristic of the *Deinococcaceae* family (Daly, 2023). Consequently, microbiologists have frequently used gamma radiation as a pre-culture treatment to enrich *Deinococcus* species from natural samples (de Groot et al., 2005; Mohseni et al., 2014; Rainey et al., 2005). On the one hand, this method is helpful because it enriches radiation-resistant isolates and eradicates other microorganisms, hence facilitating the cultivation of radioresistant *Deinococcus* members. However, the biggest drawback of this approach is that only highly resistant *Deinococcus* species are enriched, and most radiation-sensitive members of *Deinococcaceae* are eradicated before cultivation.

The ramification of this enrichment step is that radiation-sensitive *Deinococcus* species are underrepresented and have gained less attention. However, some studies did not use gamma radiation as a pre-culture treatment method and accidentally isolated radiosensitive *Deinococcus* species. Interestingly, the majority of those sensitive species were isolated from cold environments and without pre-culture irradiation exposure or low levels of gamma radiation (Callegan et al., 2008; E. B. Kim et al., 2017; Lee et al., 2016; X. P. Wang et al., 2019). In one study, four psychrophilic *Deinococcus* species were isolated from alpine environments without gamma-radiation treatment and showed D_{10} values ranging from 2.2 to 3.8 (Callegan et al., 2008), which is significantly more sensitive than *D. radiodurans* with D_{10} value of 15 kGy. The presence of some radiation-sensitive species within the *Deinococcus* genus suggests that the current approach can cause a significant bias in our understanding of the ecology and evolution of all members of this genus and enrich radiation-resistant isolates.

Another challenge that emerges from the studies of the cultivation of *Deinococcus* species from environmental samples is the lack of metadata. Many isolation and cultivation studies only briefly present the isolation source using general terms like "air", "soil", and "water" without providing any useful ecological data on the geochemical properties of their isolation habitats. The lack of metadata is also abundant

in the radiation resistance levels of isolates. For instance, *D. frigens* and *D. saxicola* were isolated from Antarctica (Hirsch et al., 2004), but the radiation resistance levels were not reported. Our review of the literature shows that from 89 validly published *Deinococcus* species based on the LPSN and NCBI taxonomy, at least 25 species lack radiation resistance data, some of which are metagenome-assembled genomes (MAGs).

1.3.3 Prediction of radiation resistance

As discussed in the literature review, radiation resistance in *D. radiodurans* is attributed to three main mechanisms: efficient DNA repair enzymes and polyploidy, ROS scavenging mechanisms, and regulatory pathways (Lim et al., 2018; Slade & Radman, 2011). An evaluation of Mn^{2+} concentration and carotenoid content in twelve *Deinococcus* strains belonging to seven species with diverse radiation resistance levels did not show any significant correlation between intracellular Mn/Fe ratios and the D_{10} values of resistant *Deinococcus* species (Shashidhar et al., 2010). This study concludes that the effective antioxidant system, which confers radiation resistance properties to *Deinococcus,* cannot be attributed to a single factor (Shashidhar et al., 2010). Therefore, it is logical to think that HGT can drive the difference in radiation resistance.

The biggest challenge in understanding the radiation resistance phenotype in the *Deinococcus* genus is the inconsistency in radiation resistance capacity among its isolates. As demonstrated in (Table 3-1), *Deinococcus* species show an extensive range of radiation resistance capability, which can be in the order of magnitude more sensitive than the type species *D. radiodurans* in some species like *D. persicinus.* Surprisingly, this diversity in the radiation resistance capacity has also been observed between closely related strains (Shashidhar et al., 2010). For instance, *D. deserti* showed a D_{10} value of 6.5 kGy for VCD117, compared to VCD115T with a D_{10} value of 8 kGy. This diversity in radiation resistance among *Deinococcus* species and strains indicates the complexity of this phenotype.

This observation prompted the idea that radiation resistance can not be predicted using genome sequence (Daly, 2023; Sharma et al., 2017; Shuryak et al., 2019). However, radiation resistance is an inheritable trait and has a genomic basis. The majority of

mechanisms we know about radiation resistance come from the model organism *D. radiodurans* and a few other studied *Deinococcus* species like *D. geothermalis, D. peraridilitoris,* and *D. deserti* (Baudet et al., 2010; Bornot et al., 2015; Daly et al., 2004, 2010; Gaidamakova et al., 2022; Jin et al., 2019; Sharma et al., 2017). It appears that the focus of microbiologists has always been on discovering a universal set of genes with a role in radiation resistance to predict this phenotype. Hence, this approach concluded that radiation resistance is not predictable with genomic data, and the answer lies in antioxidant metabolites (Daly, 2023).

It is difficult to predict radiation resistance due to three main challenges. First, the known and sequenced *Deinococcus* species have an unproportionate diversity regarding radiation resistance. Second, there is a lack of statistical studies exploring potential radiation resistance patterns using all available genomes, radiation resistance levels and even a more extended number of genomes. Third, current comparative genomics studies are limited and rely on *D. radiodurans* as a model organism.

1.3.4 Comparative genomics

The literature review demonstrated that a few comparative genomic studies, mainly in the early years of the genome sequencing era, tried to decode radiation resistance through genomic contents of *Deinococcus* species. Nevertheless, those attempts failed, and later studies promoted the idea that "despite concerted functional genomic efforts, the level of cellular IR resistance cannot be predicted by a genome sequence" (Sharma et al., 2017). Recently, Michael J Daly concluded that while the number of *Deinococcus* genomes is now sufficiently large enough to determine the core genome and pangenome of the *Deinococcaceae* family, extreme resistance persistently remains a phenotype without a genotype (Daly, 2023).

1.4 The gap in the knowledge

Despite the widespread research with a focus on the radiation resistance mechanisms of *D. radiodurans* and a few highly resistant *Deinococcus* species over the last six decades, many aspects of this genus, including their ecology and evolution, have remained vastly

unknown. Moreover, the focus on the radiation resistance phenotype overshadowed the understanding of many other aspects of this distinct taxonomic group. We discuss this gap in knowledge in three themes.

1.4.1 Conventional wisdom biases our knowledge.

The central bias on the *Deinococcaceae* family arises from the conventional wisdom that the radiation resistance trait is the defining characteristic of the *Deinococcaceae* family (Daly, 2023). The popularity of this notion among microbiologists led to the widespread usage of gamma radiation as a pre-culture treatment before cultivation to isolate *Deinococcus* species. This approach has most likely enriched radiation-resistant strains, as all radiosensitive strains were eradicated during the pre-culture treatment. Isolation of radiosensitive strains such as *D. persicinus* with a D_{10} value of 1.5 kGy (Jeon et al., 2016), which is ten-fold more sensitive than *D. geothermalis*, highlights the possibility of isolation of even more sensitive strains. As a result, there is no data on the true diversity of *Deinococcaceae* members. This view has shaped the isolated and cultivated members of the *Deinococcaceae* and, consequently, our understanding of the ecology and evolution of this family. The literature suggests a lack of comprehensive comparative genomics on the *Deinococcus* genus to address its diversity and demonstrate if the pangenome of *Deinococcus* encompasses the genomic diversity of its genus.

1.4.2 Prediction of radiation resistance

The failure of comparative genomics in finding a genomic basis promoted the idea that genomic data cannot predict radiation resistance (Sharma et al., 2017), and extreme resistance persistently remains a phenotype without a genotype (Daly, 2023). Only one study on the prediction of radiation resistance was published in the Knowledge Discovery and Data Mining Conference (Aridhi et al., 2013). This study used sequences of conserved DNA repair proteins to build a model for predicting radiation resistance but did not provide practical data.

Our literature review demonstrated that no comprehensive comparative genomic study has yet included currently available *Deinococcus* genomes to construct the pangenome of *Deinococcus*. Such a study should be unbiased by the radiation resistance trait and should consider the potential existence of radiation-sensitive species with high adaptability. Furthermore, the claim that radiation resistance is a phenotype with no genotype has been

reiterated from the perspective of *D. radiodurans*, and no statistical analysis explores all available genomes. We provide a statistically informed study to assess the predictability of radiation resistance in *Deinococcus* using their genomic content. However, it must be noted that the currently available genomes of *Deinococcus* species are highly biased and probably not representative and large enough to capture the whole pangenome repertoire of this genus.

1.4.3 Evolution of *Deinococcaceae*

We found a gap in knowledge about the evolutionary history of the *Deinococcota* phylum, especially the events that gave rise to the *Deinococcaceae* family. In the early years of the 21st century, when the genomes of *D. radiodurans* and *T. thermophilus* were sequenced, the first comparative genomic analysis compared their genomes and their divergent adaptation to radiation resistance and thermophily (Omelchenko et al., 2005). They compared evolutionary events, such as gene gains and losses and concluded that despite the abundance of horizontally transferred genes, megaplasmids of *Thermus* and *Deinococcus* originated from a common ancestor. Later, *D. geothermalis* was added to the list of compared genomes (*D. radiodurans* and *T. thermophilus*) and demonstrated that the extreme stress resistance in the *Deinococcus* lineage emerged from amassing cell-cleaning mechanisms, but not by the acquisition of novel DNA repair systems. These comparative genomic studies were highly informative about the evolution of *Deinococcus* and *Thermus,* but their primary objective was finding mechanisms for radiation resistance.

In those early years, limitations on the availability of whole genome sequences did not allow for comprehensive comparative genomics. Later, one comparative genomic study compared 11 *Deinococcus* genomes with a focus on known genes in radiation resistance (Lim et al., 2018), but there was no comprehensive study on the evolution of *Deinococcota* until 2022, when a comparative genomic analysis used 23 *Thermus* genomes and provided insights into the evolutionary history of an incomplete denitrification pathway (Jiao et al., 2022). However, the evolutionary history of *Deinococcaceae* and its evolution within the *Deinococcota* phylum has remained entirely unexplored.

With the availability of more genomes and more advanced methods, we can now gain deeper insight into the evolutionary history of the *Deinococcus* genus. In this thesis, I attempt to address these gaps in knowledge about the ecology and evolution of the *Deinococcaceae* family

using a comprehensive comparative genomic analysis of all available genomes with three different approaches. These comparative genomic analyses comprise three chapters of this thesis: chapters two, three, and four. The details of these three approaches are presented in the next section, thesis structure.

1.5 Thesis structure

1.5.1 Chapter I: Introduction

In this chapter, I present a background of the research in the field and present current knowledge about the *Deinococcaceae* family.

1.5.2 Chapter II: Pangenome analysis

This chapter aims to assess the diversity of the genomic repertoire of all available *Deinococcus* genomes. I use comparative genomics to answer the following question. What factors contribute to the widespread distribution and diverse characteristics of the *Deinococcus* genus? I hypothesise that "the ubiquity of *Deinococcus* is due to its genetic flexibility and large pangenome".

To answer the question and test the hypothesis, I will look at the pangenome of the *Deinococcus* genus using representative genomes of all available species. I also add two newly sequenced genomes to diversify our genomic dataset. Pangenome here refers to all gene families found in the *Deinococcus* genus. Some prokaryotic species have open pangenomes, meaning they have an extensive collection of genes, while others have closed pangenomes, manifesting very few gene content differences (McInerney et al., 2020).

The pangenome of a species depends on the genomes used and reflects the diversity of the studied genomes. As discussed, the diversity of the *Deinococcus* species is vastly biased due to cultivation methods. Moreover, only a few *Deinococcus* species have multiple strains, and the genome sequence is usually available for one species. Therefore, we conduct this pangenome analysis at the genus level using high-quality genomes available in the NCBI database and add two new genomes belonging to

Deinococcus altitudinis ME-04-01-32 and *Deinococcus radiomollis* PO-04-20-132 to this collection to increase the diversity of radiation-sensitive genomes.

The pangenome is classified into three groups: core genome, accessory genome, and unique genes. The core genome consists of genes shared between all genomes, the accessory genome is defined as genes not present in all but more than one genome, and unique genes are found only in one genome. We calculated the impact of adding new genomes to the analysis and estimated the openness of the pangenome based on a model proposed by Tettelin (Tettelin et al., 2008).

Ultimately, we looked into the functional potential of core, accessory and unique genes to gain deeper insight into the ecology of the *Deinococcus* genus. Functional groups were assigned to the sequences in each pangenome group, and their abundance was calculated. Finally, we sought to explore the metabolic diversities of *Deinococcus* genomes by profiling the functional capabilities of each genome and comparing genomes isolated in different habitats. KEGG pathways profiling was used to gain insight into the functional ecology of *Deinococcus* species in their habitats.

1.5.3 Chapter III: Radiation Resistance Pattern

In Chapter III, I aimed to evaluate the conventional wisdom that "radiation resistance is the defining characteristic of *Deinococcaceae"* and that "*radiation resistance cannot be predicted using genomic data and remained a phenotype without genotype* " (Daly, 2023). I hypothesise that "a large diversity of genomic samples in the *Deinococcus* genus can reveal a correlation between radiation resistance phenotype and the associated genotype, which provides a genomic basis to predict this phenotype".

This chapter uses statistical methods to explore gene presence-absence patterns among *Deinococcus* species with different radiation resistance levels. From 91 validly published *Deinococcus* species according to LPSN database, only 64 species have their genomes available in the NCBI database, and the others have not been sequenced. Moreover, only 42 out of 64 *Deinococcus* genomes in the public database have data on their D_{10} values in the literature. The two newly sequenced *Deinococcus* species are

added to this collection, and a total of 44 genomes from the *Deinococcus* species based on the NCBI taxonomy are compared.

We compiled a dataset that included the genome sequence of 44 species with radiation resistance data and metadata from GTDB (for genomic features) and literature (for D10 values and isolation habitats), genomic features, and isolated habitats of selected species. *Deinococcus* species were divided into three categories based on their resistance levels. Species with D_{10} values between 1 kGy and 5 kGy were called low-resistant (LR), species with D_{10} values between 5 kGy and 10 kGy moderately resistant (MR), and species with D_{10} values more than 10 kGy were categorised as highly resistant (HR)

We constructed a phylogenomic tree of single-copy orthogroups shared between 95% of genomes. A pattern was observed in the topology of the species tree where radiation-sensitive species belonging to (LR) category formed two distinct clades. Genomes belonging to the two sensitive clades were separately compared to MR and HR groups by filtering one sensitive clade each time. Pearson's chi-squared test was used to calculate the correlation between D_{10} values and gene presence or absence for 17,350 gene families. Benjamini–Hochberg procedure was used to correct the p-values for significantly correlated orthogroups.

1.5.4 Chapter IV Evolutionary history of *Deinococcota*

Chapter IV aimed to reconstruct the evolutionary history of the *Deinococcota* phylum to shed light on the evolution and emergence of the *Deinococcaceae* family. We conducted a phylogenomic analysis using available high-quality genomes within the *Deinococcota* phylum, including *Deinococcaceae*, *Trueperaceae*, *Thermaceae*, and *Marinithermaceae* families, to detect evolutionary events that shaped the two main families *Deinococcaceae* and *Thermaceae*. We compare the genomic characteristics of these two families and look for overall genomic differences that shaped their evolution.

We use the gene tree-species tree reconciliation method to infer the evolutionary history of the *Deinococcota* phylum and reconstruct the common ancestor of the

Deinococcaceae family. The evolutionary events, including gene duplication transfer and loss, are used to gain insights into the function of the genes that shaped the *Deinococcaceae* family.

1.5.5 Chapter V Conclusion

The final chapter of this thesis highlights general conclusions and presents future directions from this work.

1.6 References:

Albuquerque, L., Simões, C., Nobre, M. F., Pino, N. M., Battista, J. R., Silva, M. T., Rainey, F. A., & de Costa, M. S. (2005). Truepera radiovictrix gen. Nov., sp. Nov., a new radiation resistant species and the proposal of Trueperaceae fam. Nov. *FEMS Microbiology Letters*, *247*(2), 161–169. https://doi.org/10.1016/j.femsle.2005.05.002

Anderson, A., Nordon, H., Cain, R. F., Parrish, G., Duggan, D., Nordan, H., Parish, G., & Cullum-Dugan, D. (1956). Studies on a radioresistant micrococcus. I. Isolation, morphology, cultural characteristics, and resistance to gamma radiation. *Food Technology*, *10*, 575–578.

Aridhi, S., Maddouri, M., Sghaier, H., & Nguifo, E. M. (2013). Computational phenotype prediction of ionising-radiation-resistant bacteria with a multiple-instance learning model. *Proceedings of the 12th International Workshop on Data Mining in Bioinformatics - BioKDD '13*, 18–24. https://doi.org/10.1145/2500863.2500866

Baek, K., Chung, E. J., Choi, G.-G., Kim, M.-K., Lim, S., & Choi, A. (2018). *Deinococcus koreensis* sp. Nov., a gamma radiation-resistant bacterium isolated from river water. *International Journal of Systematic and Evolutionary Microbiology*, *68*(8), 2545–2550. https://doi.org/10.1099/ijsem.0.002872

Barnese, K., Gralla, E. B., Cabelli, D. E., & Selverstone Valentine, J. (2008). Manganous Phosphate Acts as a Superoxide Dismutase. *Journal of the American Chemical Society*, *130*(14), 4604–4606. https://doi.org/10.1021/ja710162n

Battista, J. R. (2015). 6 The origin of extreme ionising radiation resistance. In *6 The origin of extreme ionising radiation resistance* (pp. 111–126). De Gruyter. https://doi.org/10.1515/9783110340716-008

Battista, J. R. (2016). *Deinococcus* – Thermus Group. In *Encyclopedia of Life Sciences* (pp. 1–12). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470015902.a0021151

Baudet, M., Ortet, P., Gaillard, J.-C., Fernandez, B., Guérin, P., Enjalbal, C., Subra, G., de Groot, A., Barakat, M., Dedieu, A., & Armengaud, J. (2010). Proteomics-based Refinement of *Deinococcus deserti* Genome Annotation Reveals an Unwonted Use of Non-canonical Translation Initiation Codons*. *Molecular & Cellular Proteomics*, *9*(2), 415–426. https://doi.org/10.1074/mcp.M900359-MCP200

Bentchikou, E., Servant, P., Coste, G., & Sommer, S. (2010). A Major Role of the RecFOR Pathway in DNA Double-Strand-Break Repair through ESDSA in *Deinococcus radiodurans*. *PLOS Genetics*, *6*(1), e1000774. https://doi.org/10.1371/journal.pgen.1000774

Bornot, J., Molina-Jouve, C., Uribelarrea, J.-L., & Gorret, N. (2015). Quantitative Characterisation of the Growth of *Deinococcus geothermalis* DSM-11302: Effect of Inoculum Size, Growth Medium and Culture Conditions. *Microorganisms*, *3*(3), 441–463. https://doi.org/10.3390/microorganisms3030441

Brooks, B. W., & Murray, R. G. E. (1981). Nomenclature for 'Micrococcus *radiodurans*' and other radiation-resistant cocci: *Deinococcaceae* fam. Nov. And *Deinococcus* gen. Nov., including five species. *International Journal of Systematic Bacteriology*, *31*(3), 353–360. https://doi.org/10.1099/00207713-31-3-353

Buljubašić, M., Hlevnjak, A., Repar, J., Đermić, D., Filić, V., Weber, I., Zahradka, K., & Zahradka, D. (2019). RecBCD- RecFOR-independent pathway of homologous recombination in *Escherichia coli*. *DNA Repair*, *83*, 102670. https://doi.org/10.1016/j.dnarep.2019.102670

Callegan, R. P., Noble, M. F., McTernan, P. M., Battista, J. R., Navarro-González, R., McKay, C. P., da Costa, M. S., & Rainey, F. A. (2008). Description of four novel psychrophilic, ionising radiation-sensitive *Deinococcus* species from alpine environments. *International Journal of Systematic and Evolutionary Microbiology*, *58*(5), 1252–1258. https://doi.org/10.1099/ijs.0.65405-0

Cha, S., Srinivasan, S., Seo, T., & Kim, M. K. (2014a). *Deinococcus radiotolerans* sp. Nov., a gamma-radiation-resistant bacterium isolated from gamma ray-irradiated soil. *Antonie van Leeuwenhoek*, *105*(1), 229–235. https://doi.org/10.1007/s10482-013-0069-0

Cha, S., Srinivasan, S., Seo, T., & Kim, M. K. (2014b). *Deinococcus soli* sp. Nov., a gamma-radiation-resistant bacterium isolated from rice field soil. *Current Microbiology*, *68*(6), 777–783. https://doi.org/10.1007/s00284-014-0542-7

Chen, W., Wang, B., Hong, H., Yang, H., & Liu, S. J. (2011). *Deinococcus reticulitermitis* sp. Nov., isolated from a termite gut. *International Journal of Systematic and Evolutionary Microbiology*, *62*(1), 78–83. https://doi.org/10.1099/ijs.0.026567-0

Chevance, F. F. V., & Hughes, K. T. (2008). Coordinating assembly of a bacterial macromolecular machine. *Nature Reviews Microbiology*, *6*(6), Article 6. https://doi.org/10.1038/nrmicro1887

Cox, M. M., & Battista, J. R. (2005). *Deinococcus radiodurans*—The consummate survivor. *Nature Reviews Microbiology*, *3*(11), 882–892. https://doi.org/10.1038/nrmicro1264

Daly, M. J. (2000). Engineering radiation-resistant bacteria for environmental biotechnology. *Current Opinion in Biotechnology*, *11*(3), 280–285. https://doi.org/10.1016/s0958-1669(00)00096-3

Daly, M. J. (2009). A new perspective on radiation resistance based on *Deinococcus radiodurans*. *Nature Reviews Microbiology*, *7*(3), 237–245. https://doi.org/10.1038/nrmicro2073

Daly, M. J. (2023). The scientific revolution that unraveled the astonishing DNA repair capacity of the *Deinococcaceae*: 40 years on. *Canadian Journal of Microbiology*. https://doi.org/10.1139/cjm-2023-0059

Daly, M. J., Gaidamakova, E. K., Matrosova, V. Y., Kiang, J. G., Fukumoto, R., Lee, D. Y., Wehr, N. B., Viteri, G. A., Berlett, B. S., & Levine, R. L. (2010). Small-molecule antioxidant proteome-shields in *Deinococcus radiodurans*. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0012570

Daly, M. J., Gaidamakova, E. K., Matrosova, V. Y., Vasilenko, A., Zhai, M., Venkateswaran, A., Hess, M., Omelchenko, M. V., Kostandarithes, H. M., Makarova, K. S., Wackett, L. P., Fredrickson, J. K., & Ghosal, D. (2004). Accumulation of Mn(II) in *Deinococcus radiodurans* Facilitates Gamma-Radiation Resistance. *Science*, *306*(5698), 1025–1028. https://doi.org/10.1126/science.1103185

Daly, M. J., Ouyang, L., Fuchs, P., & Minton, K. W. (1994). In vivo damage and recA-dependent repair of plasmid and chromosomal DNA in the radiation-resistant bacterium *Deinococcus radiodurans*. *Journal of Bacteriology*, *176*(12), 3508–3517. https://doi.org/10.1128/jb.176.12.3508-3517.1994

de Groot, A., Chapon, V., Servant, P., Christen, R., Fischer-Le Saux, M., Sommer, S., & Heulin, T. (2005). *Deinococcus deserti* sp. Nov., a gamma-radiation-tolerant bacterium isolated from the Sahara Desert. *International Journal of Systematic and Evolutionary Microbiology*, *55*(6), 2441–2446. https://doi.org/10.1099/ijs.0.63717-0

De Groot, A., Dulermo, R., Ortet, P., Blanchard, L., Guérin, P., Fernandez, B., Vacherie, B., Dossat, C., Jolivet, E., Siguier, P., Chandler, M., Barakat, M., Dedieu, A., Barbe, V., Heulin, T., Sommer, S., Achouak, W., & Armengaud, J. (2009). Alliance of Proteomics and Genomics to Unravel the Specificities of Sahara Bacterium *Deinococcus deserti*. *PLoS Genetics*, *5*(3), e1000434. https://doi.org/10.1371/journal.pgen.1000434

de la Tour, C. B., Passot, F. M., Toueille, M., Mirabella, B., Guérin, P., Blanchard, L., Servant, P., de Groot, A., Sommer, S., & Armengaud, J. (2013). Comparative proteomics reveals key proteins recruited at the nucleoid of *Deinococcus* after irradiation-induced DNA damage. *Proteomics*, *13*(23–24), 3457–3469. https://doi.org/10.1002/pmic.201300249

De Maayer, P., Pillay, T., & Coutinho, T. A. (2020). Comparative genomic analysis of the secondary flagellar (flag-2) system in the order Enterobacterales. *BMC Genomics*, *21*(1), 100. https://doi.org/10.1186/s12864-020-6529-9

Devigne, A., Ithurbide, S., Bouthier de la Tour, C., Passot, F., Mathieu, M., Sommer, S., & Servant, P. (2015). DdrO is an essential protein that regulates the radiation desiccation response and the apoptotic-like cell death in the radioresistant *Deinococcus radiodurans* bacterium. *Molecular Microbiology*, *96*(5), 1069–1084. https://doi.org/10.1111/mmi.12991

Dong, N., Li, H.-R., Yuan, M., Zhang, X.-H., & Yu, Y. (2015). *Deinococcus* antarcticus sp. Nov., isolated from soil. *INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY*, *65*(Pt 2), 331–335. https://doi.org/10.1099/ijs.0.066324-0

Ekman, J. V., Raulio, M., Busse, H.-J., Fewer, D. P., & Salkinoja-Salonen, M. (2011). Deinobacterium chartae gen. Nov., sp. Nov., an extremely radiation-resistant, biofilm-forming bacterium isolated from a Finnish paper mill. *International Journal of Systematic and Evolutionary Microbiology*, *61*(3), 540–548. https://doi.org/10.1099/ijs.0.017970-0

Ferreira, A. C., Nobre, M. F., Rainey, F. A., Silva, M. T., Wait, R., Burghardt, J., Chung, A. P., & da Costa, M. S. (1997). *Deinococcus geothermalis* sp. Nov. And *Deinococcus* murrayi sp. Nov., two extremely radiation-resistant and slightly thermophilic species from hot springs. *International Journal of Systematic Bacteriology*, *47*(4), 939–947. https://doi.org/10.1099/00207713-47-4-939

Fredrickson, J. K., Li, S. W., Gaidamakova, E. K., Matrosova, V. Y., Zhai, M., Sulloway, H. M., Scholten, J. C., Brown, M. G., Balkwill, D. L., & Daly, M. J. (2008). Protein oxidation: Key to bacterial desiccation resistance? *The ISME Journal*, *2*(4), Article 4. https://doi.org/10.1038/ismej.2007.116

Gaidamakova, E. K., Sharma, A., Matrosova, V. Y., Grichenko, O., Volpe, R. P., Tkavc, R., Conze, I. H., Klimenkova, P., Balygina, I., Horne, W. H., Gostinčar, C., Chen, X., Makarova, K. S., Shuryak, I., Srinivasan, C., Jackson-Thompson, B., Hoffman, B. M., & Daly, M. J. (2022). Small-Molecule Mn Antioxidants in Caenorhabditis elegans and *Deinococcus radiodurans* Supplant MnSOD Enzymes during Aging and Irradiation. *mBio*, *13*(1), e0339421. https://doi.org/10.1128/mbio.03394-21

Gao, B., Lara-Tejero, M., Lefebre, M., Goodman, A. L., & Galán, J. E. (2014). Novel components of the flagellar system in epsilonproteobacteria. *mBio*, *5*(3), e01349-01314. https://doi.org/10.1128/mBio.01349-14

Gayen, M., Gupta, P., Morazzani, E. M., Gaidamakova, E. K., Knollmann-Ritschel, B., Daly, M. J., Glass, P. J., & Maheshwari, R. K. (2017). *Deinococcus* Mn2+-peptide complex: A novel approach to alphavirus vaccine development. *Vaccine*, *35*(29), 3672–3681. https://doi.org/10.1016/j.vaccine.2017.05.016

Gérard, E., Jolivet, E., Prieur, D., & Forterre, P. (2001). DNA protection mechanisms are not involved in the radioresistance of the hyperthermophilic archaea

Pyrococcus abyssi and P. furiosus. *Molecular Genetics and Genomics*, *266*(1), 72–78. https://doi.org/10.1007/s004380100520

Ghosal, D., Omelchenko, M. V., Gaidamakova, E. K., Matrosova, V. Y., Vasilenko, A., Venkateswaran, A., Zhai, M., Kostandarithes, H. M., Brim, H., Makarova, K. S., Wackett, L. P., Fredrickson, J. K., & Daly, M. J. (2005). How radiation kills cells: Survival of *Deinococcus radiodurans* and Shewanella oneidensis under oxidative stress⋆. *FEMS Microbiology Reviews*, *29*(2), 361–375. https://doi.org/10.1016/j.fmrre.2004.12.007

Gladyshev, E., & Meselson, M. (2008). Extreme resistance of bdelloid rotifers to ionising radiation. *Proceedings of the National Academy of Sciences*, *105*(13), 5139–5144. https://doi.org/10.1073/pnas.0800966105

Grimsley, J. K., Masters, C. I., Clark, E. P., & Minton, K. W. (1991). Analysis by pulsed-field gel electrophoresis of DNA double-strand breakage and repair in *Deinococcus radiodurans* and a radiosensitive mutant. *International Journal of Radiation Biology*, *60*(4), 613–626. https://doi.org/10.1080/09553009114552441

Hansen, M. T. (1978). Multiplicity of genome equivalents in the radiation-resistant bacterium Micrococcus *radiodurans*. *Journal of Bacteriology*, *134*(1), 71–75. https://doi.org/10.1128/jb.134.1.71-75.1978

Hirsch, P., Gallikowski, C. A., Siebert, J., Peissl, K., Kroppenstedt, R., Schumann, P., Stackebrandt, E., & Anderson, R. (2004). *Deinococcus* frigens sp. Nov., *Deinococcus* saxicola sp. Nov., and *Deinococcus* marmoris sp. Nov., Low Temperature and Draught-tolerating, UV-resistant Bacteria from Continental Antarctica. *Systematic and Applied Microbiology*, *27*(6), 636–645. https://doi.org/10.1078/0723202042370008

Horne, W. H., Volpe, R. P., Korza, G., DePratti, S., Conze, I. H., Shuryak, I., Grebenc, T., Matrosova, V. Y., Gaidamakova, E. K., Tkavc, R., Sharma, A., Gostinčar, C., Gunde-Cimerman, N., Hoffman, B. M., Setlow, P., & Daly, M. J. (2022). Effects of Desiccation and Freezing on Microbial Ionizing Radiation Survivability: Considerations for Mars Sample Return. *Astrobiology*, *22*(11), 1337–1350. https://doi.org/10.1089/ast.2022.0065

Hutchinson, F. (1966). The molecular basis for radiation effects on cells. *Cancer Research*, *26*(9), 2045–2052.

Im, W.-T., Jung, H.-M., Ten, L. N., Kim, M. K., Bora, N., Goodfellow, M., Lim, S., Jung, J., & Lee, S.-T. (2008). *Deinococcus* aquaticus sp. Nov., isolated from fresh water, and *Deinococcus* caeni sp. Nov., isolated from activated sludge. *International Journal of Systematic and Evolutionary Microbiology*, *58*(10), 2348–2353. https://doi.org/10.1099/ijs.0.64082-0

Imlay, J. A. (2006). Iron-sulphur clusters and the problem with oxygen. *Molecular Microbiology*, *59*(4), 1073–1082. https://doi.org/10.1111/j.1365-2958.2006.05028.x

Ithurbide, S., Coste, G., Lisboa, J., Eugénie, N., Bentchikou, E., Bouthier de la Tour, C., Liger, D., Confalonieri, F., Sommer, S., Quevillon-Cheruel, S., & Servant, P. (2020). Natural Transformation in *Deinococcus radiodurans*: A Genetic Analysis Reveals the Major Roles of DprA, DdrB, RecA, RecF, and RecO Proteins. *Frontiers in Microbiology*, *11*, 1253. https://doi.org/10.3389/fmicb.2020.01253

Jeon, S. H., Kang, M. S., Joo, E. S., Kim, E. B., Lim, S., Jeong, S. W., Jung, H. Y., Srinivasan, S., & Kim, M. K. (2016). *Deinococcus* persicinus sp. Nov., a radiationresistant bacterium from soil. *International Journal of Systematic and Evolutionary Microbiology*, *66*(12), 5077–5082. https://doi.org/10.1099/ijsem.0.001473

Jiao, J.-Y., Lian, Z.-H., Li, M.-M., Salam, N., Zhou, E.-M., Liu, L., Ming, H., Nie, G., Shu, W., Zhao, G., Hedlund, B. P., & Li, W.-J. (2022). Comparative genomic analysis of Thermus provides insights into the evolutionary history of an incomplete denitrification pathway. *mLife*, *1*(2), 198–209. https://doi.org/10.1002/mlf2.12009

Jin, M., Xiao, A., Zhu, L., Zhang, Z., Huang, H., & Jiang, L. (2019). The diversity and commonalities of the radiation-resistance mechanisms of *Deinococcus* and its up-to-date applications. *AMB Express*, *9*(1), 138. https://doi.org/10.1186/s13568-019-0862-x

Kaur, R., Rajesh, C., Sharma, R., Boparai, J. K., & Sharma, P. K. (2018). Metagenomic investigation of bacterial diversity of hot spring soil from Manikaran, Himachal Pradesh, India. *Ecological Genetics and Genomics*, *6*(May 2017), 16–21. https://doi.org/10.1016/j.egg.2017.11.003

Kawaguchi, Y., Shibuya, M., Kinoshita, I., Yatabe, J., Narumi, I., Shibata, H., Hayashi, R., Fujiwara, D., Murano, Y., Hashimoto, H., Imai, E., Kodaira, S., Uchihori, Y., Nakagawa, K., Mita, H., Yokobori, S., & Yamagishi, A. (2020). DNA Damage and Survival Time Course of Deinococcal Cell Pellets During 3 Years of Exposure to Outer Space. *Frontiers in Microbiology*, *11*. https://www.frontiersin.org/articles/10.3389/fmicb.2020.02050

Kim, D. U., Jang, J. H., Kang, M.-S., Kim, J.-Y., Zhang, J., Lim, S., & Kim, M. K. (2018). *Deinococcus* irradiati*soli* sp. Nov., isolated from gamma ray-irradiated soil. *International Journal of Systematic and Evolutionary Microbiology*, *68*(10), 3232–3236. https://doi.org/10.1099/ijsem.0.002968

Kim, D. U., Lee, H., Lee, J. H., Ahn, J. H., Lim, S., Jeong, S., Park, S. Y., Seong, C. N., & Ka, J. O. (2015). *Deinococcus* metallilatus sp. Nov. And *Deinococcus* carri

sp. Nov., isolated from a car air-conditioning system. *International Journal of Systematic and Evolutionary Microbiology*, *65*(9), 3175–3182. https://doi.org/10.1099/ijsem.0.000396

Kim, D. U., Lee, H., Lee, S., Park, S., Yoon, J. H., Zhao, L., Kim, M.-K., Ahn, J.-H., & Ka, J.-O. (2018). *Deinococcus* multiflagellatus sp. Nov., isolated from a car air-conditioning system. *Antonie van Leeuwenhoek*, *111*(4), 619–627. https://doi.org/10.1007/s10482-017-0982-8

Kim, E. B., Kang, M. S., Joo, E. S., Jeon, S. H., Jeong, S. W., Lim, S. Y., Jung, H. Y., Srinivasan, S., & Kim, M. K. (2017). *Deinococcus* ruber sp. Nov., a radiation-resistant bacterium isolated from soil. *International Journal of Systematic and Evolutionary Microbiology*, *67*(1), 72–76. https://doi.org/10.1099/ijsem.0.001567

Krisko, A., & Radman, M. (2013). Biology of Extreme Radiation Resistance: The Way of *Deinococcus radiodurans*. *Cold Spring Harbor Perspectives in Biology*, *5*(7), a012765. https://doi.org/10.1101/cshperspect.a012765

Krokan, H. E., & Bjørås, M. (2013). Base Excision Repair. *Cold Spring Harbor Perspectives in Biology*, *5*(4), a012583. https://doi.org/10.1101/cshperspect.a012583

Lai, W. A., Kämpfer, P., Arun, A. B., Shen, F. T., Huber, B., Rekha, P. D., & Young, C. C. (2006). *Deinococcus* ficus sp. Nov., isolated from the rhizosphere of Ficus religiosa L. *International Journal of Systematic and Evolutionary Microbiology*, *56*(4), 787–791. https://doi.org/10.1099/ijs.0.64007-0

Lee, J. H., Jung, J.-H., Kim, M.-K., & Lim, S. (2022). *Deinococcus taeanensis* sp. Nov., a Radiation-Resistant Bacterium Isolated from a Coastal Dune. *Current Microbiology*, *79*(11), 334. https://doi.org/10.1007/s00284-022-03044-8

Lee, J. J., Lee, Y. H., Park, S. J., Lim, S., Jeong, S. W., Lee, S.-Y., Park, S., Choi, H.-W., Kim, M. K., & Jung, H.-Y. (2016). *Deinococcus sedimenti* sp. Nov. Isolated from river sediment. *Journal of Microbiology*, *54*(12), 802–808. https://doi.org/10.1007/s12275-016-6361-8

Lee, J. J., Srinivasan, S., Lim, S., Joe, M., Im, S., & Kim, M. K. (2015). *Deinococcus puniceus* sp. Nov., a Bacterium Isolated from Soil-Irradiated Gamma Radiation. *Current Microbiology*, *70*(4), 464–469. https://doi.org/10.1007/s00284-014-0748-8

Lee, J. J., Lee, Y.-H., Park, S.-J., Lim, S., Jeong, S.-W., Lee, S.-Y., Cho, Y.-J., Kim, M. K., & Jung, H.-Y. (2016). *Deinococcus seoulensis* sp. Nov., a bacterium isolated from sediment at Han River in Seoul, Republic of Korea. *Journal of Microbiology*, *54*(8), 537–542. https://doi.org/10.1007/s12275-016-6253-y

Lemee, L., Peuchant, E., Clerc, M., Brunner, M., & Pfander, H. (1997). Deinoxanthin: A new carotenoid isolated from *Deinococcus radiodurans*. *Tetrahedron*, *53*(3), 919–926. https://doi.org/10.1016/S0040-4020(96)01036-8

Li, X., & Heyer, W.-D. (2008). Homologous recombination in DNA repair and DNA damage tolerance. *Cell Research*, *18*(1), Article 1. https://doi.org/10.1038/cr.2008.1

Lim, S., Jung, J.-H., Blanchard, L., & de Groot, A. (2018). Conservation and diversity of radiation and oxidative stress resistance mechanisms in *Deinococcus* species. *FEMS Microbiology Reviews*, *43*(1), 19–52. https://doi.org/10.1093/femsre/fuy037

Liu, Z., Kim, M. C., Wang, L., Zhu, G., Zhang, Y., Huang, Y., Wei, Z., Danzeng, W., & Peng, F. (2017). *Deinococcus* taklimakanensis sp. Nov., isolated from desert soil. *International Journal of Systematic and Evolutionary Microbiology*, *67*(11), 4311–4316. https://doi.org/10.1099/ijsem.0.002168

Ludanyi, M., Blanchard, L., Dulermo, R., Brandelet, G., Bellanger, L., Pignol, D., Lemaire, D., & de Groot, A. (2014). Radiation response in *Deinococcus deserti*: IrrE is a metalloprotease that cleaves repressor protein DdrO. *Molecular Microbiology*, *94*(2), 434–449. https://doi.org/10.1111/mmi.12774

Makarova, K. S., Aravind, L., Wolf, Y. I., Tatusov, R. L., Minton, K. W., Koonin, E. V., & Daly, M. J. (2001). Genome of the Extremely Radiation-Resistant Bacterium *Deinococcus radiodurans* Viewed from the Perspective of Comparative Genomics. *Microbiology and Molecular Biology Reviews*, *65*(1), 44–79. https://doi.org/10.1128/mmbr.65.1.44-79.2001

Makarova, K. S., & Daly, M. J. (2014). Comparative Genomics of Stress Response Systems in *Deinococcus* Bacteria. In *Bacterial Stress Responses* (pp. 445–457). ASM Press. https://doi.org/10.1128/9781555816841.ch27

Makarova, K. S., Omelchenko, M. V., Gaidamakova, E. K., Matrosova, V. Y., Vasilenko, A., Zhai, M., Lapidus, A., Copeland, A., Kim, E., Land, M., Mavromatis, K., Pitluck, S., Richardson, P. M., Detter, C., Brettin, T., Saunders, E., Lai, B., Ravel, B., Kemner, K. M., … Daly, M. J. (2007). *Deinococcus geothermalis*: The Pool of Extreme Radiation Resistance Genes Shrinks. *PLoS ONE*, *2*(9), e955. https://doi.org/10.1371/journal.pone.0000955

Makarova, K. S., Wolf, Y. I., White, O., Minton, K., & Daly, M. J. (1999). Short repeats and IS elements in the extremely radiation-resistant bacterium *Deinococcus radiodurans* and comparison to other bacterial species. *Research in Microbiology*, *150*(9–10), 711–724. https://doi.org/10.1016/S0923-2508(99)00121-7

Makk, J., Enyedi, N. T., Tóth, E., Anda, D., Szabó, A., Felföldi, T., Schumann, P., Mádl-Szőnyi, J., & Borsodi, A. K. (2019). *Deinococcus* fonticola sp. Nov.,

isolated from a radioactive thermal spring in Hungary. *International Journal of Systematic and Evolutionary Microbiology*, *69*(6), 1724–1730. https://doi.org/10.1099/ijsem.0.003383

Makk, J., Tóth, E. M., Anda, D., Pál, S., Schumann, P., Kovács, A. L., Mádl-Szőnyi, J., Márialigeti, K., & Borsodi, A. K. (2016). *Deinococcus* budaensis sp. Nov., a mesophilic species isolated from a biofilm sample of a hydrothermal spring cave. *International Journal of Systematic and Evolutionary Microbiology*, *66*(12), 5345–5351. https://doi.org/10.1099/ijsem.0.001519

Mattimore, V., & Battista, J. R. (1996). Radiation resistance of *Deinococcus radiodurans*: Functions necessary to survive ionising radiation are also necessary to survive prolonged desiccation. *Journal of Bacteriology*, *178*(3), 633–637. https://doi.org/10.1128/jb.178.3.633-637.1996

McInerney, J. O., Whelan, F. J., Domingo-Sananes, M. R., McNally, A., & O'Connell, M. J. (2020). Pan-genomes and Selection: The Public Goods Hypothesis. In H. Tettelin & D. Medini (Eds.), *The* Pan-genome*: Diversity, Dynamics and Evolution of Genomes* (pp. 151–167). Springer International Publishing. https://doi.org/10.1007/978-3-030-38281-0_7

Mohseni, M., Abbaszadeh, J., & Nasrollahi Omran, A. (2014). Radiation resistant of native *Deinococcus* spp. Isolated from the Lout desert of Iran "the hottest place on Earth". *International Journal of Environmental Science and Technology*, *11*(7), 1939–1946. https://doi.org/10.1007/s13762-014-0643-7

Narasimha, A., & Basu, B. (2021). New insights into the activation of Radiation Desiccation Response regulon in *Deinococcus radiodurans*. *Journal of Biosciences*, *46*(1), 10. https://doi.org/10.1007/s12038-020-00123-5

Omelchenko, M. V., Wolf, Y. I., Gaidamakova, E. K., Matrosova, V. Y., Vasilenko, A., Zhai, M., Daly, M. J., Koonin, E. V., & Makarova, K. S. (2005). Comparative genomics of Thermus thermophilus and *Deinococcus radiodurans*: Divergent routes of adaptation to thermophily and radiation resistance. *BMC Evolutionary Biology*, *5*, 57. https://doi.org/10.1186/1471-2148-5-57

Oyaizu, H., Stackebrandt, E., & Schleifer, K. H. (1987). A radiation-resistant rod-shaped bacterium, Deinobacter grandis gen. Nov., sp. Nov., with peptidoglycan containing ornithine. *International Journal of Systematic Bacteriology*, *37*(1), 62–67. https://doi.org/10.1099/00207713-37-1-62

Pavlov, A. K., Kalinin, V. L., Konstantinov, A. N., Shelegedin, V. N., & Pavlov, A. A. (2006). Was Earth Ever Infected by Martian Biota? Clues from Radioresistant Bacteria. *Astrobiology*, *6*(6), 911–918. https://doi.org/10.1089/ast.2006.6.911

Peng, F., Zhang, L., Luo, X., Dai, J., An, H., Tang, Y., & Fang, C. (2009). *Deinococcus* xinjiangensis sp. Nov., isolated from desert soil. *International Journal of*

Systematic and Evolutionary Microbiology, *59*(4), 709–713. https://doi.org/10.1099/ijs.0.004564-0

Rainey, F. A., Ferreira, M., Nobre, M. F., Ray, K., Bagaley, D., Earl, A. M., Battista, J. R., Gómez-Silva, B., McKay, C. P., & da Costa, M. S. (2007). *Deinococcus* peraridilitoris sp. Nov., isolated from a coastal desert. *International Journal of Systematic and Evolutionary Microbiology*, *57*(7), 1408–1412. https://doi.org/10.1099/ijs.0.64956-0

Rainey, F. A., Ray, K., Ferreira, M., Gatz, B. Z., Nobre, M. F., Bagaley, D., Rash, B. A., Park, M. J., Earl, A. M., Shank, N. C., Small, A. M., Henk, M. C., Battista, J. R., Kämpfer, P., & Da Costa, M. S. (2005). Extensive diversity of ionising-radiation-resistant bacteria recovered from Sonoran Desert soil and description of nine new species of the genus *Deinococcus* obtained from a single soil sample. *Applied and Environmental Microbiology*, *71*(9), 5225–5235. https://doi.org/10.1128/AEM.71.9.5225-5235.2005

Robinson, C. K., Webb, K., Kaur, A., Jaruga, P., Dizdaroglu, M., Baliga, N. S., Place, A., & DiRuggiero, J. (2011). A major role for nonenzymatic antioxidant processes in the radiation resistance of Halobacterium salinarum. *Journal of Bacteriology*, *193*(7), 1653–1662. https://doi.org/10.1128/JB.01310-10

Rosenberg, E. (2014). The family deinococcaceae. In *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea* (pp. 613–615). Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-642-38954-2_127

Santos, S. P., Cuypers, M. G., Round, A., Finet, S., Narayanan, T., Mitchell, E. P., & Romão, C. V. (2017). SAXS Structural Studies of Dps from *Deinococcus radiodurans* Highlights the Conformation of the Mobile N-Terminal Extensions. *Journal of Molecular Biology*, *429*(5), 667–687. https://doi.org/10.1016/j.jmb.2017.01.008

Santos, S. P., Yang, Y., Rosa, M. T. G., Rodrigues, M. A. A., De La Tour, C. B., Sommer, S., Teixeira, M., Carrondo, M. A., Cloetens, P., Abreu, I. A., & Romão, C. V. (2019). The interplay between Mn and Fe in *Deinococcus radiodurans* triggers cellular protection during paraquat-induced oxidative stress. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-53140-2

Sharma, A., Gaidamakova, E. K., Grichenko, O., Matrosova, V. Y., Hoeke, V., Klimenkova, P., Conze, I. H., Volpe, R. P., Tkavc, R., Gostinčar, C., Gunde-Cimerman, N., Diruggiero, J., Shuryak, I., Ozarowski, A., Hoffman, B. M., Daly, M. J., Designed, M. J. D., & Performed, A. O. (2017). *Across the tree of life, radiation resistance is governed by antioxidant Mn 2+, gauged by paramagnetic resonance*. https://doi.org/10.1073/pnas.1713608114

Shashidhar, R., & Bandekar, J. R. (2009). *Deinococcus* piscis sp. Nov., a radiation-resistant bacterium isolated from a marine fish. *International Journal of*

Systematic and Evolutionary Microbiology, *59*(11), 2714–2717. https://doi.org/10.1099/ijs.0.003046-0

Shashidhar, R., Kumar, S. A., Misra, H. S., & Bandekar, J. R. (2010). Evaluation of the role of enzymatic and nonenzymatic antioxidant systems in the radiation resistance of *Deinococcus* . *Canadian Journal of Microbiology*, *56*(3), 195–201. https://doi.org/10.1139/w09-118

Shuryak, I., Tkavc, R., Matrosova, V. Y., Volpe, R. P., Grichenko, O., Klimenkova, P., Conze, I. H., Balygina, I. A., Gaidamakova, E. K., & Daly, M. J. (2019). Chronic gamma radiation resistance in fungi correlates with resistance to chromium and elevated temperatures, but not with resistance to acute irradiation. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-47007-9

Slade, D., Lindner, A. B., Paul, G., & Radman, M. (2009). Recombination and Replication in DNA Repair of Heavily Irradiated *Deinococcus radiodurans*. *Cell*, *136*(6), 1044–1055. https://doi.org/10.1016/j.cell.2009.01.018

Slade, D., & Radman, M. (2011). Oxidative Stress Resistance in *Deinococcus radiodurans*. *Microbiology and Molecular Biology Reviews*, *75*(1), 133–191. https://doi.org/10.1128/mmbr.00015-10

Sleytr, U. B., Silva, M. T., Kocur, M., & Lewis, N. F. (1976). The fine structure of Micrococcus radiophilus and Micrococcus radioproteolyticus. *Archives of Microbiology*, *107*(3), 313–320. https://doi.org/10.1007/BF00425346

Soppa, J. (2014). Polyploidy in archaea and bacteria: About desiccation resistance, giant cell size, long-term survival, enforcement by a eukaryotic host and additional aspects. *Journal of Molecular Microbiology and Biotechnology*, *24*(5–6), 409–419. https://doi.org/10.1159/000368855

Srinivasan, S., Lee, J.-J., Lim, S., Joe, M., & Kim, M. K. (2012). *Deinococcus* humi sp. Nov., isolated from soil. *International Journal of Systematic and Evolutionary Microbiology*, *62*, 2844–2850. https://doi.org/10.1099/ijs.0.037234-0

Sun Joo, E., Jin Lee, J., Kang, M.-S., Lim, S., Jeong, S., Bit Kim, E., Hwa Jeon, S., Srinivasan, S., & Kyum Kim, M. (2016). *Deinococcus* actinosclerus sp. Nov., a novel bacterium isolated from soil of a rocky hillside. *International Journal of Systematic and Evolutionary Microbiology*, *66*(2), 1003–1008. https://doi.org/10.1099/ijsem.0.000825

Suresh, K., Reddy, G. S. N., Sengupta, S., & Shivaji, S. (2004). *Deinococcus* indicus sp. Nov., an arsenic-resistant bacterium from an aquifer in West Bengal, India. *International Journal of Systematic and Evolutionary Microbiology*, *54*(2), 457–461. https://doi.org/10.1099/ijs.0.02758-0

Tanaka, M., Earl, A. M., Howell, H. A., Park, M.-J., Eisen, J. A., Peterson, S. N., & Battista, J. R. (2004). Analysis of *Deinococcus radiodurans*'s transcriptional response to ionising radiation and desiccation reveals novel proteins that

contribute to extreme radioresistance. *Genetics*, *168*(1), 21–33. https://doi.org/10.1534/genetics.104.029249

Tettelin, H., Riley, D., Cattuto, C., & Medini, D. (2008). Comparative genomics: The bacterial pangenome. *Current Opinion in Microbiology*, *11*(5), 472–477. https://doi.org/10.1016/j.mib.2008.09.006

Tian, J., Wang, L., Liu, P., Geng, Y., Zhu, G., Zheng, R., Liu, Z., Zhao, Y., Yang, J., & Peng, F. (2019). *Deinococcus* psychrotolerans sp. Nov., isolated from soil on the South Shetland Islands, Antarctica. *International Journal of Systematic and Evolutionary Microbiology*, *69*(12), 3696–3701. https://doi.org/10.1099/ijsem.0.003484

Vaishampayan, P., Roberts, A. H. ayden, Augustus, A., Pukall, R., Schumann, P., Schwendner, P., Mayilraj, S., Salmassi, T., & Venkateswaran, K. (2014). *Deinococcus* phoenicis sp. Nov., an extreme ionising-radiation-resistant bacterium isolated from the Phoenix Lander assembly facility. *International Journal of Systematic and Evolutionary Microbiology*, *64*(Pt 10), 3441–3446. https://doi.org/10.1099/ijs.0.063107-0

Wang, J.-J., Wu, S.-G., Chen, Q., Sheng, D.-H., Du, Z.-J., & Li, Y.-Z. (2020). *Deinococcus* terrestris sp. Nov., a gamma ray- and ultraviolet-resistant bacterium isolated from soil. *International Journal of Systematic and Evolutionary Microbiology*, *70*(9), 4993–5000. https://doi.org/10.1099/ijsem.0.004369

Wang, X. P., Li, C.-M., Yu, Y., Li, H.-R., Du, Z.-J., & Mu, D. (2019). *Deinococcus* arcticus sp. Nov., isolated from Silene acaulis rhizosphere soil of the Arctic tundra. *International Journal of Systematic and Evolutionary Microbiology*, *69*(11), 3437–3442. https://doi.org/10.1099/ijsem.0.003636

White, O., Eisen, J. A., Heidelberg, J. F., Hickey, E. K., Peterson, J. D., Dodson, R. J., Haft, D. H., Gwinn, M. L., Nelson, W. C., Richardson, D. L., Moffat, K. S., Qin, H., Jiang, L., Pamphile, W., Crosby, M., Shen, M., Vamathevan, J. J., Lam, P., McDonald, L., … Fraser, C. M. (1999). Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. *Science (New York, N.Y.)*, *286*(5444), 1571–1577. https://doi.org/10.1126/science.286.5444.1571

Xu, X., Jiang, L., Zhang, Z., Shi, Y., & Huang, H. (2013). Genome Sequence of a Gamma- and UV-Ray-Resistant Strain, *Deinococcus* wulumuqiensis R12. *Genome Announcements*, *1*(3), 10.1128/genomea.00206-13. https://doi.org/10.1128/genomea.00206-13

Yang, Y., Itoh, T., Yokobori, S., Itahashi, S., Shimada, H., Satoh, K., Ohba, H., Narumi, I., & Yamagishi, A. (2009). *Deinococcus* aerius sp. Nov., isolated from the high atmosphere. *International Journal of Systematic and Evolutionary Microbiology*, *59*(8), 1862–1866. https://doi.org/10.1099/ijs.0.007963-0

Yuan, M., Zhang, W., Dai, S., Wu, J., Wang, Y., Tao, T., Chen, M., & Lin, M. (2009). *Deinococcus* gobiensis sp. Nov., an extremely radiation-resistant bacterium. *INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY*, *59*(6), 1513–1517. https://doi.org/10.1099/ijs.0.004523-0

Zahradka, K., Slade, D., Bailone, A., Sommer, S., Averbeck, D., Petranovic, M., Lindner, A. B., & Radman, M. (2006). Reassembly of shattered chromosomes in *Deinococcus radiodurans*. *Nature*, *443*(7111), 569–573. https://doi.org/10.1038/nature05160

Zhang, K., Zhu, J., Li, S., Rao, M. P. N., Li, N.-M., Guo, A.-Y., Jiang, Z., Tang, Q.-Y., Wan, Y., Zhang, Z.-D., & Li, W.-J. (2020a). *Deinococcus* detaillensis sp. Nov., isolated from humus soil in Antarctica. *Archives of Microbiology*, *202*(9), 2493–2498. https://doi.org/10.1007/s00203-020-01920-0

Zhang, K., Zhu, J., Li, S., Rao, M. P. N., Li, N.-M., Guo, A.-Y., Jiang, Z., Tang, Q.-Y., Wan, Y., Zhang, Z.-D., & Li, W.-J. (2020b). *Deinococcus* detaillensis sp. Nov., isolated from humus soil in Antarctica. *Archives of Microbiology*, *202*(9), 2493–2498. https://doi.org/10.1007/s00203-020-01920-0

Zhou, Z., Zhang, W., Su, S., Chen, M., Lu, W., Lin, M., Molnár, I., & Xu, Y. (2015). CYP287A1 is a carotenoid 2-β-hydroxylase required for deinoxanthin biosynthesis in *Deinococcus radiodurans* R1. *Applied Microbiology and Biotechnology*, *99*(24), 10539–10546. https://doi.org/10.1007/s00253-015-6910-9

Zhu, J., Li, S.-H., Tang, Q.-Y., Chu, M., Wang, W., Salam, N., Li, L., Hozzein, W. N., Zhang, Z.-D., & Li, W.-J. (2017). *Deinococcus* malanensis sp. Nov., isolated from radiation-polluted soil. *Archives of Microbiology*, *199*(4), 621–626. https://doi.org/10.1007/s00203-016-1335-0

2 Chapter II: Genomic Diversity and Ubiquity of the *Deinococcus* Genus

Question:

What factors contribute to the widespread distribution and diverse characteristics of the *Deinococcus* genus?

Hypothesis:

The ubiquity of *Deinococcus* is due to its genetic flexibility and large pangenome.

2.1 Abstract:

The *Deinococcus* genus is widely known for the resistance of its type species, *Deinococcus radiodurans*, to extreme doses of ionising radiation. A few *Deinococcus* species have been extensively studied for this outstanding phenotype, and this focus has led to the conventional wisdom that radiation resistance is the defining characteristic of the *Deinococcaceae* family. Microbiologists have used this concept and widely used gamma radiation as a pre-cultivation treatment method to eliminate radiosensitive cells from environmental samples. However, a few cultivation studies accidentally isolated radio-sensitive *Deinococcus* species without applying gamma radiation. In addition to the radiation resistance phenotype, *Deinococcus* species are ubiquitous and isolated from diverse habitats. This feature has been associated with their genomic flexibility and acquiring diverse functions to adapt to new environments.

Nonetheless, focusing on radiation resistance phenotype and the cultivation methods has shaped a biased population of known *Deinococcus* species, which might not represent all members of this diverse genus. As a result, this approach has obscured our understanding of the ecology of the *Deinococcus* genus, transforming it into an enigmatic taxon in microbiology.

This thesis aims to address this gap in knowledge by using comparative genomics methods and focusing on the genomic diversity of all available *Deinococcus* members instead of some species with high radiation resistance levels to provide insight into the poorly understood ecology of this genus. We use the pangenome concept to construct the genomic repertoire of *Deinococcus* members and look into the functional properties that shape the highly diverse members of the *Deinococcus* genus. We show that *Deinococcus* members have highly diverse genomic features which shape their open pangenome and explain their ubiquity. This thesis revealed that amino acid and carbohydrate transporters, as well as different families of transposases, constitute a large proportion of the accessory genome, which are found in some but not all members. Comparing the metabolic capabilities of *Deinococcus* species revealed the absence of many cofactors and vitamins and iron and manganese oxidation metabolisms. We also

show that nitrite reduction to ammonia is common in *Deinococcus* species isolated from freshwater but less common in other environments. Sulfate reduction is conserved among most *Deinococcus* species except for a few host-associated species.

We conclude that viewing the *Deinococcus* genus only through the lens of radiation resistance phenotype has biased our understanding of the ecology and diversity of this genus. Moreover, the genomes of the known *Deinococcus* species can capture only a fraction of genomic diversity in all *Deinococcus,* and it is necessary to develop novel methods for the isolation of *Deinococcus* species to avoid the disproportionate known *Deinococcus* species in terms of their ecological features.

2.2 Introduction

The *Deinococcus* genus is known for the extraordinary capability of some of its members to survive high doses of gamma radiation and prolonged desiccation (Rosenberg, 2014). The type species of this genus, *Deinococcus radiodurans,* was initially isolated from irradiated meat cans in 1956 and named *Micrococcus radiodurans* because of its similarities with the *Micrococcus* genus (Anderson et al., 1956). Later, in 1981, Brooks and Murray proposed a new nomenclature for five known radioresistant species in the *Micrococcus* genus, which formed a polyphyletic cluster with other members based on 16S rRNA gene classification. They proposed a new genus, *Deinococcus,* within the new family, *Deinococcaceae* (Brooks & Murray, 1981). *Deinococcaceae* family comprises two genera, *Deinococcus* and *Deinobacterium*. Since the *Deinobacterium* genus has only one member, *D. chartae* (Ekman et al., 2011), the main focus of this chapter will be on the *Deinococcus* genus.

D. radiodurans is the type species of the *Deinococcus* genus and is resistant to extremely high doses of gamma radiation. While only 5 to 10 grays (Gy) of gamma radiation can kill humans within days, 5000 Gy of gamma radiation does not decrease the populations of *D. radiodurans.* In comparison, *Escherichia coli* and *Thermus thermophilus* populations diminish to 10% after exposure to only 800 Gy of gamma radiation. The D_{10} value (to the dose required to kill 90% of the population) for *D. radiodurans* has been reported as 12 kGy (Daly et al., 2010). This extraordinary radiation resistance capability was also observed in other *Deinococcus* species, such as *D. geothermalis, D. proteolyticus,* and *D. gobiensi*s (Slade & Radman, 2011). However, *D. radiodurans* has been extensively studied, becoming the radiobiology model organism.

In the early years of isolation and characterisation of *Deinococcus* species, high levels of resistance to ionising radiation observed in isolated species led to the notion that all *Deinococcus* species are radiation resistant (Brooks & Murray, 1981; Murray, 1992). This view has remained popular to date to the extent that radiation resistance was defined as the defining characteristic of the *Deinococcaceae* family (Daly, 2023). This

conventional wisdom has always viewed the *Deinococcus* members through the lens of extreme radiation resistance (Baudet et al., 2010; Daly, 2023; Lim et al., 2018; Makarova et al., 2007; Panitz et al., 2019).

Contrary to the popular notion that radiation resistance is the defining characteristic of *Deinococcaceae*, a diverse range of radiation resistance capacities have been reported among its members. The D_{10} value among some *Deinococcus* members varies in order of magnitude. For instance, *D. persicinus* isolated from soil has a D_{10} value of 1.5 kGy, which is ten times more sensitive to gamma radiation than *D. radiodurans* (Jeon et al., 2016). Other examples of radiation-sensitive species are *D. alpinitundrae*, *D. altitudinis*, *D. claudionis*, *D. radiomollis* (Callegan et al., 2008), *D. detaillensis*, *D. sedimenti*, *D. seoulensis*, *D. indicus*, and *D. arcticus* (J. Lee et al., 2016; Suresh et al., 2004; X.-P. Wang et al., n.d.; Zhang et al., 2020a)(Table 3-1).

As of writing this thesis in January 2024, more than 300 *Deinococcus* strains have been cultivated and identified based on the 16S rRNA gene sequences and deposited to NCBI (https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=1298). However, only 92 have been named and described according to the List of Prokaryotic Names with Standing in Nomenclature (LSPN) (https://www.bacterio.net/genus/*Deinococcus*). Members of the *Deinococcus* genus are ubiquitous and have been isolated from diverse habitats. Their habitats range from extremely hot and hyper arid environments like the Sahara and Lut deserts to extremely cold environments like Antarctica and alpine environments (Callegan et al., 2008; de Groot et al., 2005; Hirsch et al., 2004; Mohseni et al., 2014). *Deinococcus* species have also been found in geothermal pools (Ferreira et al., 1997), freshwater sediments (J. Lee et al., 2016), atmospheric dust (Yang et al., 2009), the gut of marine and terrestrial animals, and the surface of metals in urban regions (W. Chen et al., 2011; Lai et al., 2006). This widespread distribution of *Deinococcus* members can be associated with their genomic flexibility and their evolution pathway in the direction of gaining new functions to adapt to extreme environments (Krisko & Radman, 2013; Omelchenko et al., 2005)

Even though *Deinococcus* species are widespread in the environment, they are relatively rare and have a patchy distribution in diverse habitats (Krisko & Radman, 2013). This low abundance of members of this genus in the environment and their chemoorganotrophic lifestyle make their cultivation difficult. Therefore, microbiologists have followed the described conventional wisdom that radiation resistance is the defining characteristic of *Deinococcaceae* (Daly, 2023) and used gamma radiation as a pre-cultivation treatment method to eliminate sensitive populations from the environmental sample (de Groot et al., 2005; Ferreira et al., 1997; Jin et al., 2019; Mohseni et al., 2014; Zhang et al., 2020a). This method has proved helpful in isolating highly radiation-resistant strains, but the indiscriminate elimination of all radiation-sensitive cells in the environment can also eradicate *Deinococcus* species with potentially lower radiation resistance capability. This trend in isolation and cultivation of *Deinococcus* species could have led to a disproportionate population of known species, specifically the under-representation of radiation-sensitive species. Consequently, despite the extensive research on the radiation resistance of *D. radiodurans* and the cultivation of hundreds of isolates, their ecology and phenotypic diversity have remained poorly understood. It is crucial to consider diverse genomic and environmental data to address this gap in knowledge.

In the past decade, and with the exponential availability of genomic and environmental data, comparative genomics has enabled a comprehensive study of many prokaryotic genera and species (Chen et al., 2021; Fang et al., 2022; Liao et al., 2021; Tettelin et al., 2008). The genome of *D. radiodurans* was sequenced in 1999 due to its extraordinary radiation resistance phenotype and potential application in biotechnology (White et al., 1999). Shortly after sequencing, researchers began to compare its genome with the model organism *E. coli* (Makarova et al., 2001) and a few years later with *Thermus thermophilus* as a close relative (Omelchenko et al., 2005). Later, when more *Deinococcus* genomes became available, scientists started to compare multiple *Deinococcus* species, and the first such study was the comparative genomics of *D. geothermalis* (Makarova et al., 2007). The latest comparative study of 11 *Deinococcus* species has revealed a complex interplay of conserved and diverse mechanisms

contributing to radiation resistance within the *Deinococcus* genus (Lim et al., 2018). However, the focus of all these studies has been merely on radiation resistance phenotype, and no comparative genomic study focused on the genomic flexibility of *Deinococcus* species and their metabolic diversity.

Pangenome analysis has emerged as a valuable method in comparative genomics for capturing the complete genetic diversity within a species and reducing bias in genetic analysis inherent in using a single reference genome (Wang et al., 2023). The pangenome approach has been instrumental in understanding the genetic basis of microbial adaptation to diverse ecological niches and the functional implications of genetic diversity within microbial populations (Jonkheer et al., 2021). This study uses comparative genomics methods with a pangenome approach to answer three main questions. (i). Does available genomes represent the whole pangenome of the *Deinococcus* genus? (ii). What is the functional foundation for the genetic diversity and genome size variability in the *Deinococcus* genus? (iii). Is there any significant difference in the metabolic capability of *Deinococcus* species isolated from different habitats? We constructed the pangenome of the *Deinococcus* genus using all available high-quality genomes in the public database and two newly sequenced radiation-sensitive *Deinococcus* species. Then, we annotated the *Deinococcus* pangenome and explored the enrichment of COG functional categories within the core, accessory, and unique genes. Finally, we mapped the metabolic pathways of *Deinococcus* members to their habitats to gain insight into the ecology and metabolic diversity of the *Deinococcus* members.

2.3 Methods

2.3.1 Bacterial strains and culture conditions

In this study, we sequenced two radiation-sensitive *Deinococcus* species to address the under-representation of this group and diversify our genomic dataset. *D. radiomollis,* with a D_{10} value of 2.2 kGy, and *D. altitudinis,* with a D_{10} value of 3.8 kGy, are two psychrophilic species that were isolated from Pico de Orizaba, Mexico, and Mount Evans, CO, USA, respectively, without gamma radiation as a pre-culture treatment. Both isolates showed sensitivity to desiccation and ionising radiation (Callegan et al., 2008). Strains were purchased from the Belgian Coordinated Collections of Microorganisms (BCCM) and used for sequencing. Lyophilised cells were revived following the instruction of BCCM and cultivated on ten-fold diluted Plate Count Agar (PCA) (Tryptone 0.5 g, Yeast extract 0.25 g, Glucose 0.1 g, Agar 1.5 g, Distilled water 1 L, pH 7,2). The plates were incubated at optimum growth temperature, 10°C, for seven days, and single colonies were inoculated in a liquid state of the same culture medium. Cell pellets were collected and used for genomic DNA extraction.

2.3.2 Genome sequencing:

We combined the long-read sequencing method using Minion with short-read sequencing Illumina to generate a circular genome. High molecular weight gDNA for long-read sequencing was extracted using Nanobind® CBB kit for Gram-positive bacteria (then Circulomics) and following the manufacturer's protocol. The DNA QC was carried out by measuring the DNA concentration using the Qubit dsDNA HS kit (ThermoFisher Scientific, USA). The purity of the DNA was determined using a NanoDrop Spectrophotometer (Denovix DS-11 FX), and the absorbance ratio at 260/280 nm was ~ 1.8. We prepared the library using the ligation sequencing kits SQK-LSK109 (Oxford Nanopore Technologies) and NEBNext® Companion Module for Oxford Nanopore Technologies® Ligation Sequencing. The prepared library was loaded into R.9.4.1 flowcell on a MinION device (Oxford Nanopore Technologies), and sequencing data were generated using the MinKNOW v 4.1.22 software package (https://community.nanoporetech.com/downloads). For short-read sequencing, the

Illumina libraries were prepared using NEB Ultra II DNA Library Prep Kit for Illumina® (NewEngland BioLabs) and following the facility's protocols. Libraries were prepared and sequenced at the Genewiz sequencing facility (GENEWIZ China & Suzhou Lab).

2.3.3 De novo assembly

For the *de-novo* assembly, we prepared MinIon and Illumina raw reads as follows. We performed the base-calling step later (not real-time) using GPU-accelerated guppy basecaller v 5.0.7 (Wick et al., 2019) on the New Zealand eScience Infrastructure (NeSI) HPC cluster. We used FastQC to evaluate the quality of reads from Illumina and MinIOn. Then, Porechop (v 0.2.4) was used to find and remove MinION adapters. Nanofilt was used to filter out reads that were too short and too long (500 bp and 75,000 bp). We used the Trycycler pipeline (v 0.5.4) (Wick et al., 2021) to assemble the MinIOn nanopore sequence. This tool uses multiple long read assemblers, including Flye (v.2.9.1), miniasm_and_minipolish(0.3), and raven (1.5.0), to generate consensus sequence scripts provided in the code availability. Then, we used quality-filtered Illumina reads to polish the long read to increase the genome's accuracy using Polypolish (Wick & Holt, 2022). We deposited genome sequences to the NCBI under Bioproject number PRJNA991629 and accession numbers SAMN36308717 and SAMN36368281 for *D. altitudinis* and *D. radiomollis*, respectively.

2.3.4 Dataset compilation

According to the NCBI taxonomy, the *Deinococcaceae* family has two genera, *Deinococcus* and *Deinobacterium* (Rosenberg, 2014). However, based on the GTDB taxonomy, the *Deinococcaceae* is categorised into seven genera, including *Deinobacterium*, *Deinococcus*, *Deinococcus_A*, *Deinococcus_B*, and *Deinococcus_C* represented by cultured species and two uncultured genera JACMOA01 and JAJZIR01 (Parks et al., 2018). As of the writing of this thesis, the genus *Deinococcus* has 89 classified and published species (https://www.bacterio.net/genus/*Deinococcus*). Among these genomes, only 69 representative genomes, including MAGs, were included in the GTDB database. We used the metadata from GTDB v214 (Parks et al.,

2022) to compile a dataset for 69 *Deinococcus* species with their metadata on genome size, protein counts, CheckM completeness, and GC content. The isolation habitat for each species was extracted from the literature (Supplementary Data 1-1). Genome quality was estimated based on GTDB criteria (completeness – 5 × contamination) > 90 (Parks et al., 2018). After quality assessment, representative genomes for 64 *Deinococcus* species were selected, and the two sequenced genomes in this study were added to the dataset. A total of 66 *Deinococcus* species based on NCBI taxonomy were used for pangenome analysis (Supplementary Data S3-1). Open reading frames (ORFs) were predicted from the genomic DNA sequences using the Prokka package v 1.14.5 (Seemann, 2014), and the output proteome sequences were used for phylogenomic and pangenome analyses.

2.3.5 Pangenome analysis

2.3.5.1 Pangenome construction

A pangenome refers to the whole genomic repertoire of a taxonomic group like a genus or species (Tettelin et al., 2008). To construct the pangenome of the *Deinococcus* genus, we used OrthoFinder v.2.5.2 with default parameters to infer gene families or orthogroups (Emms & Kelly, 2019). In addition to the 66 *Deinococcus* genomes selected for this study, three genomes from close relatives, *Deinobacterium chartae, Truepera radiovictrix* and *Thermus thermophilus*, were added as outgroups. OrthoFinder generates more accurate orthogroups when the outgroup is used in the analysis (Emms & Kelly, 2020). The outgroup species were filtered out later and excluded from pangenome construction. A pangenome includes the core genome (softcore genome), accessory genome, and unique genes. The core genome refers to genes shared between all genomes. Soft-core is defined as genes that are present in 95% of genomes (Contreras-Moreira & Vinuesa, 2013). Accessory genome refers to genes shared between more than one genome, but not all, and unique genes are genes present only in one genome. We categorised orthogroups into core, accessory, and unique genes using the pandas package (Reback et al., 2020) with Python 3.9.

2.3.5.2 Pangenome openness assessment

We sought to determine the openness of the pangenome by calculating the response of adding new genomes to the increase or decrease in the core, accessory, and unique groups based on the models proposed by Tettelin (Tettelin et al., 2008). We started with three genomes and calculated 10,000 permutations for adding new genomes. Then, we reclassified pangenome groups in each iteration until all genomes were added. Then, we calculated the median value of random samples in each datapoint (number of genomes) for changes in pangenome groups (decrease in core and increase in accessory and unique) and fitted Heaps' law using the median values to create the rarefaction curves. Heaps law is a power law introduced by Tettelin (Tettelin et al., 2008) and suggests that the chance of discovering new genes becomes increasingly less as the number of new genome samples increases. This model is used to determine the openness of the pangenome (Jonkheer et al., 2021; Liao et al., 2021; Tettelin et al., 2008). The Heaps' law model was calculated using the Microplane R package v2.1 (Snipen & Liland, 2015) by the formula $n = k \times N^{-\alpha}$ where n is the newly observed genes, N is the total number of genomes, α, and k are the fitting parameters. A pangenome follows Heaps' law and is open when $\alpha < 1$, meaning more genome samples are required for pangenome size to reach a constant rate. However, $\alpha > 1$ means a pangenome is closed, and adding new genome samples does not increase the rate of observing new genes (Tettelin et al., 2008).

2.3.6 Phylogenomic tree reconstruction

To generate the species tree for the *Deinococcus* genus, we used the 66 *Deinococcus* species in our dataset and the neighbouring taxa, *Deinobacterium chartae, Truepera radiovictrix* and *Thermus thermophilus,* were used as outgroup species. We used 476 single-copy orthologous groups shared between all genomes (identified from the OrthoFinder). Multiple sequence alignment (MSA) files were generated for each gene family using MAFFT (Katoh & Standley, 2013) and trimmed using trimAl (Capella-Gutiérrez et al., 2009). A concatenated supermatrix file was used for species tree reconstruction. IQ-TREE v.2.2.2.2 (Minh et al., 2020) was used for the generation of

species tree with the model finder function and 1000 ultra-fast bootstraps (iqtree2 -m MFP -madd LG+C20, LG+C60). The best substitution model was chosen as LG+F+R8 based on the Bayesian Information Criterion (BIC). Tree manipulation, visualisation, and annotation were done in R and using the ggtree and Treeio packages (L.-G. Wang et al., 2020; Yu, 2020)

2.3.7 Functional annotation of pangenome groups:

To determine the functional categories of the pangenome groups in *Deinococcus*, we used eggNOG mapper v 2.0.1 (Cantalapiedra et al., 2021) to annotate each gene in orthogroups. All orthogroups were concatenated and used as input to the EggNOG mapper. Then, we calculated the abundance of genes belonging to each COG category and compared the presence of core accessory and unique genes. We also calculated the composition of pangenome groups in each genome.

2.3.8 Radiation Desiccation Response genes

To closely investigate the genomic basis of radiation-resistance phenotype variability among *Deinococcus* species, we curated a custom database comprising a list of genes identified through transcriptomics and gene knockout studies and shown to be involved in response to ionising radiation and desiccation. These genes are part of a regulon called Radiation Desiccation Response (RDR) in *D. radiodurans*. A recent study used a combination of ChIP-seq, RNA-seq, and in-silico methods and extended the RDR regulon genes to 37 (Eugénie et al., 2021). We used those genes as a marker for radiation resistance to compare the presence and absence of those genes in the pangenome.

We curated our database using a collection of homologous genes to those found in *D. radiodurans*. All *Deinococcus* genomes were searched against our custom protein databases using DIAMOND v2.1.6 (query cover > 80%) (Buchfink et al., 2015). Finally, we used DIAMOND blast to search for the presence and abundance of those genes among three resistance groups: low resistance, moderate resistance, and high resistance.

2.3.9 Metabolic profile and functional networks

To gain insight into the ecological potentials of the *Deinococcus* species under study, we used our general dataset genomes to investigate the presence or absence of metabolic pathways. We used the METABOLIC toolkit v4.0 (METabolic And BiogeOchemistry anaLyses In miCrobes) (Zhou et al., 2022). The "METABOLIC-G" option was used to profile metabolic and biogeochemical traits and functional networks based on prior biochemical validation and the presence or absence of metabolic pathways. METABOLIC toolkit uses an integrated annotation of proteins using KEGG (Kanehisa & Goto, 2000), TIGRfam (Selengut et al., 2007), Pfam (Finn et al., 2014), the custom hidden Markov model (HMM) databases. We used the results of the custom database from the METABOLIC toolkit as well as KEGG pathways to investigate the metabolic pathways.

2.4 Results:

2.4.1 Genome sequencing of sensitive *Deinococcus*

To address the genomic under-representation of radiation-sensitive *Deinococcus* species, we sequenced two *Deinococcus* species, *D. altitudinis* LMG 24022 and *D. radiomollis* LMG 24019 and included them in our comparative genomic analysis. These two species were isolated from alpine environments and showed phylogenomic and phenotypic similarities despite inhabiting geographically distant environments (Pico de Orizaba, Mexico and Mount Evans, CO, USA) (Callegan et al., 2008). The two isolates showed ANI 86.29% and formed a single clade in the phylogenomic tree. (Figure 2-3) *D. altitudinis* has six replicons, a 3,046,749 bp chromosome and five plasmids ranging from 731,747 bp to 144,095. Ten replicons constitute the genomic content of *D. radiomollis,* including one 3,359,780 bp chromosome and nine plasmids ranging from 905,485 bp to 51,515 bp (Table 2-1)

Table 2-1. Genom statistics or genomic features of sequenced sensitive *Deinococcus* species

Features	*D. altitudinis*	*D. radiomollis*
Status	Complete Circular	Complete Circular
Genome size bp	5,041,628	5,756,693
Chromosome numbers	1	1
Plasmid numbers	5	7
Proteins	4611	5607
GC-Content	66%	64.10%
Accession	SAMN36308717	SAMN36368281

2.4.2 Genomic features of *Deinococcus*

Our results indicate that the genomic features of the *Deinococcus* species are highly diverse. Genome size ranges from 2.7 Mbp in thermophilic *D. murrayi* to 6.65 Mbp in the *D. hopiensis* isolated from Sonoran desert soil**.** The GC content is also highly variable among *Deinococcus* species and spans from 55% in *D. misasensis* to 70% in *D. actinosclerus* (Supplementary Data S2-1 and

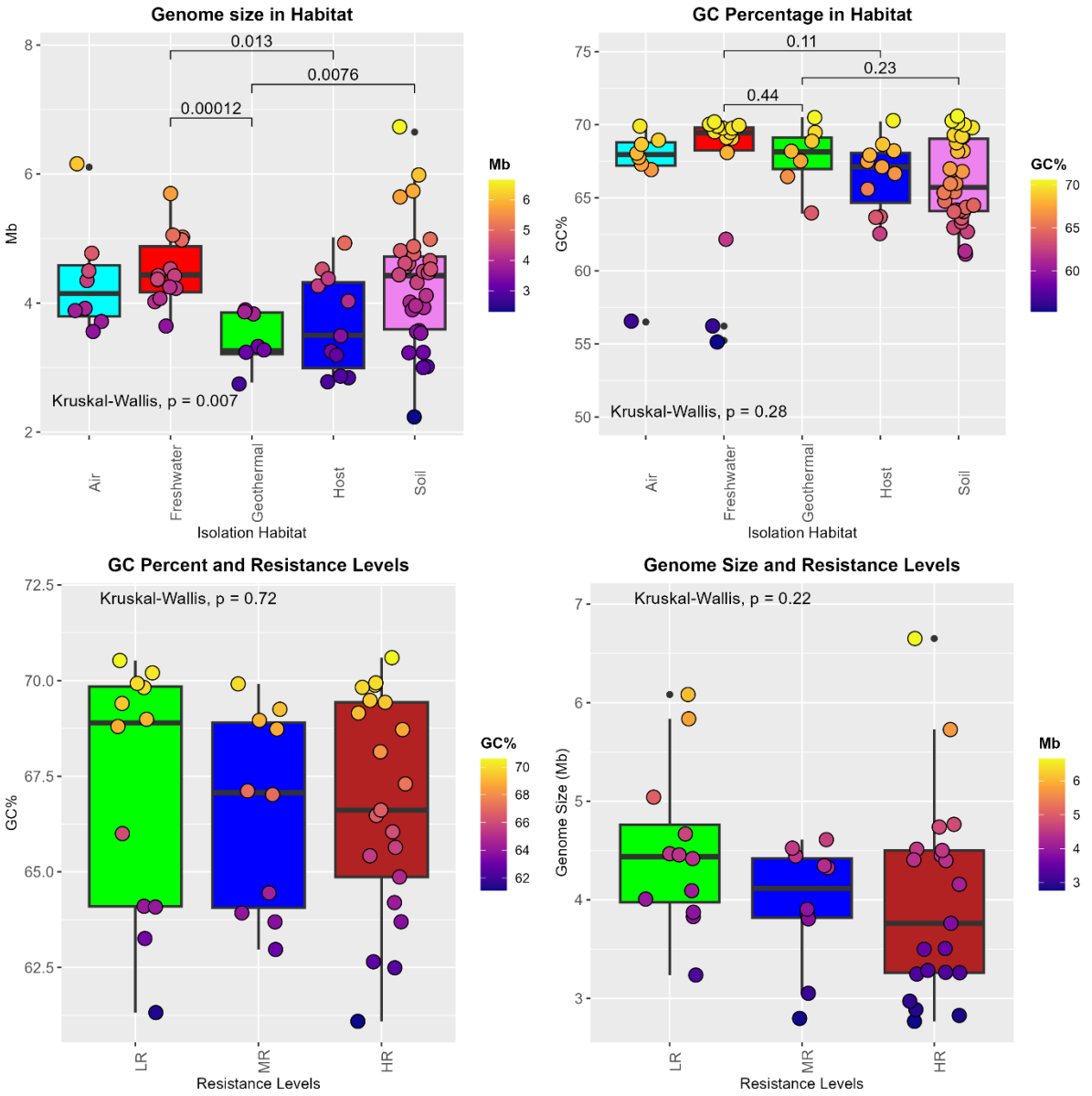

Figure 2-1). We used our dataset to investigate whether there is a correlation between the genomic features of *Deinococcus* members, their isolation habitats, and radiation resistance levels. A subset of our dataset with available radiation resistance data was used (Supplementary Data 2-1). Kruskal–Wallis test showed a significant correlation between genome size and isolation habitat, but no statistical correlation was detected between GC content and different habitats (

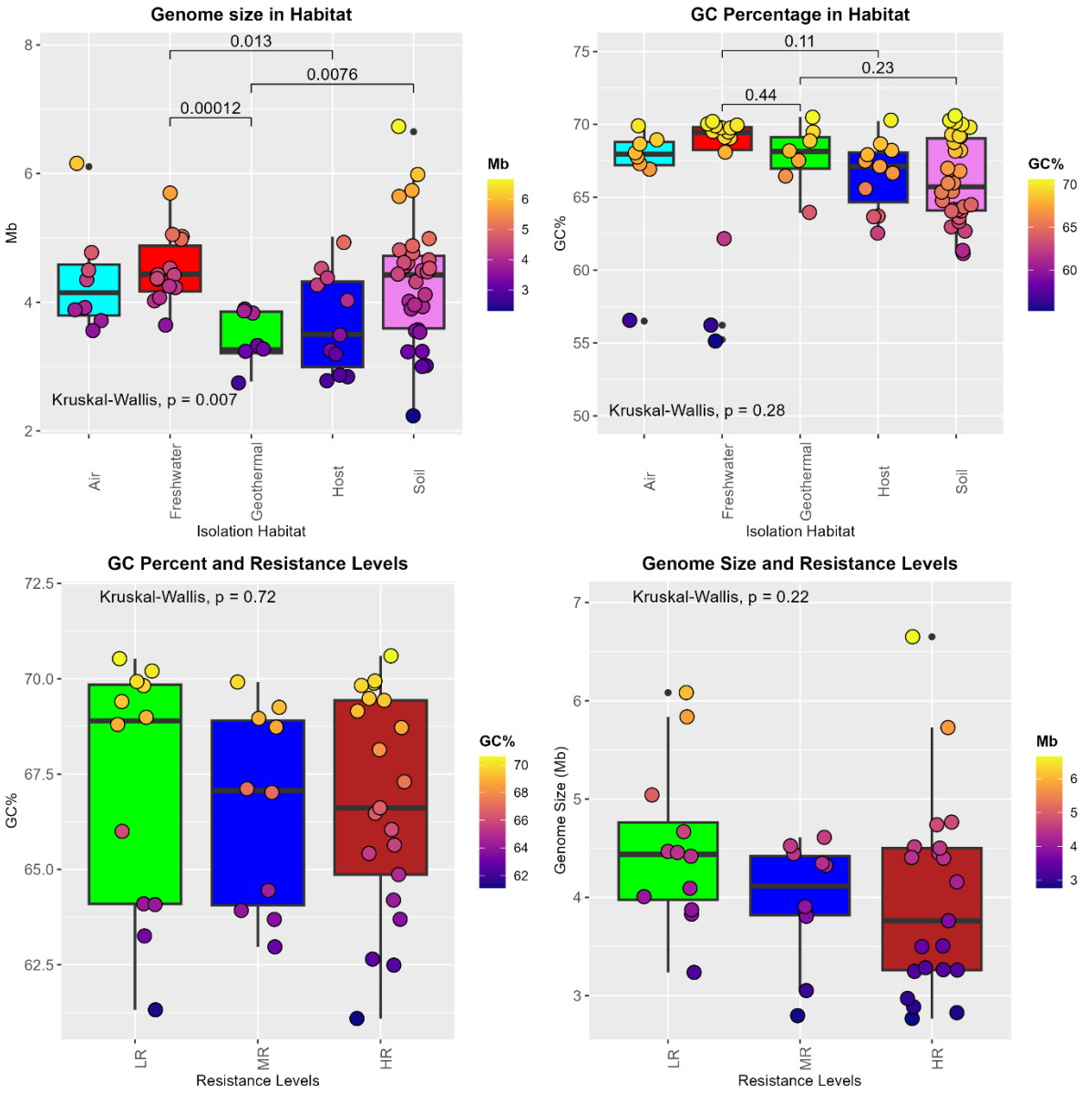
Genome size in Habitat
0.013
0.0076
0.00012
Mb
Kruskal-Wallis, p = 0.007
Air
Freshwater
Geothermal
Host
Soil
Isolation Habitat
GC Percentage in Habitat
0.11
0.23
0.44
GC%
Kruskal-Wallis, p = 0.28
GC Percent and Resistance Levels
Kruskal-Wallis, p = 0.72
LR
MR
HR
Resistance Levels
Genome Size and Resistance Levels
Kruskal-Wallis, p = 0.22
Genome Size (Mb)

Figure 2-1a

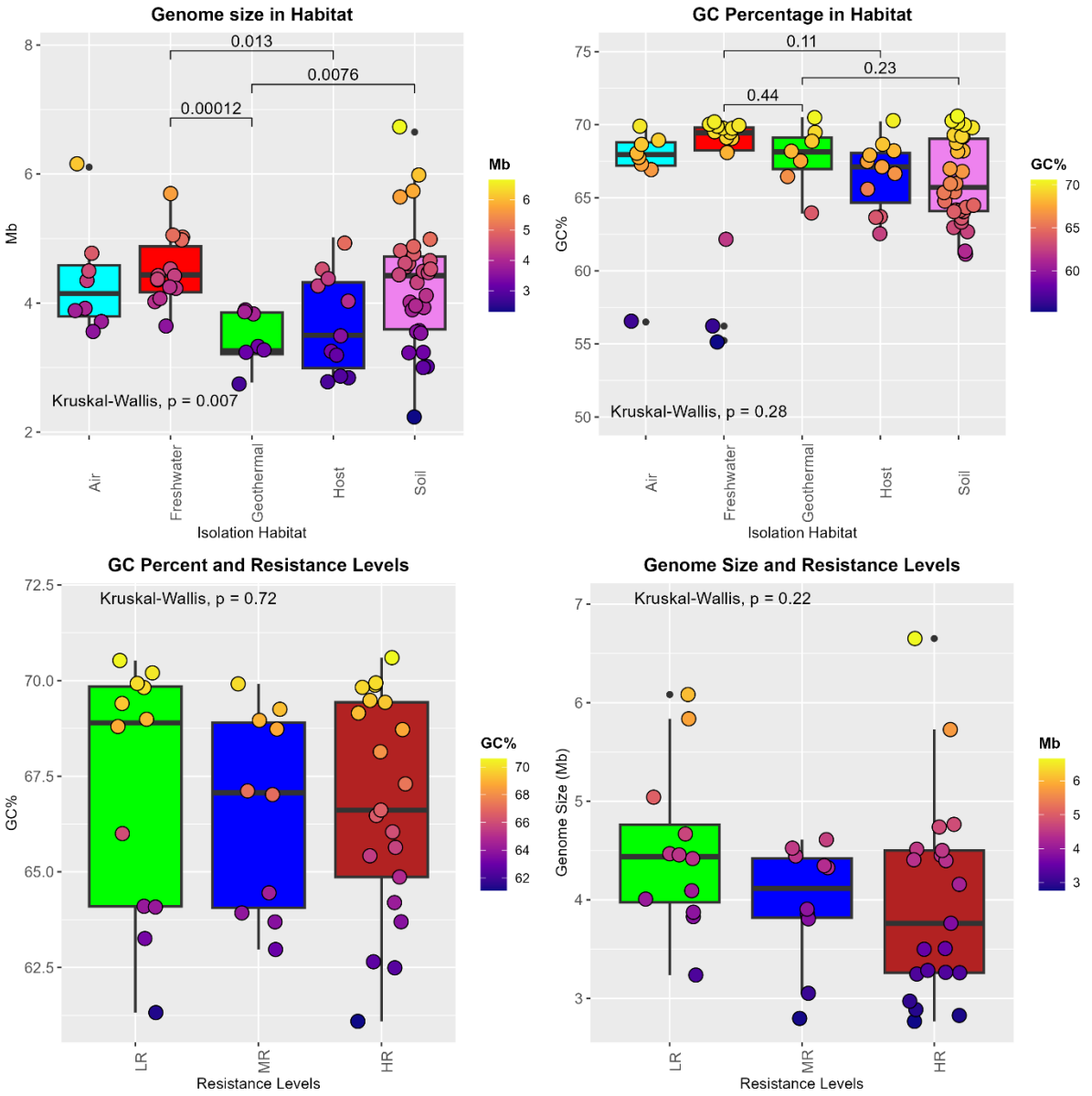

Figure 2-1). On the other hand, neither genome size nor GC content had any significant correlation with reported radiation resistance levels among *Deinococcus* members (

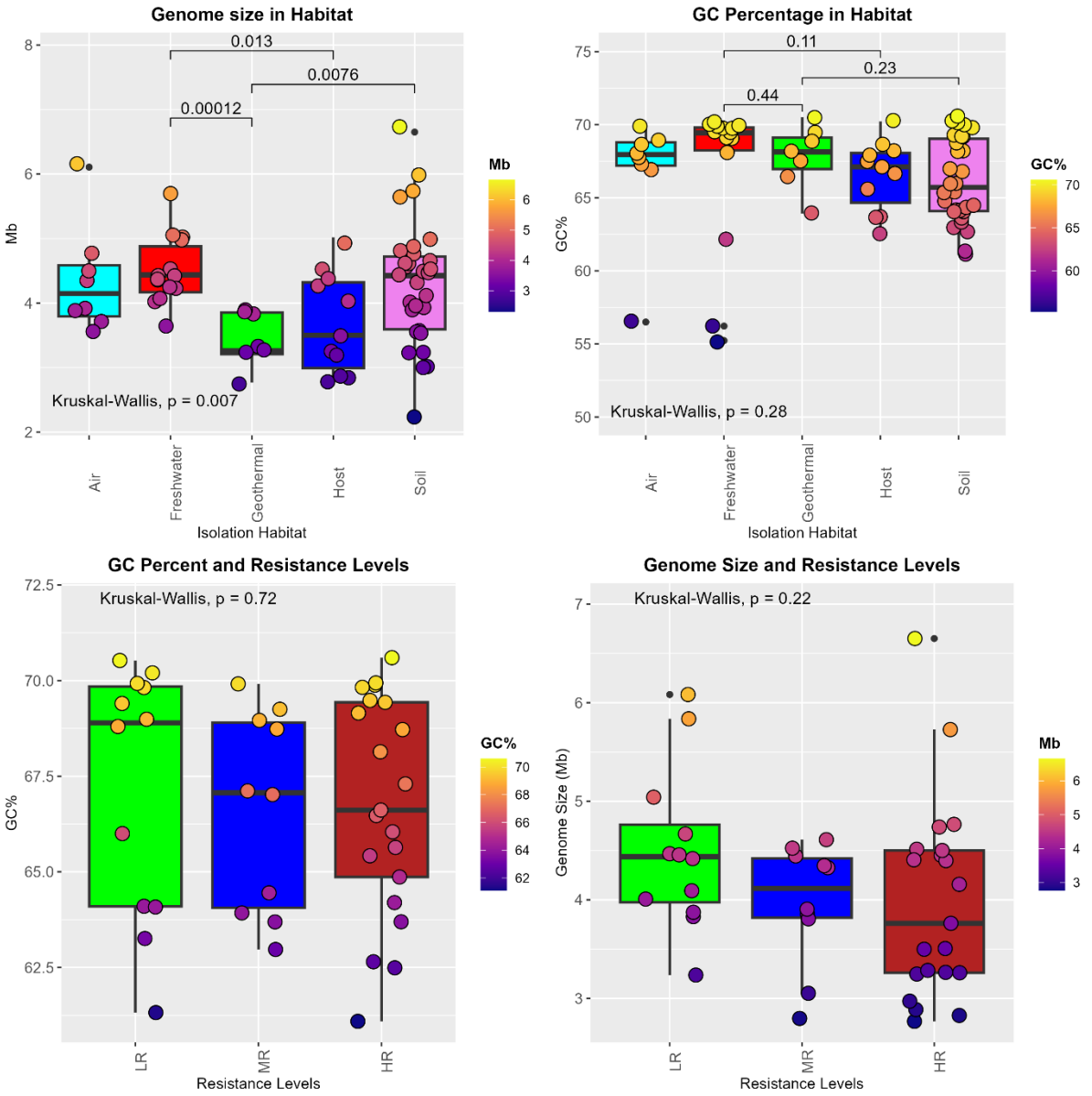

Figure 2-1b).

2.4.3 Environmental distribution of *Deinococcus* species

We categorised isolation habitats of *Deinococcus* species extracted into five categories—air, freshwater, geothermal water, host, and soil. (Supplementary Data S2-1 and

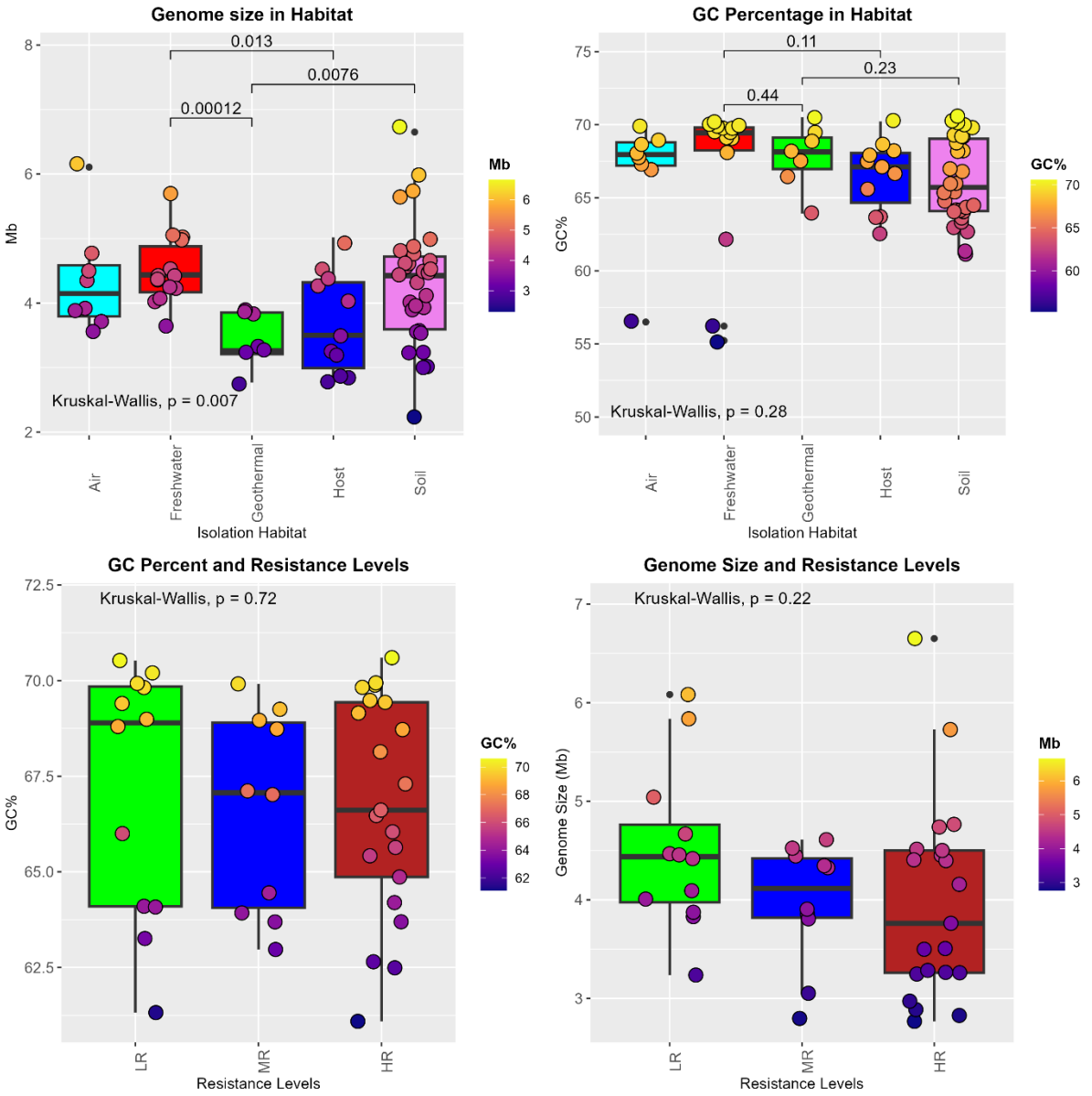

Figure 2-1). Most *Deinococcus* species included in this study were isolated from soil samples, followed by freshwater, host environments, air and geothermal pools.

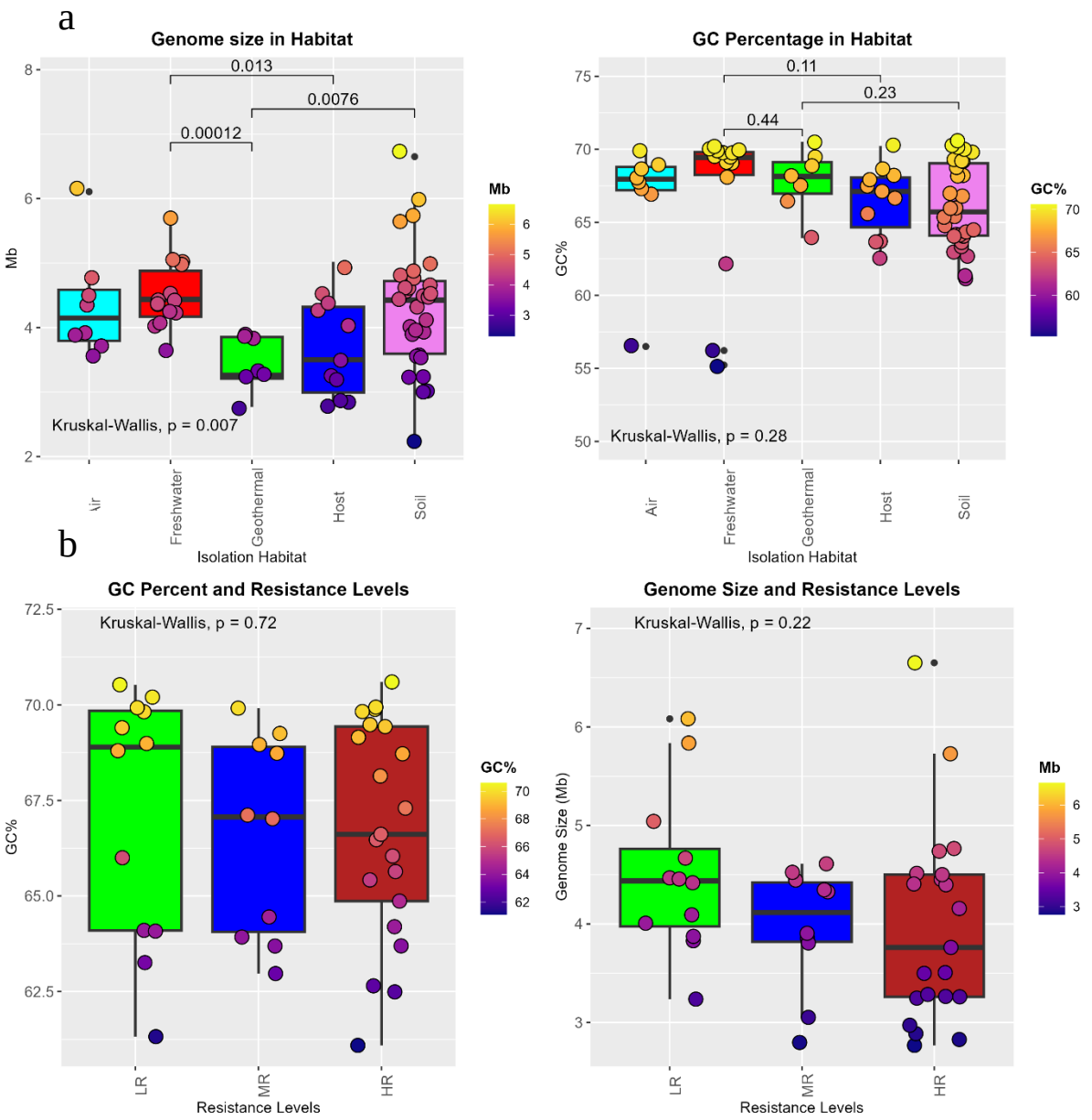

Figure 2-1 Variability of genomic features among *Deinococcus* species.

The genomic features of the *Deinococcus* species used in the pangenome analysis are shown in regard to their ecological data, such as isolation habitat and resistance levels a. Significant correlation between genomic features and isolation habitat is indicated with a p-value < 0.05, but no significant correlation was observed between GC content and isolation habitat. b. There is no correlation between genome size and GC content among *Deinococcus* with different IR resistance levels with all p-values > 0.05.

2.4.4 Pangenome characterisation:

We constructed the pangenome of the *Deinococcus* genus to highlight the diversity of its genomic features and answer the question about the genomic flexibility in the *Deinococcus* genus. We used 66 high-quality genomes, including 64 representative genomes of known and sequenced species in the public database and two from this

study. A total number of 266,658 genes were analysed, of which 256,229 (96.1%) were assigned to 12,055 orthogroups (OGs), and 10,429 were detected as unique genes (not clustered into any OGs). Using this dataset, we defined the pangenome of the *Deinococcus* genus and identified the core, accessory, and unique gene groups. The pangenome of the *Deinococcus* genus consists of 970 orthogroups (4.08%) as the core genome, 1080 orthogroups as soft core genome (4.21%), 11,085 orthogroups (47.26%) as accessory, and 10429 OGs (44.5%) as unique genes (

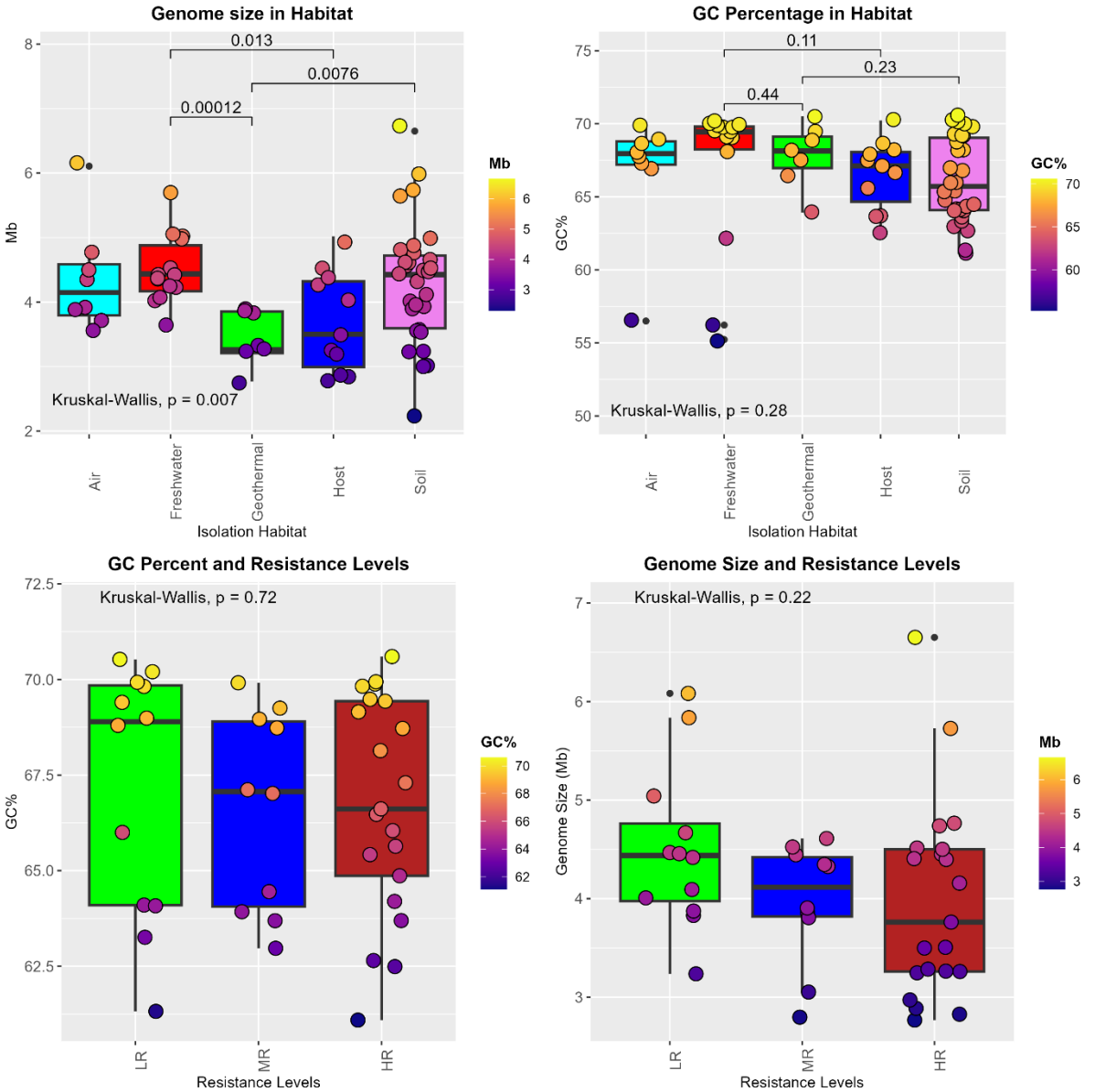

Figure 2-12 and Supplementary Data S2-2).

2.4.5 Pangenome openness

To determine whether currently available genomes are sufficiently comprehensive to define the genomic diversity of the *Deinococcus* genus, we used the constructed pangenome to explore shifts in the number of core, accessory, and unique genes upon the rarefaction curves (

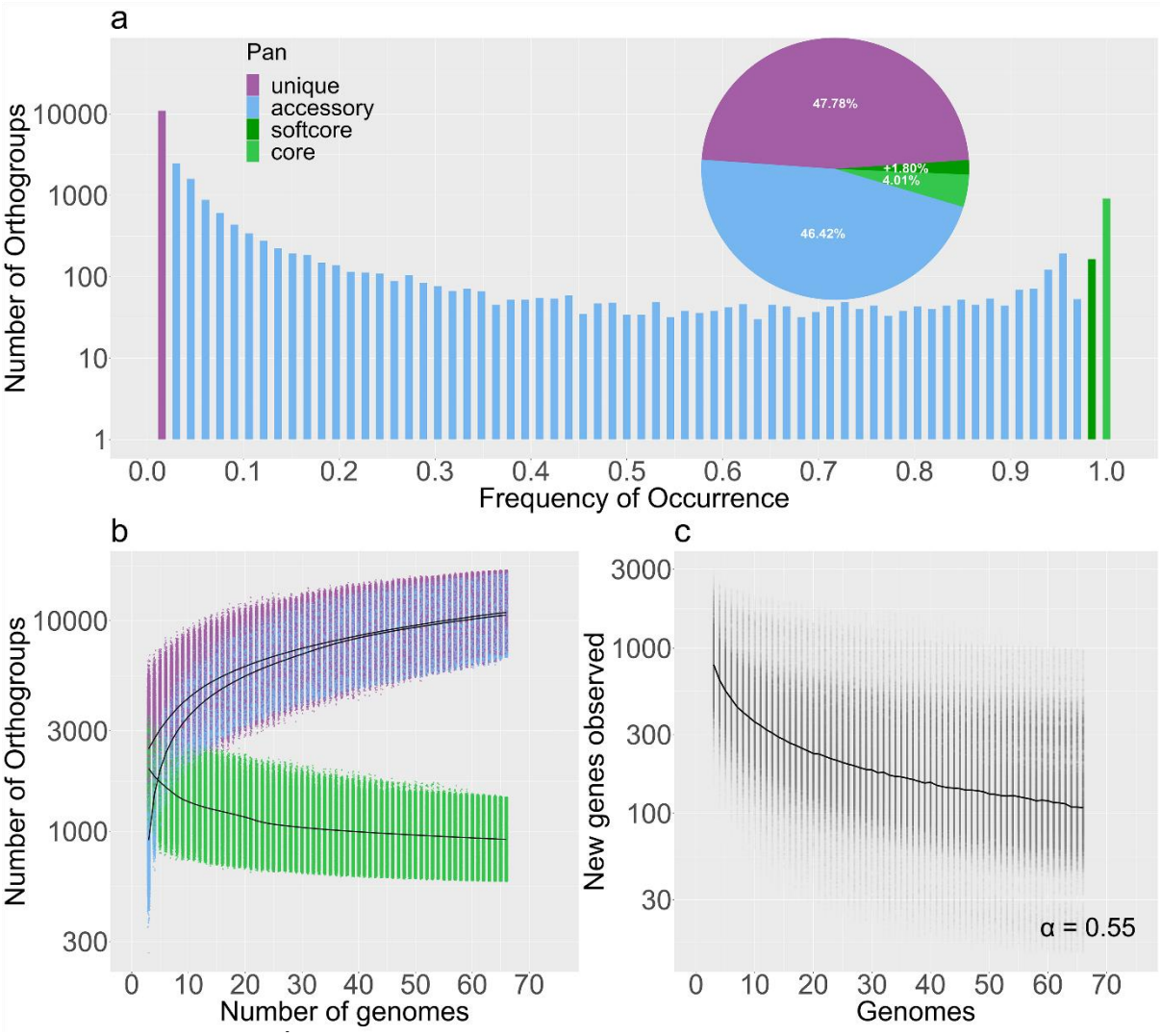

Figure 2-2a). Our analysis showed that the core genome stabilised after approximately 30 genomes and reached a plateau. However, it slightly decreased for every genome added. This pattern was different for the unique and accessory genes. The rarefaction curve for the accessory genome showed a sharp increase until 40 genomes were added, and then rates of gene detection slowed down, but the number of observed genes was

significantly high even when the last genome was added for the pangenome size (

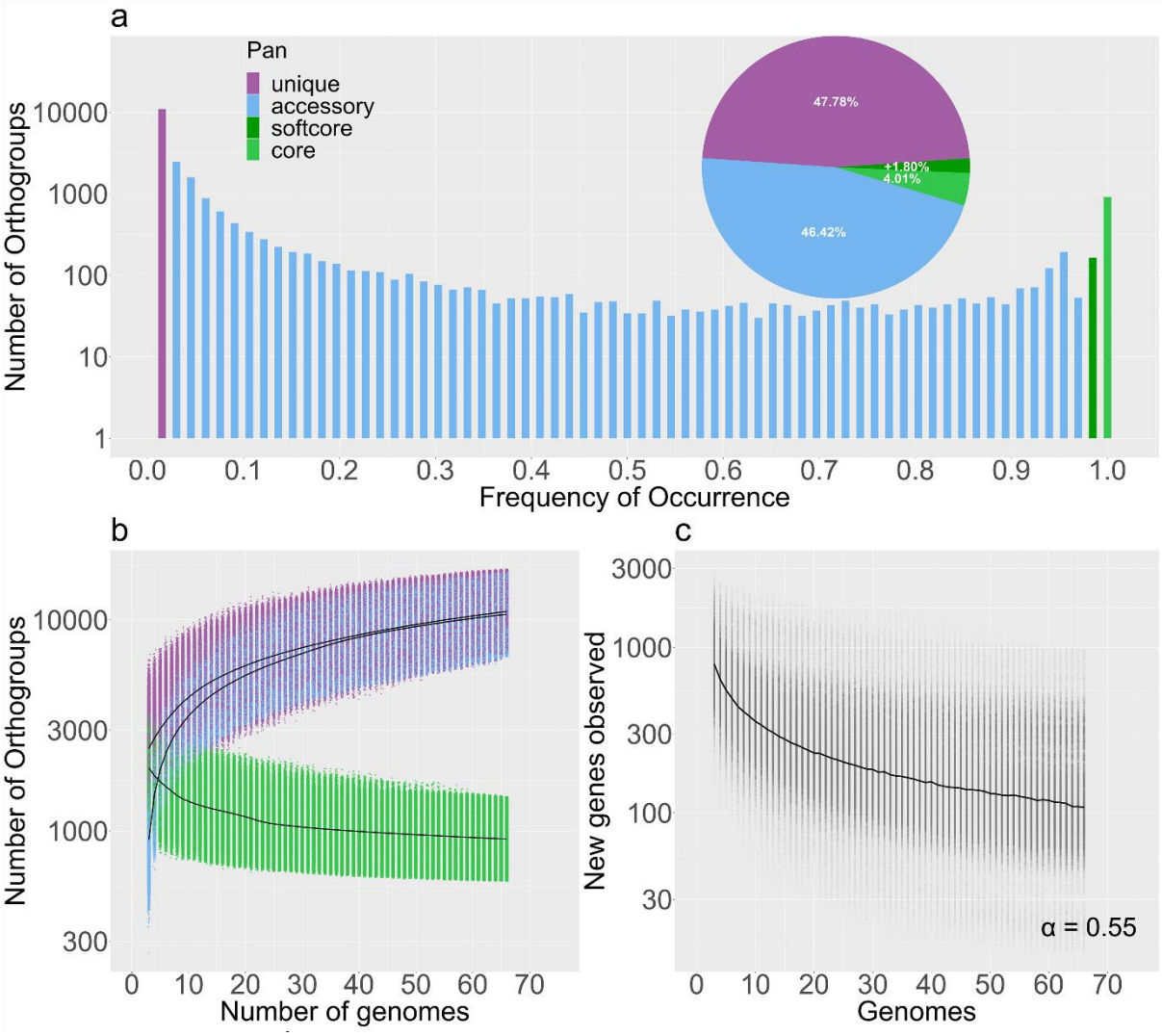

Figure 2-2b). Adding the final genome to the pangenome of the remaining 65 genomes leads to a loss of 1 orthogroup in the core genome and an increase of 47 orthogroups in

accessory and 313 genes in the unique genes (

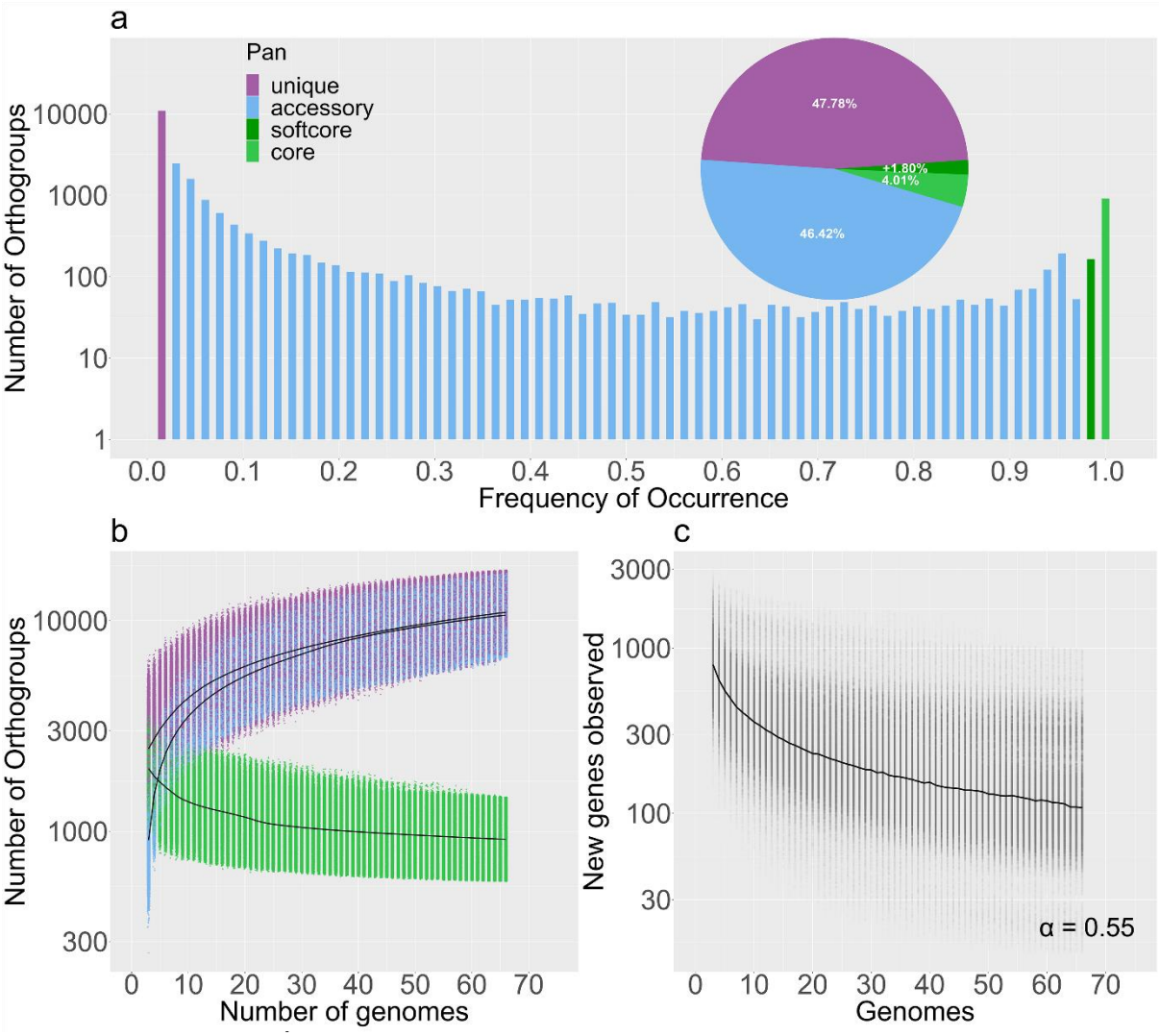

Figure 2-2b).

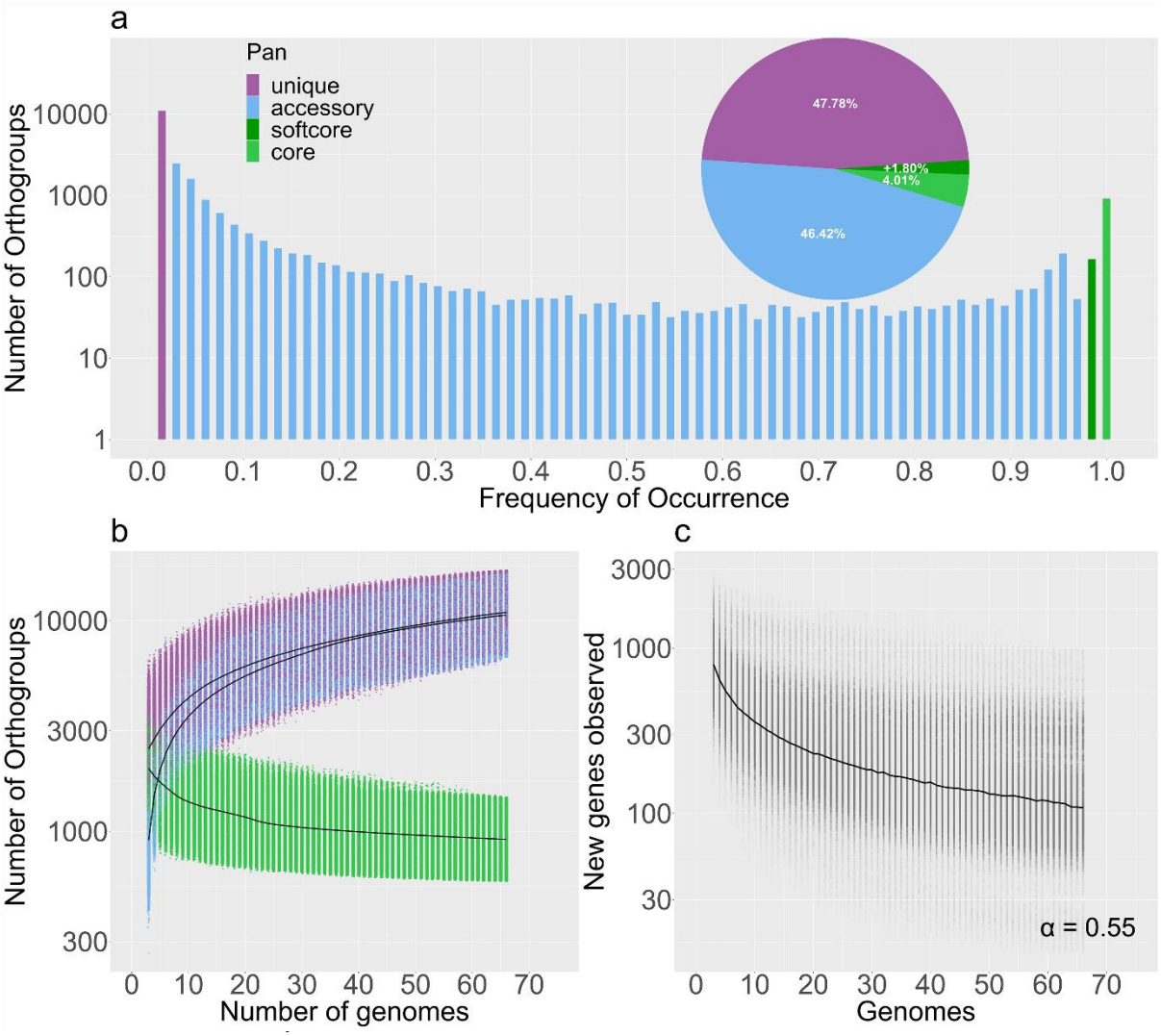

Figure 2-2. The *Deinococcus* pangenome.

a, The percentage and distribution of pangenome groups. The number of genomes determines the number of orthogroups. b, Simulation of genome sampling with 10,000 permutations shows a decrease in the core genome and an increase in the accessory and unique gene size when genome number increases. Coloured dots represent the number of classified homology groups: violet is unique genes, cyan shows accessory genes, light green softcore, and dark green core genes. The black line indicates the median of each category. c, The number of new genes decreases with adding more genomes. The curve is the least squares fit of the power law $n = k \times N^{-\alpha}$ to medians with $\alpha = 0.55$. d, The pangenome is mapped on the phylogenomic tree of the *Deinococcus* genus

To estimate whether the pangenome of the *Deinococcus* genus is open or closed, we used the Heaps law model approach with 10,000 permutations. We fitted Heaps' law by calculating the medians of the number of newly discovered unique genes per genome addition. The Heaps law was calculated using the formula $n = k \times N^{-\alpha}$ where n is the newly observed genes, N is the total number of genomes, α is the decay rate, and k is the fitting parameter. The calculated decay rate (α) for the *Deinococcus* pangenome is 0.44, meaning the pangenome follows Heaps' Law and is open. In other words, currently, available genes may be able to demonstrate the core genome of the *Deinococcaceae* family but are not comprehensive enough to include the whole diversity of the *Deinococcus* genus, and there will be new gene clusters with the

addition of new genomes (

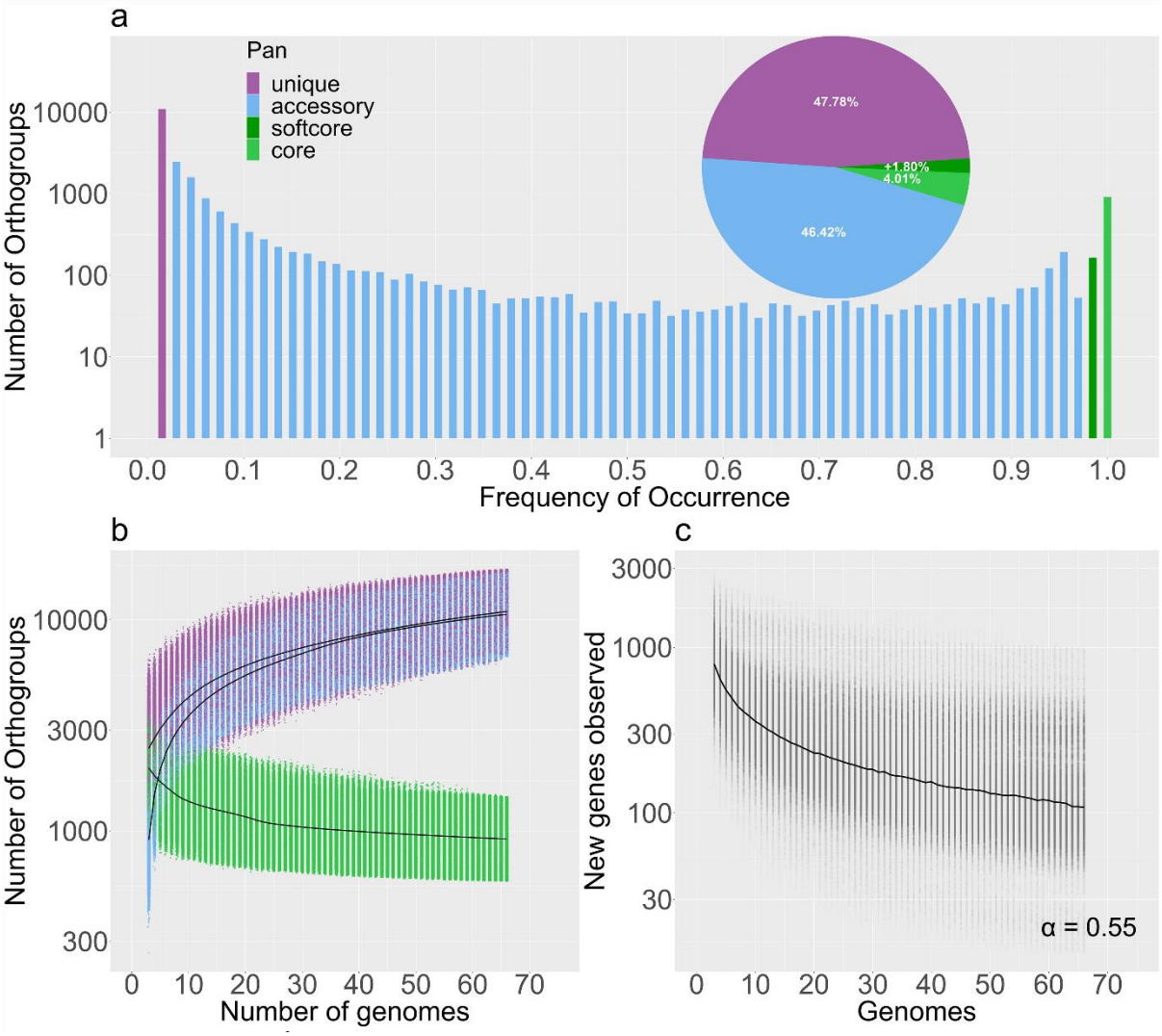

Figure 2-2c).

2.4.6 Phylogenomics based on single-copy orthogroups

To demonstrate the evolutionary relationships between members of the *Deinococcus* genus, we reconstructed a phylogenomic tree using a concatenated supermatrix from 476 single-copy orthogroups shared between all genomes. *Deinobacterium chartae*, *Truepera radiovictrix*, and *Thermus thermophilus* were used as outgroup species to root the tree. Next, we mapped the composition of each pangenome group for each *Deinococcus* species on the phylogenomic tree (Figure 2-3). Our maximum likelihood tree calculated with 1000 ultrafast bootstraps showed a robust phylogenomic relationship between species with most branch support values > 95. The phylogenomic

analysis shows that the *Deinococcus* genus is a paraphyletic group. A paraphyletic group is defined as a taxonomic group that includes the common ancestor but does not contain all descendants from this ancestor (Sturmbauer, 2013).

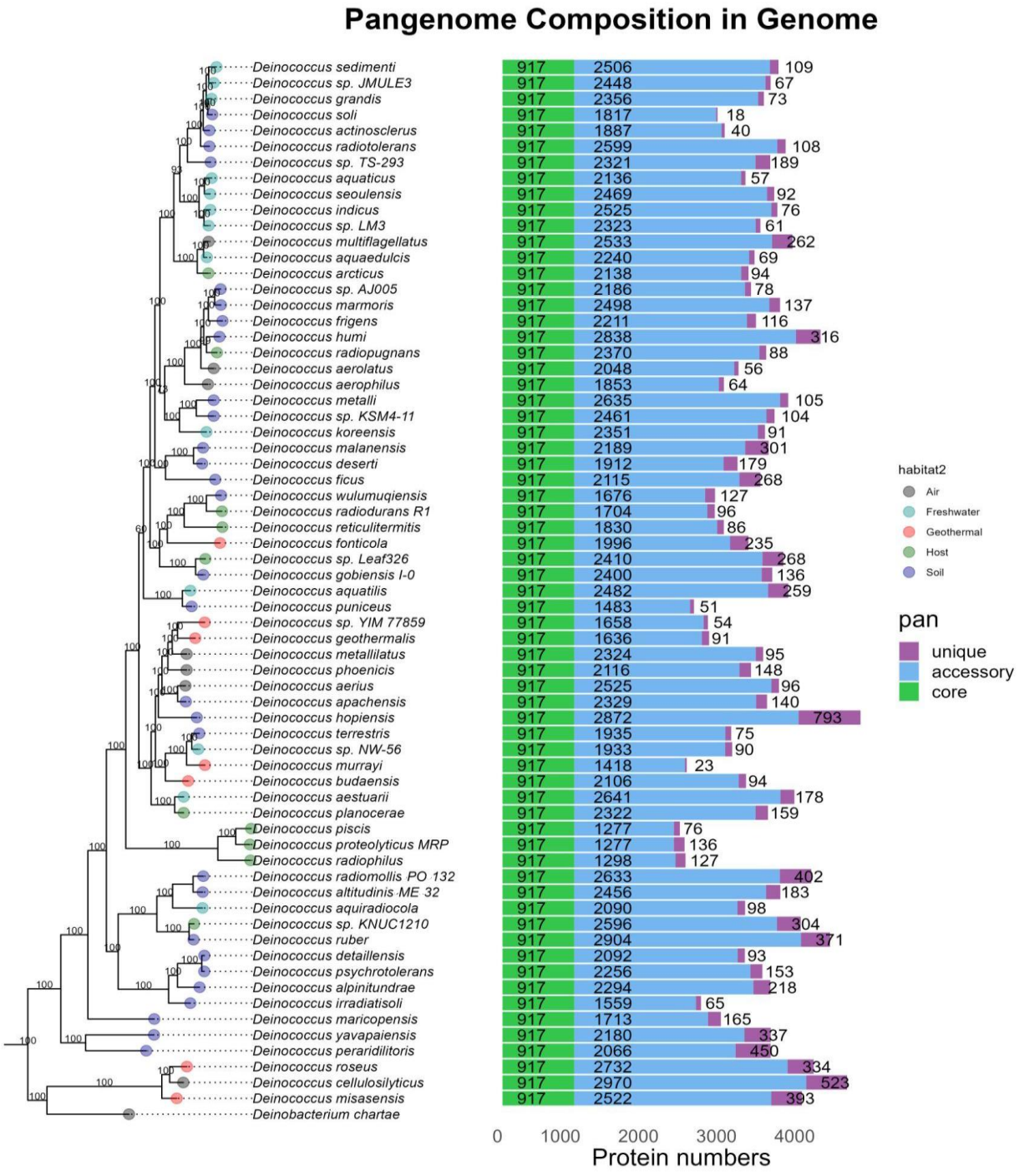

Figure 2-3. Phylogenomic tree and pangenome composition of *Deinococcaceae*.

The maximum likelihood tree of the *Deinococcaceae* family members was reconstructed using 476 concatenated single-copy orthogroups. *Thermus thermophilus, Truepera radiovictrix, and Deinobacterium chartae* are used as

outgroups. *Thermus thermophilus* and *Truepera radiovictrix* were removed from the tree for a clear visualisation in shorter branches. Bar plots show the composition of proteins belonging to different pangenome groups, and tip colours show isolated habitats.

2.4.7 Functional annotation

To enable the biological interpretation of the pangenome groups and demonstrate the enrichment of functional categories, we annotated the pangenome orthogroups on COG categories using EggNOG mapper v2.1.12. Each *Deinococcus* species has 917 orthogroups as core, an average of 2192 orthogroups as accessory, and 165 OGs as unique genes, with a high variation in accessory and unique genes with standard deviations of 403 and 136 (Fig1a), respectively. From 266,658 analysed sequences in the orthology analysis performed by OrthoFinder, 218,973 genes (about 82%) were assigned to one or more functions or unknown functions. Of 12,055 detected orthogroups, only 8,674 orthogroups (72%) had annotations and 3381 (28%) orthogroups did not match any function in the EggNOG database. Among them, 89% of the core genome, 50% of the accessory genome, and only 9% of unique genes were annotated (Figure 2-4b).

The abundance of each pangenome group in COG categories was used to identify functional differences in the enrichment of the core, accessory, and unique genes (Figure 2-4b). Orthogrops with unknown functions comprised the highest abundance throughout the whole pangenome, at 23.6%. Since the orthogroups that belong to the core genome are mostly conserved genes, they showed the lowest rates of unknown functions. As expected, the frequency of occurrence of OGs in the accessory genome was much higher than core and unique genes in most cases, while core and unique genes had a close abundance in most categories except for translation, ribosomal structure and biogenesis, and signal transduction mechanisms. Orthogroups related to translation, ribosomal structure and biogenesis showed the highest abundance in the core genome with 2% of the pangenome, slightly higher than accessory genes at 1.9% and much higher than unique genes at 0.04% of the pangenome.

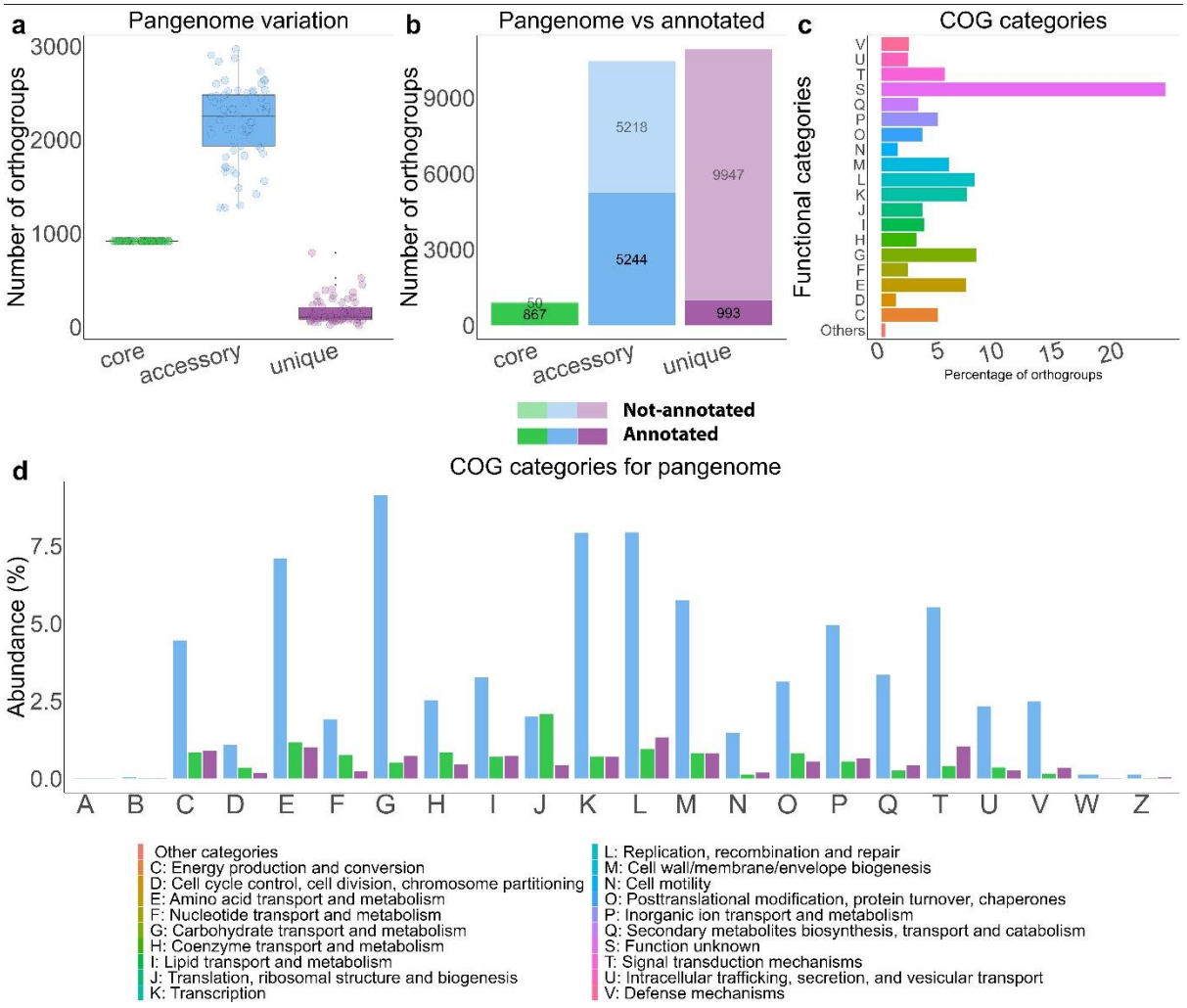

Figure 2-4. Functional annotation based on COG categories.

a. shows the variation of core, accessory and unique genes within *Deinococcus* members. b. The proportion of total orthogroups numbers in pangenome versus the number of annotated orthogroups with COG categories. The lighter colour represents unannotated portion of the pangenome c. Percentage of enrichment of each functional category in the pangenome. d. Enrichment of functional categories between pangenome groups.

Genes related to carbohydrate transport and metabolism (G category), enabling the chemoorganotrophic lifestyle of *Deinococcus*, showed the highest abundance within the pangenome at 7.90%. OGs belonging to the accessory genome comprise most of it, 614

OGs out of 697. The most diverse gene families in this category are epimerase, glycosyl hydrolase, and binding-protein-dependent transport systems. The second enriched category within the *Deinococcus* pangenome is replication, recombination, and repair mechanisms (L category), which account for about 7.79%. Genes related to different families of transposases, which facilitate DNA integration through transposition, were the most abundant orthogroups within this group. In this category, unique genes had a higher frequency than core genes (Figure 2-4, Table 2-2).

Orthogroups related to transcription (K), with, are the third most abundant category, and accessory genomes constitute most of them. Regulatory proteins, such as the LacI–family transcription factors, which regulate carbohydrate utilisation genes, sequence-specific DNA binding proteins, and other sugar regulatory proteins are the most abundant orthogroups in the transcription category. Amino acid transport and metabolism, with 7%, was the next abundant orthogroup and was predominantly present in the accessory genome. Highly abundant gene families within the amino acid transport and metabolism system (E) ABC-type dipeptide oligopeptide nickel transport systems, extracellular solute-binding proteins, dipeptide transporters, LysE type translocator, and ABC transporter complex of CysAWTP genes (Supplementary Data S2-3, Table 2-2).

Table 2-2. Number and abundance of OGs in each COG functional category

COG Category Description	OGs in Accessory genome	OGs in core genome	OGs in unique genome	Relative Abundance of total OGs
J: Translation, ribosomal structure and biogenesis	134	140	29	3.43
A: RNA processing and modification	1	0	0	0.01
K: Transcription	533	48	47	7.12
L: Replication, recombination and repair	534	64	89	7.79
B: Chromatin structure and dynamics	3	1	1	0.05
D: Cell cycle control, cell division, chromosome partitioning	73	23	12	1.22
V: Defence mechanisms	167	10	23	2.26
T: Signal transduction mechanisms	371	27	70	5.30
M: Cell wall/membrane/envelope biogenesis	386	55	54	5.61
N: Cell motility	99	8	13	1.36
Z: Cytoskeleton	9	1	3	0.14
W: Extracellular structures	9	0	1	0.11
U: Intracellular trafficking, secretion, and vesicular transport	156	24	18	2.24
O: Posttranslational modification, protein turnover, chaperones	211	54	37	3.42
C: Energy production and conversion	300	57	60	4.73
G: Carbohydrate transport and metabolism	614	34	49	7.90
E: Amino acid transport and metabolism	477	78	68	7.06
F: Nucleotide transport and metabolism	128	51	15	2.20
H: Coenzyme transport and metabolism	170	56	30	2.90
I: Lipid transport and metabolism	220	47	49	3.58
P: Inorganic ion transport and metabolism	333	37	44	4.69
Q: Secondary metabolites biosynthesis, transport and catabolism	225	17	29	3.074
S: Function unknown	1,617	146	326	23.69
Total	6,770	978	1,067	8,81

2.4.8 Genes related to ionising radiation resistance

To investigate the presence or absence patterns of genes with a known role in radiation resistance in *D. radiodurans*, we collected a set of literature-vetted genes associated with RDR regulon. These genes are upregulated after exposure to ionising radiation and desiccation in *D. radiodurans* (Table 2-3). Of 37 OGs belonging to the RDR regulon in *D. radiodurans*, 15 OGs comprised the core genome, and 22 OGs were identified as the accessory genes (Figure 2-5)**.** Most of the RDR regulon genes within the core genome were related to replication, recombination and repair categories (L) (Figure 2-5). Genes in this category include (single-stranded DNA-binding protein - *Ssb*, DNA ligase (NAD+) - *ligA*, DNA topoisomerase IV subunit B, exodeoxyribonuclease V, Holliday junction DNA, helicase *ruvB*, DNA gyrase subunit A - *gyrA*, DNA mismatch repair protein - *mutS*, ATP-dependent DNA helicase - *recG*, excinuclease ABC subunit B). Other core RDR regulon OGs include one gene in the category carbohydrate transport and metabolism (transketolase - *tkt*), two genes in translation, ribosomal structure and biogenesis (leucine-tRNA ligase - *leuS*, acetyltransferase - *ddrN*), one gene in transcription (transcriptional regulator - *ddrO*), and one gene in the inorganic ion transport and metabolism category (thiosulfate sulfurtransferase - Rhodanese).

Accessory genes within the RDR regulon include two families of transposase, IS3 and Tn3. IS3 transposase was present in most *Deinococcus* members except for 10. However, Tn3 family transposase was present in 21 out of 66 genomes. Different families of transporters, such as peptide ABC transporter substrate-binding protein, peptide ABC transporter, and multidrug resistance transporters, were observed in the accessory RDR regulon orthogroups.

Table 2-3. List of genes belonging to the RDR regulon in *D. radiodurans*

locus_tag	genes	description
DRO_1288	DR_1297	ABC transporter
DRO_1551	peptide transporter	peptide ABC transporter
DRO_2416	argB	acetylglutamate kinase -argB
DRO_2415	ddrN	acetyltransferase - ddrN
DRO_1894	recG	ATP-dependent DNA helicase - RecG
DRO_2308	operon recA	damage-inducible protein - recA
DRO_A0273	CccA	cytochrome C6 - CccA
DRO_1891	gyrA	DNA gyrase subunit A - gyrA
DRO_1552	helD	DNA helicase UvrD
DRO_2042	ligA	DNA ligase (NAD(+)) LigA
DRO_1033	mutS	DNA mismatch repair protein MutS
DRO_A0342	pprA	DNA repair protein
DRO_0899	gyrB	DNA topoisomerase IV subunit B
DRO_2249	uvrB	excinuclease ABC subunit B
DRO_1880	recD	exodeoxyribonuclease V
DRO_2229	GNAT family N-acetyltransferase	GNAT family N-acetyltransferase
DRO_0139	hflX	GTPase HflX
DRO_0596	ruvB	Holliday junction DNA helicase RuvB
DRO_C0021	DR_C0023	hypothetical protein - DR_C0023
DRO_0003	ddrC	hypothetical protein - ddrC
DRO_0681	DR_0686	hypothetical protein - DR_0686
DRO_1140	DR1143	hypothetical protein - DR)1143
DRO_2144	DR_2174	hypothetical protein - DR_2174
DRO_A0167	DR_A0166	hypothetical protein - DR_A0166
DRO_A0168	DR_A0167	hypothetical protein - DR_A0167
DRO_0219	ddrF	hypothetical protein - DR0219
DRO_2145	leuS	leucine-tRNA ligase - leuS
DRO_0323	ddrD	multidrug DMT transporter - ddrD
DRO_0070	ddrB	single-stranded DNA-binding protein -ddrB
DRO_0099	ssb	single-stranded DNA-binding protein - ssb
DRO_0421	ddrA	single-stranded DNA-binding protein - ddrA
DRO_0559	MalE	substrate-binding protein - MalE
DRO_0217	Rhodanese-like protein	thiosulfate sulfurtransferase - Rhodanese
DRO_2545	ddrO	transcriptional regulator - ddrO
DRO_2230	tkt	transketolase - tkt
DRO_C0017	Tn3 Family	transposase Tn3 family
DRO_1287	IS5_3	transposase IS5_3

These genes were reported in different transcriptomicsstudies, but recently re-evaluated by (Eugénie et al., 2021) and the number of genes in the RDR regulon were extended to 37.

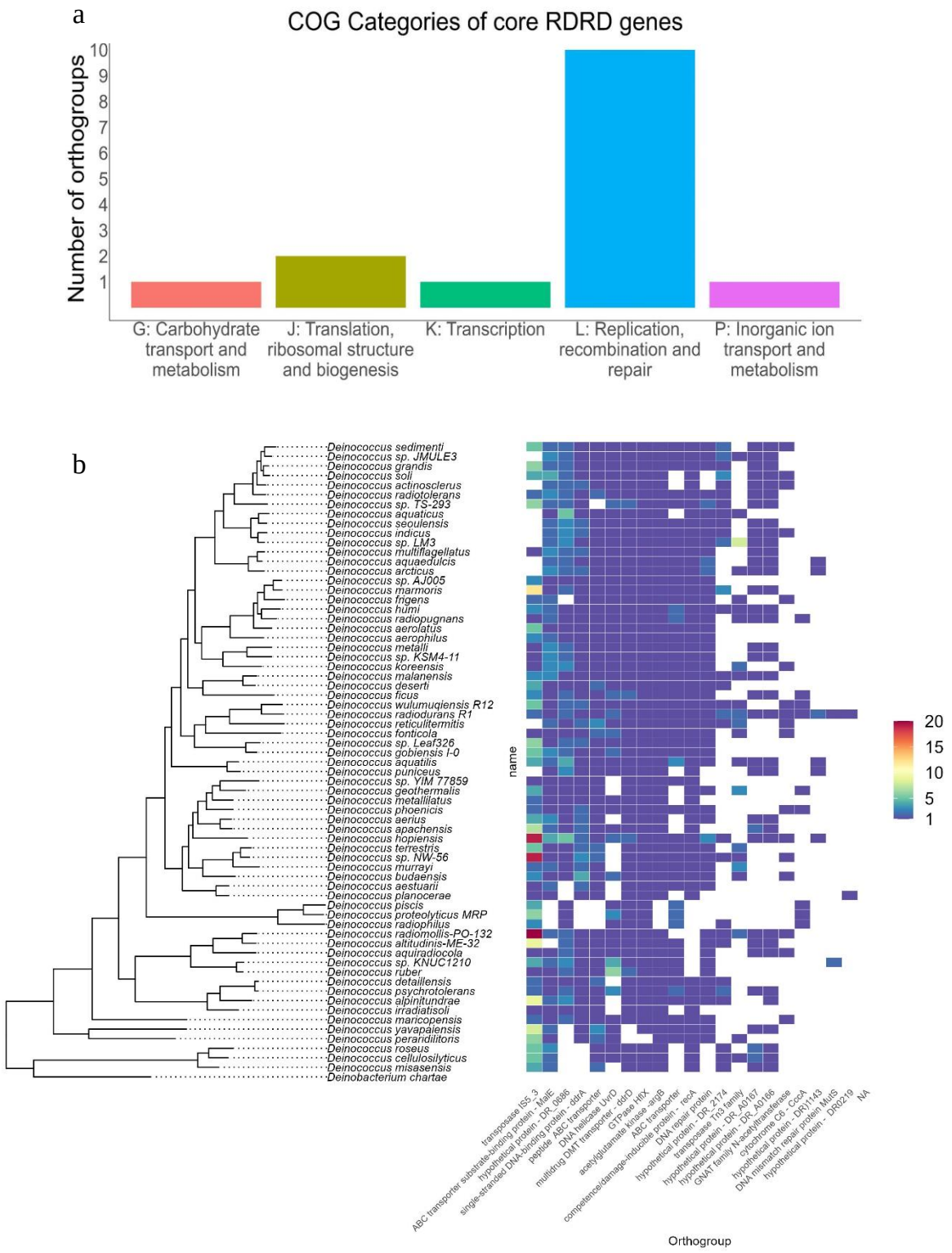

Figure 2-5. Enrichment of *D. radiodurans* RDR regulon among the pangenome.

Functional profiling of genes belonging to RDR regulon present in the pangenome. **a**. Bar plot showing the number of RDR regulon genes belonging to the core orthogroups. **b** Heatmap depicting the presence and absence patterns of *D. radiodurans* RDR regulon genes within the *Deinococcus* accessory genes.

2.4.9 Functional ecology of *Deinococcus* genomes:

To enhance our understanding of the poorly understood and under-studied ecology of the *Deinococcus* genus, we used the METABOLIC tool to link metabolic pathways of available *Deinococcus* genomes to their isolation habitat. The METABOLIC tool benefits from a broad range of metabolic pathways by integrating KEGG, TIGRfam, Pfam, and custom HMM profiles and generates a standardised characterisation of metabolic pathway predictions, metabolite exchanges, microbial interactions, and microbial contributions to biogeochemical cycling (Zhou et al., 2022).

Our results suggest that *Deinococcus* species have a chemoorganotrophic lifestyle in all habitats. All species possessed a minimal pentose phosphate pathway, which lacks key steps for synthesising ribose 5-phosphate and carbon fixation (Figure 2-6). Only a few amino acid biosynthesis pathways were present in all species, including ornithine, threonine, shikimate, methionine, valine/isoleucine, leucine, arginine, histidine, and proline. Lysine biosynthesis was missing in most of the host-associated *Deinococcus* species, *D. piscis,* isolated from marine fish. Moreover, de-novo nucleotide metabolism, including purine and pyrimidine, is conserved among all *Deinococcus* species and indicates that *Deinococcus* members have a free-living lifestyle of *Deinococcus* species (Figure 2-6).

Additionally, many cofactors and vitamins that promote bacterial growth (Mu et al., 2020), such as heme, biotin, thiamin, ubiquinone, VB12, and VB6, were absent in most genomes. Notably, all the genomes had pathways for fatty acid biosynthesis and degradation.

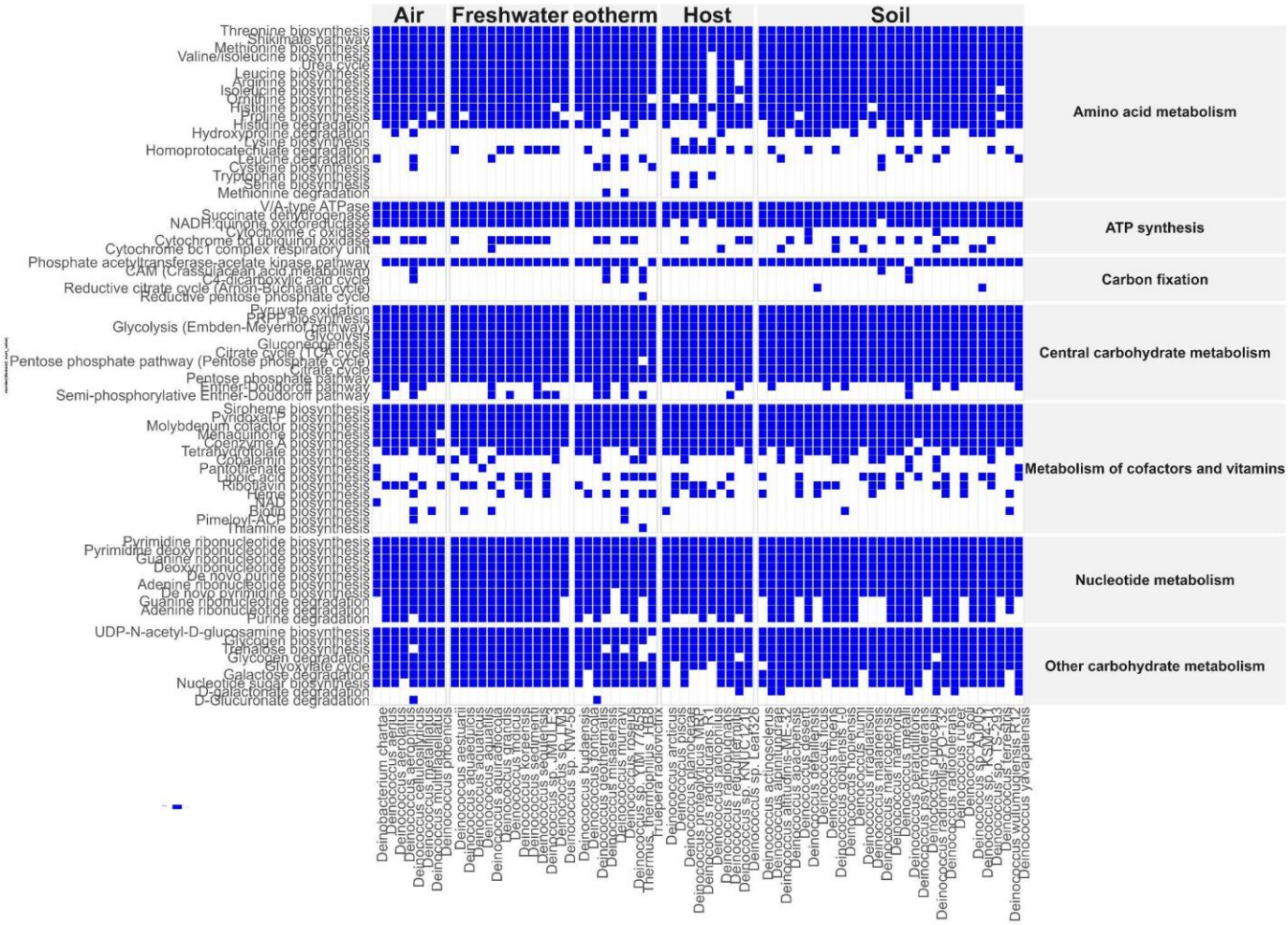

Figure 2-6. Heatmap of KEGG metabolic pathways for *Deinococcus* genomes

The heatmap represents metabolic pathways based on the presence or absence of genes determined by the METABOLIC toolkit v4.0. White represents the absence of a pathway, and blue represents the presence of a pathway.

The results from the METABOLIC custom database provided another dimension of metabolic diversity among *Deinococcus* species. Iron cycling metabolisms were sporadic and observed in some species. Iron reduction was rare and only observed in *D. phoenicis.* However, iron oxidation was more abundant and observed in some species isolated from diverse habitats except geothermal pools (Figure 2-7). Manganese oxidation and conversion of Mn(II) to Mn(IV) were rare and only observed in *D. ficus* isolated from rhizosphere soil. Methane oxidation was also sporadic and present only in 6 species.

Nitrogen cycling metabolisms, especially nitrite reduction to ammonia, showed a pattern in different environments. While most organisms isolated from freshwater and soil had essential genes for reducing nitrite to ammonia, most host-associated species did not have this pathway. Except for sulfate reduction, sulfur cycling functions were rare and observed in a few species. Sulfate reduction was highly conserved in most environments except for host-associated species. Dissimilatory sulfur metabolisms were mostly conserved but only missing in three host-associated species. The same three species, however, had assimilatory sulfur metabolisms instead of dissimilatory sulfur metabolisms, which was absent in all but two other *Deinococcus* species, *D. fonticola* and *D. ficus* (Figure 2-7).

Oxidative phosphorylation and production of ATP in *Deinococcus* members happens through complexes I, II, IV, and V. Complex III oxidative phosphorylation was absent in all species. Complex II and V were conserved among all *Deinococcus* species, but other complexes were missing in some. Cytochrome c oxidase caa3-type was highly conserved in complex V, while cytochrome quinone (bd-type) was observed only in some. Interestingly, type cbb3 cytochrome c oxidase was observed in only one species, *D. radiomollis,* which is sensitive to gamma radiation. Even though *D. geothermalis* and *D. murray* are known as slightly thermophilic species, they did not have the marker for hyper-themophily reverse gyrase, and the only species in our analysis with this marker was *T. thermophilus* (Figure 2-7).

Figure 2-7. Heatmap of METABOLIC data for the *Deinococcaceae*

Presence/absence heatmap of metabolic pathways in *Deinococcus* across isolation habitats. A heatmap classifying 19 categories and 54 rows of reactions demonstrates the distribution of metabolic pathways in *Deinococcus* species based on their isolation habitats. This heatmap provides insights into the variation of metabolic pathways within the *Deinococcus* genus with respect to their isolation habitats.

2.5 Discussion

In the last decade, the increasing amount of publicly available genomic data has enabled the study of the evolutionary histories of many organisms across the Tree of Life. Members of the *Deinococcus* genus are highly versatile and ubiquitous and have been isolated from diverse habitats. Despite numerous publications on the IR-resistance properties of some of its members, the ecology of this diverse genus has remained under-explored. Nonetheless, hundreds *Deinococcus* strains have been isolated from diverse environments. From the soil of hot and cold deserts and alpine environments to the air dust of the high atmosphere, freshwater environments, geothermal pools, and the guts of marine fishes as well as terrestrial fishes. But the has not been any coimprehesive study to collectively investigate the ecology and diversity of available genomes. In this study, we construct the pangenome of the *Deinococcus* genus to approximate its genomic repertoire. Then, we use the functional potentials of the *Deinococcus* species with their isolation habitats to provide insights into their ecology.

2.5.1 Genomic features and ecological distribution

We generated a complete genome sequence for two radiation-sensitive *Deinococcus* species to increase the diversity of under-represented genomes for radiosensitive species in the public database. Next, we constructed a dataset comprising 66 high-quality *Deinococcus* genomes and their metadata, including genome size and isolation habitats. Our results demonstrated that genome size and GC content are highly variable among *Deinococcus* species and span between 2.7 to 6.65 Mbp for genome size and 55% to 70% for GC content **(**

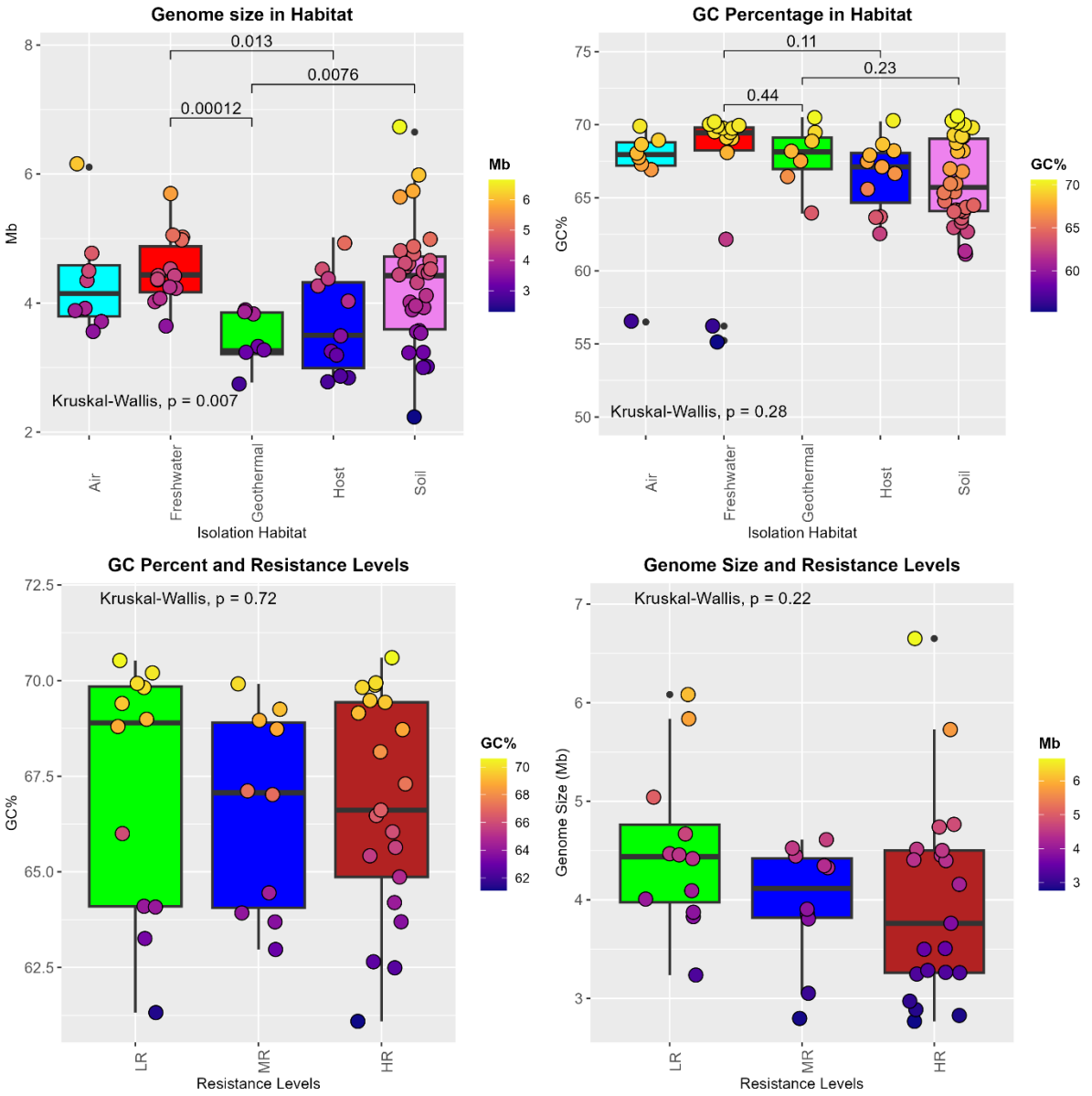

Figure 2-1**)**. We also observed that the genome size is significantly correlated to the isolation habitat, with host-associated *Deinococcus* and species isolated from geothermal pool and air having a smaller genome size than species isolated from soil **(**

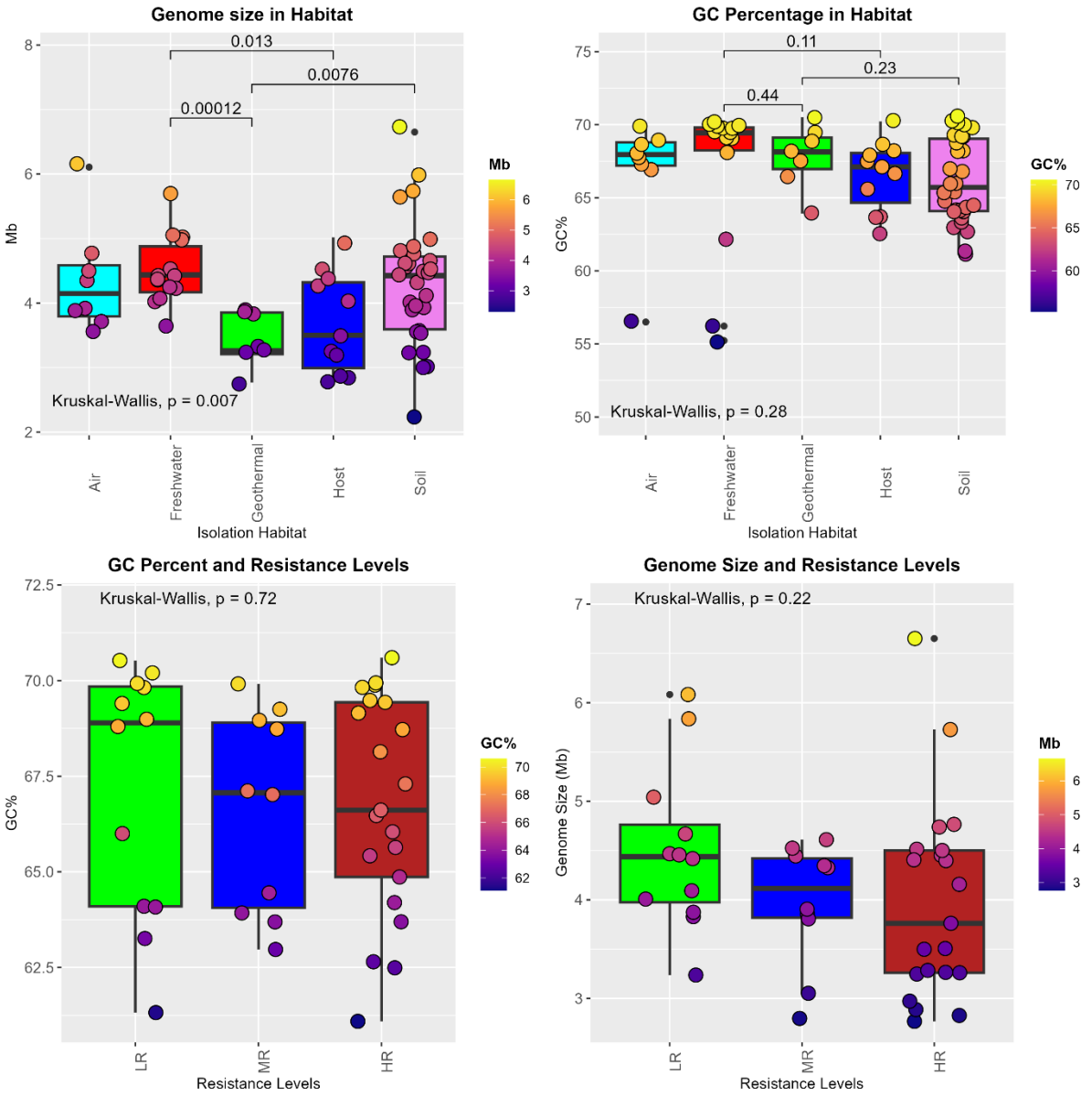

Figure 2-1a**).** However, GC content did not show any correlation with isolation habitat (

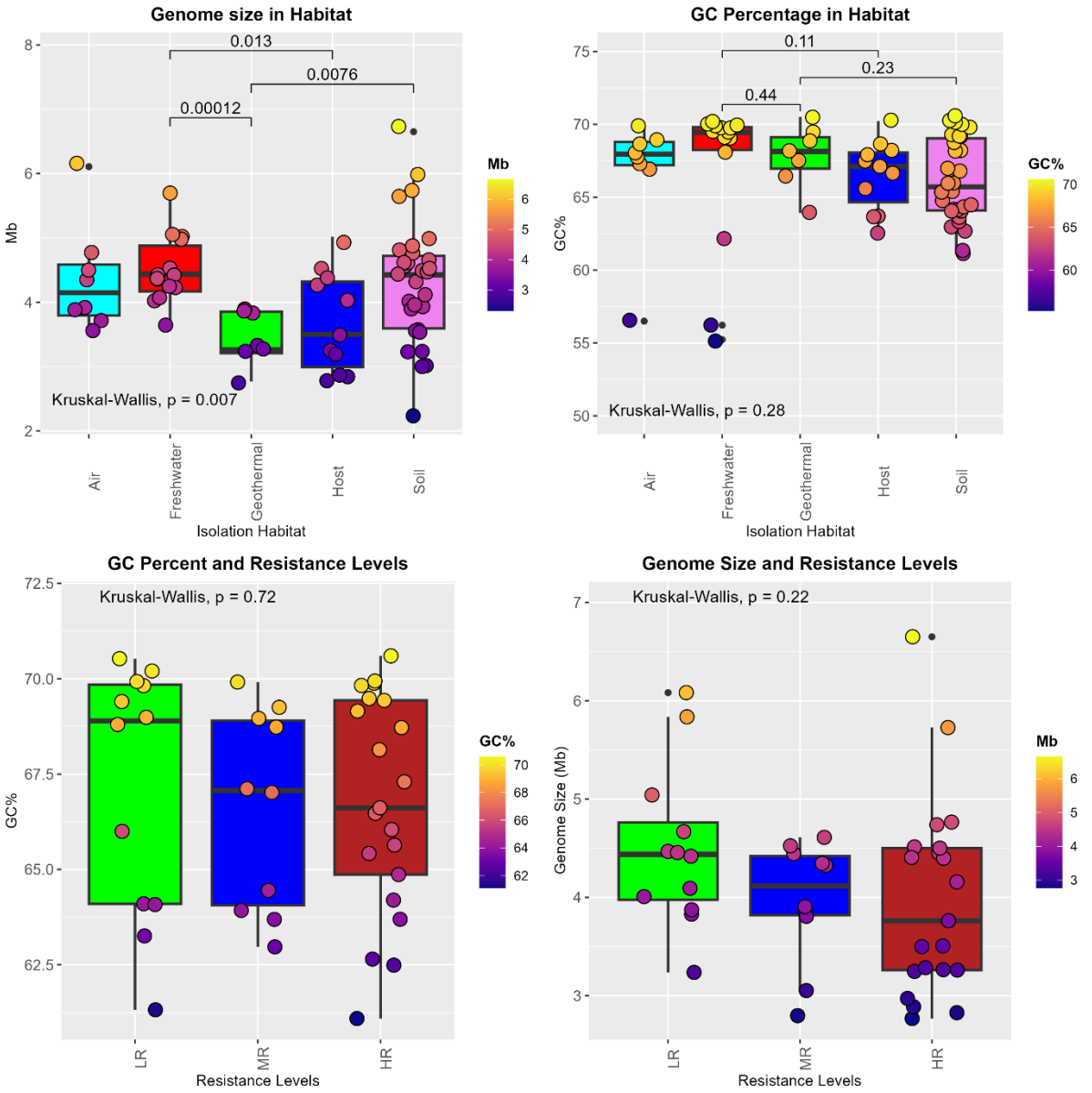

Figure 2-1b). An interpretation of this result is that *Deinococcus* members have a flexible genome and are prone to high levels of evolutionary events like HGT in the new environment. This genomic flexibility enables a higher metabolic diversity, making *Deinococcus* species ubiquitous and adaptable to diverse environments. Our findings indicate a striking difference with the closely related genus *Thermus*, which showed a very different trend with having a homogeneous genome size in a range of 2.02 to 2.56 Mbp and GC content between 64.81% and 69.50% among 23 *Thermus* genomes (Jiao et al., 2022a). Members of the *Thermus* genus are known as specialist organisms and can only colonise and survive high temperatures. This finding aligns with one of the first comparative genomics analyses between *T. thermophilus* and *D. radiodurans*

(Omelchenko et al., 2005), which found that despite sharing a common ancestor with *Thermus*, the *Deinococcus* genus had a different evolutionary pathway and in the direction of surviving in diverse habitats, and this genomic flexibility can provide the underlying requirements for their adaptation.

A subset of our data with 44 *Deinococcus* species with available data on radiation resistance levels (D_{10} values) did not show any significant correlation between genome size and D_{10} values among *Deinococcus* members (

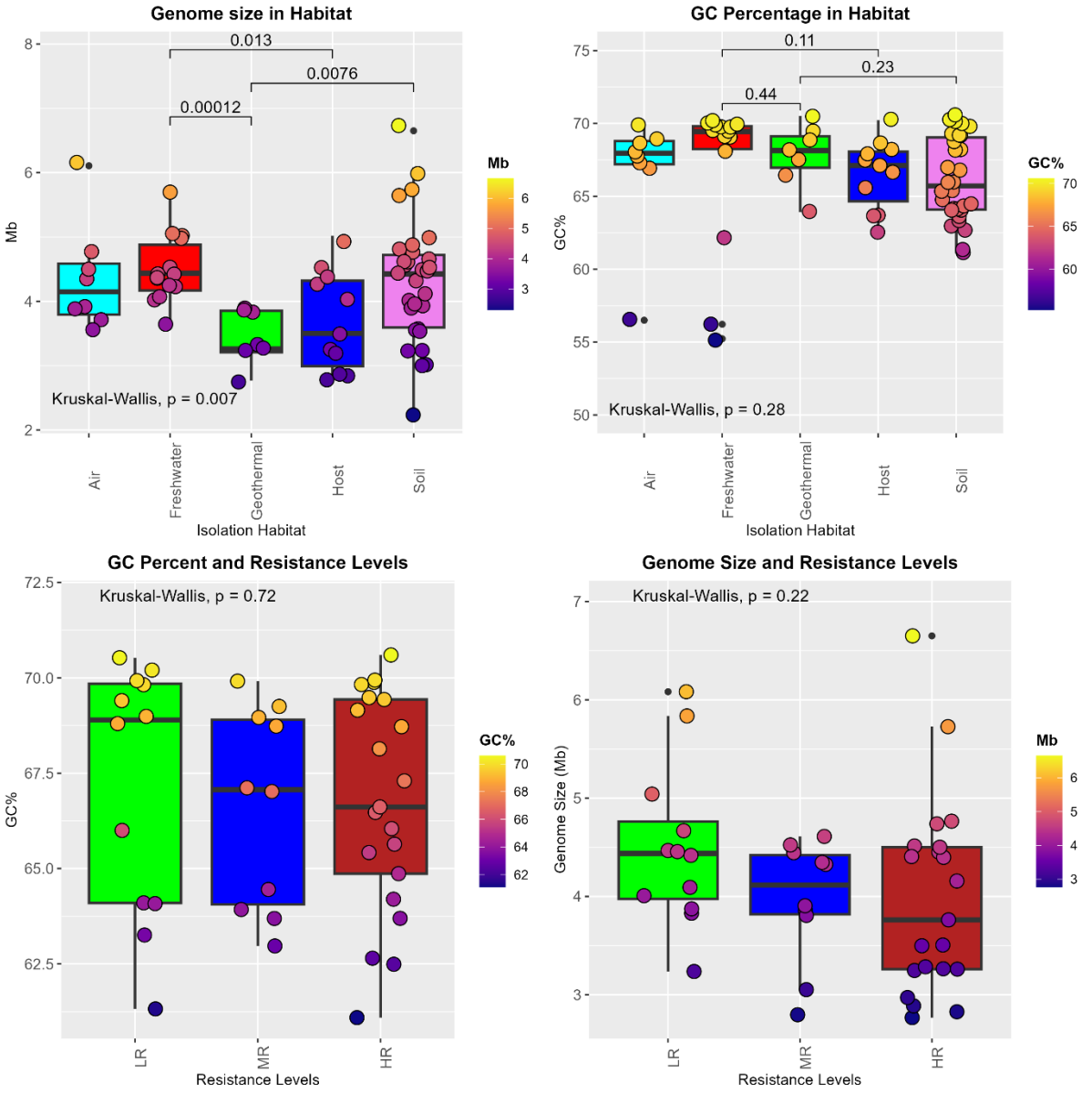

Figure 2-1c,d). This lack of correlation between genome size and radiation resistance levels is in contrast with a long-standing theory that the number of DNA double-strand

breaks (DSBs) increases linearly with an increase in the dose of IR, and therefore, radiation resistance is directly correlated to genome size (Daly, 2012; Daly et al., 1994; Horne et al., 2022). This discrepancy could be explained by the theory that ionising radiation causes more DSBs in larger DNA molecules (Daly, 2023), and the number of replicons, instead of the whole genome sizes, determines the sensitivity of *Deinococcus* species. *Deinococcus* members do not always experience ionising radiation as an evolutionary pressure, so their priority lies in adapting to new habitats or niches they encounter through acquiring adaptive genes via HGT. Hence, acquiring new genes leads to the enlargement of replicons, which in some species can be detrimental for the cell when exposed to double-strand break caused by oxidative stress damage.

2.5.2 *Deinococcus* has a large genomic repertoire

We constructed the pangenome of the *Deinococcus* genus using 66 high-quality genomes representing all sequenced species. Our data indicates that the pangenome of the *Deinococcaceae* includes 22,484 orthogroups, from which 12,055 comprise the core genome (970 orthogroups) and accessory genome (11,085 orthogroups) and 10,429 orthogroups are unique genes (not clustered into any OGs). The Heap's law model fitted with the number of discovered novel genes per genome addition resulted in the decay

rate (α) of 0.55 (

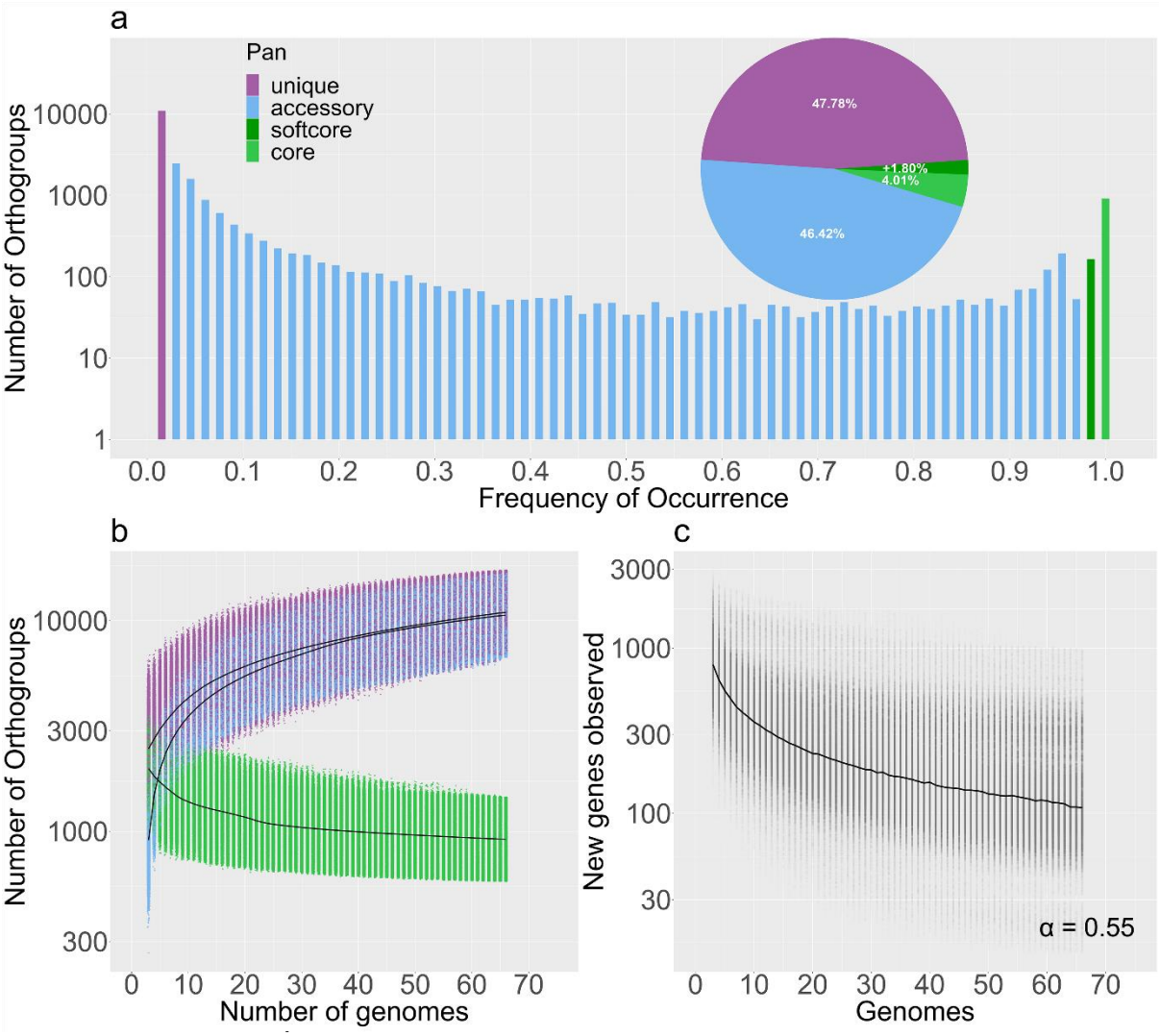

Figure 2-2c), which means the pangenome is open because it is less than one. Our findings highlight the striking difference between the *Deinococcus* pangenome and its close relative, *Thermus*. Jiao and colleagues showed that the pan-genome of the *Thermus* genus is smaller than *Deinococcus* and includes 6992 gene clusters, including 1027 gene families in the core genome (Jiao et al., 2022a). Despite the big difference in the pangenome, the core genome of these two genera has a similar number of OGs.

We note that using genomes at the genus level contributes to this large and open pangenome, but the lack of diversity within the *Deinococcus* species is a limitation in the construction of the pangenome of each species. This trend also means that currently available *Deinococcus* genomes do not capture the entire diversity of this genus.

However, pangenome studies are not always limited to the species levels. A study by Fang et al. (2022) looked at the pangenome of the *Clostridium* genus in an Artificial Ecosystem with the same approach (Fang et al., 2022). Other studies have looked at the pangenome of multiple species in the *Thermus* genus (Jiao et al., 2022a), *Listeria* (Liao et al., 2021), and the plant pathogen *Pectobacterium* (Jonkheer et al., 2021).

One limitation of implementing the pangenome approach in the *Deinococcus* genus is the lack of intraspecies diversity. The majority of *Deinococcus* species have only one known member. Therefore, we used representative genomes for the pangenome reconstruction and downstream analyses. This lack of intraspecies diversity in the *Deinococcus* genus can be interpreted as follows: (i) *Deinococcus* species are under-sequenced in metagenomic data because *Deinococcaceae* usually have a low relative abundance. (ii) They are hard to cultivate and are usually cultivated using Gamma radiation as a pre-culture treatment method, which enriches highly resistant species and precludes the discovery and sequencing of IR-sensitive species. (iii) *Deinococcus* species are monotypic or oligotypic. A monotypic species has a homogeneous distribution of genetic similarities across their entire population (Van Rossum et al., 2020). It is known that monotypic species with low diversity are more likely to be specialists, with narrow geographic distributions or host ranges, or are the product of recent speciation (Bobay & Ochman, 2018; Sheppard et al., 2018; Van Rossum et al., 2020). However, *Deinococcus* members are generalists and are found in diverse environments despite their apparent monotypic nature.

2.5.3 Functional annotation

To gain insight into the function of gene families in the core, accessory, and unique genes, we assigned OGs in each group to COG categories. As expected, the analysis of COG annotation revealed that the highest proportion of assigned genes within the pangenome belonged to the unknown function category (S). Accessory genome and unique genes constituted the majority of unknown genes. The five most abundant categories are carbohydrate transport and metabolism (G), replication, recombination and repair (L), transcription (K), amino acid transport and metabolism (E), and signal

transduction mechanism (T) (Figure 2-3c,d). The enrichment of carbohydrate transport and metabolism and amino acid transport and metabolism aligns with the heterotrophic lifestyle of *Deinococcus* and their reliance on organic substrates. This finding is consistent with previous studies on *D. radiodurans*, which demonstrated its organotrophic and proteolytic lifestyle (Slade & Radman, 2011).

The second abundant category within the *Deinococcus* pangenome is the replication, recombination, and repair mechanisms (L), with a relative abundance of 10% in the whole pangenome. In this category, accessory and unique genes comprise the highest proportion of OGs. It might be unexpected to see core genes as the least abundant genes in the replication, recombination, and repair mechanisms as DNA repair genes are among housekeeping genes. However, the majority of accessory and unique OGs in this category belong to different families of transposase, such as the Tn3-family (tnpA), IS1, IS4, and IS605 OrfB family, constituting approximately 24% of the category (Figure 2-4d). The prevalence of transposase families, coupled with efficient recombination repair in *Deinococcus*, may account for the higher rates of HGT, enabling genome size variability and genetic flexibility observed within the *Deinococcus* genus (C. Lee et al., 2020; Vigil-Stenman et al., 2017). This finding aligns with the theory that genome size reflects not just the number of genes or organismal complexity but strongly correlates with the abundance of transposable elements (Suh, 2019).

The high abundance and diversity of transposase-like enzymes in the *Deinococcus* pangenome supports the findings of previous studies on genome plasticity under oxidative stress, particularly in *D. radiodurans* and *D. geothermalis*, where a high abundance of transposase was reported, suggesting increased genome plasticity under oxidative stress (C. Lee et al., 2020; Pasternak et al., 2010). Furthermore, the association between genome size and the density of transposable elements in both prokaryotes and eukaryotes (Suh, 2019; Touchon & Rocha, 2007) supports the notion that the prevalence of transposase families may contribute to the genome size variability and genetic flexibility observed within the *Deinococcus* genus.

The OGs associated with the core genome in the replication, recombination, and repair category include genes such as *hup* (Histone-like DNA-binding protein), *rnhA* (Endonuclease degrading the RNA of RNA- DNA hybrids), *uvrA*, *xseA*, *topA*, *ssb*, *ddrB*, *recA*, *ligA*, *sbcC*, *yprA*, *ruvB*, *sbcD*, *gyrB*, *mutS*, and numerous other repair enzymes. Extensive research on *D. radiodurans* and *D. geothermalis* has demonstrated the direct involvement of genes like *recA*, *uvrA*, and *ssb* in conferring radiation resistance properties, as their elimination significantly diminishes radiation resistance (Blanchard & de Groot, 2021; Jin et al., 2019; Makarova et al., 2007; Slade et al., 2009). The conservation of these repair genes in the core genome underscores their fundamental role in the survival and basic biology and phenotypes of *Deinococcus,* particularly under extreme conditions such as radiation exposure.

Overall, the enrichment of replication, recombination, and repair mechanisms in the core genome of *Deinococcus* is supported by the conservation of these repair enzymes across *Deinococcus* members. This finding aligns with the theory that the core genome consists of genes shared by all the strains and probably encodes functions related to the basic biology and phenotypes of the species. The accessory and unique genes contribute to the species' diversity and provide functions that are not essential to its basic lifestyle but confer selective advantages, including niche adaptation, antibiotic resistance, and the ability to colonise new hosts (Tettelin et al., 2008).

2.5.4 Radiation and Desiccation Response genes

The presence/absence analysis of 37 genes representing literature-vetted genes belonging to RDR regulon in the *D. radiodurans* (Eugénie et al., 2021; Narasimha & Basu, 2021) showed that all those genes were assigned to OGs. From those 37 OGs, 15 gene families were in the core genome and present in all *Deinococcus* species (Figure 2-5a**).** This observation was expected and in line with the previous findings that DNA repair mechanisms are conserved in all *Deinococcus* species, many of which are unremarkable and homologous to *E. coli*. (Daly, 2023; Makarova et al., 2007).

The other 22 OGs constituted the accessory genes and most likely one factor in the variability of radiation resistance in the *Deinococcus* genus. Most of those gene families

were transposases and transporters, which might potentially impact the radiation resistance levels in different *Deinococcus* species.

2.5.5 Functional ecology of *Deinococcus* genus

The data presented here demonstrated that carbon fixation was not observed in any *Deinococcus* species, and diverse metabolic pathways for complex carbohydrates were present. This finding supports the previous studies that *Deinococcus* species have a chemoorganotrophic lifestyle, as shown for some *Deinococcus* species and the *Thermus* genus (Jiao et al., 2022a; Makarova et al., 2007; Omelchenko et al., 2005).

Moreover, *Deinococcus* species are missing metabolic pathways for synthesising essential amino acids, and they must uptake them from their environments (Krisko & Radman, 2013). Ornithine, threonine, shikimate, methionine, valine/isoleucine, leucine, arginine, histidine, and proline are amino acids that are synthesised by almost all species except two host-associated *Deinococcus* species. Lysine biosynthesis was missing in most host-associated species *D. radiophilus*, isolated from a marine fish, and *Deinococcus sp* KNUC1210, isolated from the rhizome of a plant. We note that this absence could be because of this genome's lower CheckM completeness (87.97%). However, lysine biosynthesis was missing in all environments except three of the host-associated *Deinococcus* species, *D. piscis,* isolated from marine fish. Our data revealed an absence of pathways related to cofactors and vitamin metabolisms like heme, biotin, thiamin, ubiquinone, VB12, and VB6 among *Deinococcus* genomes. This observation leads to a similar conclusion where *Deinococcus* species have evolved to reduce redox factors such as cofactors and proteins (Slade & Radman, 2011).

2.6 Conclusion:

This chapter aimed to identify the factors contributing to the widespread distribution and diverse characteristics of the *Deinococcus* genus. To address the under-representation of radiation-sensitive *Deinococcus* species, we sequenced two previously isolated and reported as radiation-sensitive isolates, added them to publicly available genomes, and constructed the pangenome of the *Deinococcus* genus.

Our results indicate that the *Deinococcus* genus has an open pangenome, meaning the currently available genomes cannot capture the entire diversity of this genus. We found that genes related to carbohydrate and amino acid transport and metabolism were the most abundant in the pangenome, highlighting the chemoorganotrophic lifestyle of *Deinococcus*. Additionally, genes related to transcription and signal transduction mechanisms, which play known roles in habitat adaptation, were also abundant, explaining the adaptability of this genus. Interestingly, genes associated with replication, recombination, and repair were highly diverse and abundant. The presence of diverse families of transposases, coupled with efficient recombination repair mechanisms in *Deinococcus*, may account for the higher rates of horizontal gene transfer (HGT), enabling genome size variability and genetic flexibility within the *Deinococcus* genus.

We conclude that the ubiquity of *Deinococcus* species is mainly due to their genetic flexibility, facilitated by the enrichment of diverse families of transposases. This combination of efficient homologous recombination and abundant transposases contributes to the evolutionary flexibility of *Deinococcus* species, allowing them to adapt to various stressors.

2.7 References

Akita, H., Itoiri, Y., Ihara, S., Takeda, N., Matsushika, A., & Kimura, Z. (2020). *Deinococcus* kurensis sp. Nov., isolated from pond water collected in Japan. *Archives of Microbiology*, *202*(7), 1757–1762. https://doi.org/10.1007/s00203-020-01845-8

Albuquerque, L., Simões, C., Nobre, M. F., Pino, N. M., Battista, J. R., Silva, M. T., Rainey, F. A., & de Costa, M. S. (2005). Truepera radiovictrix gen. Nov., sp. Nov., a new radiation resistant species and the proposal of Trueperaceae fam. Nov. *FEMS Microbiology Letters*, *247*(2), 161–169. https://doi.org/10.1016/j.femsle.2005.05.002

Asgarani, E., Soudi, M. R., Borzooee, F., & Dabbagh, R. (2012). Radio-resistance in psychrotrophic *Kocuria* sp. ASB 107 isolated from Ab-e-Siah radioactive spring. *Journal of Environmental Radioactivity*, *113*, 171–176. https://doi.org/10.1016/j.jenvrad.2012.04.009

Battista, J. R. (2016). *Deinococcus* – Thermus Group. In *Encyclopedia of Life Sciences* (pp. 1–12). https://doi.org/10.1002/9780470015902.a0021151

Baudet, M., Ortet, P., Gaillard, J.-C., Fernandez, B., Guérin, P., Enjalbal, C., Subra, G., de Groot, A., Barakat, M., Dedieu, A., & Armengaud, J. (2010). Proteomics-based Refinement of *Deinococcus* deserti Genome Annotation Reveals an Unwonted Use of Non-canonical Translation Initiation Codons*. *Molecular & Cellular Proteomics*, *9*(2), 415–426. https://doi.org/10.1074/mcp.M900359-MCP200

Bentchikou, E., Servant, P., Coste, G., & Sommer, S. (2010). A Major Role of the RecFOR Pathway in DNA Double-Strand-Break Repair through ESDSA in *Deinococcus radiodurans*. *PLOS Genetics*, *6*(1), e1000774. https://doi.org/10.1371/journal.pgen.1000774

Blanchard, L., & de Groot, A. (2021). Coexistence of SOS-Dependent and SOS-Independent Regulation of DNA Repair Genes in Radiation-Resistant *Deinococcus* Bacteria. *Cells*, *10*(4), Article 4. https://doi.org/10.3390/cells10040924

Bobay, L.-M., & Ochman, H. (2018). Factors driving effective population size and pan-genome evolution in bacteria. *BMC Evolutionary Biology*, *18*(1), 153. https://doi.org/10.1186/s12862-018-1272-4

Brooks, B. W., & Murray, R. G. E. (1981). Nomenclature for 'Micrococcus radiodurans' and other radiation-resistant cocci: *Deinococcaceae* fam. Nov. And *Deinococcus* gen. Nov., including five species. *International Journal of Systematic Bacteriology*, *31*(3), 353–360. https://doi.org/10.1099/00207713-31-3-353

Bruch, E. M., de Groot, A., Un, S., & Tabares, L. C. (2015). The effect of gamma-ray irradiation on the Mn(II) speciation in *Deinococcus radiodurans* and the potential role of Mn(II)-orthophosphates. *Metallomics : Integrated Biometal Science*, *7*(5), 908–916. https://doi.org/10.1039/c5mt00009b

Callegan, R. P., Noble, M. F., McTernan, P. M., Battista, J. R., Navarro-González, R., McKay, C. P., da Costa, M. S., & Rainey, F. A. (2008). Description of four novel psychrophilic, ionizing radiation-sensitive *Deinococcus* species from alpine

environments. *International Journal of Systematic and Evolutionary Microbiology*, *58*(5), 1252–1258. https://doi.org/10.1099/ijs.0.65405-0

Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P., & Huerta-Cepas, J. (2021). eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Molecular Biology and Evolution*, *38*(12), 5825–5829. https://doi.org/10.1093/molbev/msab293

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, *25*(15), 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Chen, M.-Y., Teng, W.-K., Zhao, L., Hu, C.-X., Zhou, Y.-K., Han, B.-P., Song, L.-R., & Shu, W.-S. (2021). Comparative genomics reveals insights into cyanobacterial evolution and habitat adaptation. *The ISME Journal*, *15*(1), 211–227. https://doi.org/10.1038/s41396-020-00775-z

Chen, W., Wang, B., Hong, H., Yang, H., & Liu, S. J. (2011). *Deinococcus* reticulitermitis sp. Nov., isolated from a termite gut. *International Journal of Systematic and Evolutionary Microbiology*, *62*(1), 78–83. https://doi.org/10.1099/ijs.0.026567-0

Coleman, G. A., Davín, A. A., Mahendrarajah, T. A., Szánthó, L. L., Spang, A., Hugenholtz, P., Szöllsi, G. J., & Williams, T. A. (2021). A rooted phylogeny resolves early bacterial evolution. *Science*, *372*(6542). https://doi.org/10.1126/science.abe0511

Comte, N., Morel, B., Hasić, D., Guéguen, L., Boussau, B., Daubin, V., Penel, S., Scornavacca, C., Gouy, M., Stamatakis, A., Tannier, E., & Parsons, D. P. (2020). Treerecs: An integrated phylogenetic tool, from sequences to reconciliations. *Bioinformatics*, *36*(18), 4822–4824. https://doi.org/10.1093/bioinformatics/btaa615

Contreras-Moreira, B., & Vinuesa, P. (2013). GET_HOMOLOGUES, a Versatile Software Package for Scalable and Robust Microbial Pangenome Analysis. *Applied and Environmental Microbiology*, *79*(24), 7696–7701. https://doi.org/10.1128/AEM.02411-13

Daly, M. J. (2012). Death by protein damage in irradiated cells. *DNA Repair*, *11*(1), 12–21. https://doi.org/10.1016/j.dnarep.2011.10.024

Daly, M. J. (2023). The scientific revolution that unraveled the astonishing DNA repair capacity of the Deinococcaceae: 40 years on. *Canadian Journal of Microbiology*. https://doi.org/10.1139/cjm-2023-0059

Daly, M. J., Gaidamakova, E. K., Matrosova, V. Y., Kiang, J. G., Fukumoto, R., Lee, D. Y., Wehr, N. B., Viteri, G. A., Berlett, B. S., & Levine, R. L. (2010). Small-molecule antioxidant proteome-shields in *Deinococcus radiodurans*. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0012570

Daly, M. J., Gaidamakova, E. K., Matrosova, V. Y., Vasilenko, A., Zhai, M., Venkateswaran, A., Hess, M., Omelchenko, M. V., Kostandarithes, H. M., Makarova, K. S., Wackett, L. P., Fredrickson, J. K., & Ghosal, D. (2004). Accumulation of Mn(II) in *Deinococcus radiodurans* Facilitates Gamma-

Radiation Resistance. *Science*, *306*(5698), 1025–1028. https://doi.org/10.1126/science.1103185

Daly, M. J., Ouyang, L., Fuchs, P., & Minton, K. W. (1994). In vivo damage and recA-dependent repair of plasmid and chromosomal DNA in the radiation-resistant bacterium *Deinococcus radiodurans*. *Journal of Bacteriology*, *176*(12), 3508–3517. https://doi.org/10.1128/jb.176.12.3508-3517.1994

David, L. A., & Alm, E. J. (2011). Rapid evolutionary innovation during an Archaean genetic expansion. *Nature*, *469*(7328), 93–96. https://doi.org/10.1038/nature09649

de Groot, A., Chapon, V., Servant, P., Christen, R., Fischer-Le Saux, M., Sommer, S., & Heulin, T. (2005). *Deinococcus* deserti sp. Nov., a gamma-radiation-tolerant bacterium isolated from the Sahara Desert. *International Journal of Systematic and Evolutionary Microbiology*, *55*(6), 2441–2446. https://doi.org/10.1099/ijs.0.63717-0

de Groot, A., Siponen, M. I., Magerand, R., Eugénie, N., Martin-Arevalillo, R., Doloy, J., Lemaire, D., Brandelet, G., Parcy, F., Dumas, R., Roche, P., Servant, P., Confalonieri, F., Arnoux, P., Pignol, D., & Blanchard, L. (2019). Crystal structure of the transcriptional repressor DdrO: Insight into the metalloprotease/repressor-controlled radiation response in *Deinococcus* . *Nucleic Acids Research*, *47*(21), 11403–11417. https://doi.org/10.1093/nar/gkz883

Dewar, A. E., Thomas, J. L., Scott, T. W., Wild, G., Griffin, A. S., West, S. A., & Ghoul, M. (2021). Plasmids do not consistently stabilize cooperation across bacteria but may promote broad pathogen host-range. *Nature Ecology & Evolution*, 1–13. https://doi.org/10.1038/s41559-021-01573-2

Dharamshi, J. E., Köstlbacher, S., Schön, M. E., Collingro, A., Ettema, T. J. G., & Horn, M. (2023). Gene gain facilitated endosymbiotic evolution of Chlamydiae. *Nature Microbiology*, *8*(1), Article 1. https://doi.org/10.1038/s41564-022-01284-9

Ekman, J. V., Raulio, M., Busse, H.-J., Fewer, D. P., & Salkinoja-Salonen, M. (2011). Deinobacterium chartae gen. Nov., sp. Nov., an extremely radiation-resistant, biofilm-forming bacterium isolated from a Finnish paper mill. *International Journal of Systematic and Evolutionary Microbiology*, *61*(3), 540–548. https://doi.org/10.1099/ijs.0.017970-0

Emms, D. M., & Kelly, S. (2019). OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biology*, *20*(1), 238. https://doi.org/10.1186/s13059-019-1832-y

Emms, D. M., & Kelly, S. (2020). Benchmarking Orthogroup Inference Accuracy: Revisiting Orthobench. *Genome Biology and Evolution*, *12*(12), 2258–2266. https://doi.org/10.1093/gbe/evaa211

Eugénie, N., Zivanovic, Y., Lelandais, G., Coste, G., Bouthier de la Tour, C., Bentchikou, E., Servant, P., & Confalonieri, F. (2021). Characterization of the Radiation Desiccation Response Regulon of the Radioresistant Bacterium *Deinococcus radiodurans* by Integrative Genomic Analyses. *Cells*, *10*(10), 2536. https://doi.org/10.3390/cells10102536

Fang, G.-Y., Chai, L.-J., Zhong, X.-Z., Lu, Z.-M., Zhang, X.-J., Wu, L.-H., Wang, S.-T., Shen, C.-H., Shi, J.-S., & Xu, Z.-H. (2022). Comparative Genomics Unveils the Habitat Adaptation and Metabolic Profiles of Clostridium in an Artificial Ecosystem for Liquor Production. *mSystems*, *7*(3), e00297-22. https://doi.org/10.1128/msystems.00297-22

Ferreira, A. C., Nobre, M. F., Rainey, F. A., Silva, M. T., Wait, R., Burghardt, J., Chung, A. P., & da Costa, M. S. (1997). *Deinococcus* geothermalis sp. Nov. And *Deinococcus* murrayi sp. Nov., two extremely radiation-resistant and slightly thermophilic species from hot springs. *International Journal of Systematic Bacteriology*, *47*(4), 939–947. https://doi.org/10.1099/00207713-47-4-939

Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E. L. L., Tate, J., & Punta, M. (2014). Pfam: The protein families database. *Nucleic Acids Research*, *42*(D1), D222–D230. https://doi.org/10.1093/nar/gkt1223

Gaidamakova, E. K., Sharma, A., Matrosova, V. Y., Grichenko, O., Volpe, R. P., Tkavc, R., Conze, I. H., Klimenkova, P., Balygina, I., Horne, W. H., Gostinčar, C., Chen, X., Makarova, K. S., Shuryak, I., Srinivasan, C., Jackson-Thompson, B., Hoffman, B. M., & Daly, M. J. (2022). Small-Molecule Mn Antioxidants in Caenorhabditis elegans and *Deinococcus radiodurans* Supplant MnSOD Enzymes during Aging and Irradiation. *mBio*, *13*(1), e0339421. https://doi.org/10.1128/mbio.03394-21

Gayen, M., Gupta, P., Morazzani, E. M., Gaidamakova, E. K., Knollmann-Ritschel, B., Daly, M. J., Glass, P. J., & Maheshwari, R. K. (2017). *Deinococcus* Mn2+-peptide complex: A novel approach to alphavirus vaccine development. *Vaccine*, *35*(29), 3672–3681. https://doi.org/10.1016/j.vaccine.2017.05.016

Ghiassi-nejad, M., Mortazavi, S. M. J., Cameron, J. R., Niroomand-rad, A., & Karam, P. A. (2002). Very high background radiation areas of Ramsar, Iran: Preliminary biological studies. *Health Physics*, *82*(1), 87–93. https://doi.org/10.1097/00004032-200201000-00011

Ghosal, D., Omelchenko, M. V., Gaidamakova, E. K., Matrosova, V. Y., Vasilenko, A., Venkateswaran, A., Zhai, M., Kostandarithes, H. M., Brim, H., Makarova, K. S., Wackett, L. P., Fredrickson, J. K., & Daly, M. J. (2005). How radiation kills cells: Survival of *Deinococcus radiodurans* and Shewanella oneidensis under oxidative stress⋆. *FEMS Microbiology Reviews*, *29*(2), 361–375. https://doi.org/10.1016/j.fmrre.2004.12.007

Gupta, P., Gayen, M., Smith, J. T., Gaidamakova, E. K., Matrosova, V. Y., Grichenko, O., Knollmann-Ritschel, B., Daly, M. J., Kiang, J. G., & Maheshwari, R. K. (2016). MDP: A *Deinococcus* Mn2+-Decapeptide Complex Protects Mice from Ionizing Radiation. *PloS One*, *11*(8), e0160575. https://doi.org/10.1371/journal.pone.0160575

Hirsch, P., Gallikowski, C. A., Siebert, J., Peissl, K., Kroppenstedt, R., Schumann, P., Stackebrandt, E., & Anderson, R. (2004). *Deinococcus* frigens sp. Nov., *Deinococcus* saxicola sp. Nov., and *Deinococcus* marmoris sp. Nov., Low Temperature and Draught-tolerating, UV-resistant Bacteria from Continental

Antarctica. *Systematic and Applied Microbiology*, *27*(6), 636–645. https://doi.org/10.1078/0723202042370008

Horne, W. H., Volpe, R. P., Korza, G., DePratti, S., Conze, I. H., Shuryak, I., Grebenc, T., Matrosova, V. Y., Gaidamakova, E. K., Tkavc, R., Sharma, A., Gostinčar, C., Gunde-Cimerman, N., Hoffman, B. M., Setlow, P., & Daly, M. J. (2022). Effects of Desiccation and Freezing on Microbial Ionizing Radiation Survivability: Considerations for Mars Sample Return. *Astrobiology*, *22*(11), 1337–1350. https://doi.org/10.1089/ast.2022.0065

Jeon, S. H., Kang, M. S., Joo, E. S., Kim, E. B., Lim, S., Jeong, S. W., Jung, H. Y., Srinivasan, S., & Kim, M. K. (2016). *Deinococcus* persicinus sp. Nov., a radiationresistant bacterium from soil. *International Journal of Systematic and Evolutionary Microbiology*, *66*(12), 5077–5082. https://doi.org/10.1099/ijsem.0.001473

Jiao, J.-Y., Lian, Z.-H., Li, M.-M., Salam, N., Zhou, E.-M., Liu, L., Ming, H., Nie, G., Shu, W., Zhao, G., Hedlund, B. P., & Li, W.-J. (2022a). Comparative genomic analysis of Thermus provides insights into the evolutionary history of an incomplete denitrification pathway. *mLife*, *1*(2), 198–209. https://doi.org/10.1002/mlf2.12009

Jiao, J.-Y., Lian, Z.-H., Li, M.-M., Salam, N., Zhou, E.-M., Liu, L., Ming, H., Nie, G., Shu, W., Zhao, G., Hedlund, B. P., & Li, W.-J. (2022b). Comparative genomic analysis of Thermus provides insights into the evolutionary history of an incomplete denitrification pathway. *mLife*, *1*(2), 198–209. https://doi.org/10.1002/mlf2.12009

Jin, M., Xiao, A., Zhu, L., Zhang, Z., Huang, H., & Jiang, L. (2019). The diversity and commonalities of the radiation-resistance mechanisms of *Deinococcus* and its up-to-date applications. *AMB Express*, *9*(1), 138. https://doi.org/10.1186/s13568-019-0862-x

Jonkheer, E. M., Brankovics, B., Houwers, I. M., van der Wolf, J. M., Bonants, P. J. M., Vreeburg, R. A. M., Bollema, R., de Haan, J. R., Berke, L., Smit, S., de Ridder, D., & van der Lee, T. A. J. (2021). The Pectobacterium pangenome, with a focus on Pectobacterium brasiliense, shows a robust core and extensive exchange of genes from a shared gene pool. *BMC Genomics*, *22*(1), 265. https://doi.org/10.1186/s12864-021-07583-5

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), Article 6. https://doi.org/10.1038/nmeth.4285

Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, *28*(1), 27–30.

Katoh, K., & Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kawaguchi, Y., Shibuya, M., Kinoshita, I., Yatabe, J., Narumi, I., Shibata, H., Hayashi, R., Fujiwara, D., Murano, Y., Hashimoto, H., Imai, E., Kodaira, S., Uchihori, Y., Nakagawa, K., Mita, H., Yokobori, S., & Yamagishi, A. (2020). DNA Damage and Survival Time Course of Deinococcal Cell Pellets During 3 Years of

Exposure to Outer Space. *Frontiers in Microbiology*, *11*. https://www.frontiersin.org/articles/10.3389/fmicb.2020.02050

Kawaguchi, Y., Yang, Y., Kawashiri, N., Shiraishi, K., Takasu, M., Narumi, I., Satoh, K., Hashimoto, H., Nakagawa, K., Tanigawa, Y., Momoki, Y., Tanabe, M., Sugino, T., Takahashi, Y., Shimizu, Y., Yoshida, S., Kobayashi, K., Yokobori, S., & Yamagishi, A. (2013). 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. *Origins of Life and Evolution of Biospheres*, *43*(4), 411–428. https://doi.org/10.1007/s11084-013-9346-1

Kim, E. B., Kang, M. S., Joo, E. S., Jeon, S. H., Jeong, S. W., Lim, S. Y., Jung, H. Y., Srinivasan, S., & Kim, M. K. (2017). *Deinococcus* ruber sp. Nov., a radiation-resistant bacterium isolated from soil. *International Journal of Systematic and Evolutionary Microbiology*, *67*(1), 72–76. https://doi.org/10.1099/ijsem.0.001567

Krisko, A., & Radman, M. (2013). Biology of Extreme Radiation Resistance: The Way of *Deinococcus radiodurans*. *Cold Spring Harbor Perspectives in Biology*, *5*(7), a012765. https://doi.org/10.1101/cshperspect.a012765

Lai, W. A., Kämpfer, P., Arun, A. B., Shen, F. T., Huber, B., Rekha, P. D., & Young, C. C. (2006). *Deinococcus ficus* sp. Nov., isolated from the rhizosphere of Ficus religiosa L. *International Journal of Systematic and Evolutionary Microbiology*, *56*(4), 787–791. https://doi.org/10.1099/ijs.0.64007-0

Lee, C., Choo, K., & Lee, S.-J. (2020). Active Transposition of Insertion Sequences by Oxidative Stress in *Deinococcus* geothermalis. *Frontiers in Microbiology*, *0*. https://doi.org/10.3389/fmicb.2020.558747

Lee, J. H., Jung, J.-H., Kim, M.-K., & Lim, S. (2022). *Deinococcus* taeanensis sp. Nov., a Radiation-Resistant Bacterium Isolated from a Coastal Dune. *Current Microbiology*, *79*(11), 334. https://doi.org/10.1007/s00284-022-03044-8

Lee, J., Lee, Y. H., Park, S. J., Lim, S., Jeong, S. W., Lee, S.-Y., Park, S., Choi, H.-W., Kim, M. K., & Jung, H.-Y. (2016). *Deinococcus* sedimenti sp. Nov. Isolated from river sediment. *Journal of Microbiology*, *54*(12), 802–808. https://doi.org/10.1007/s12275-016-6361-8

Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, *49*(W1), W293–W296. https://doi.org/10.1093/nar/gkab301

Liao, J., Guo, X., Weller, D. L., Pollak, S., Buckley, D. H., Wiedmann, M., & Cordero, O. X. (2021). Nationwide genomic atlas of soil-dwelling Listeria reveals effects of selection and population ecology on pangenome evolution. *Nature Microbiology*, *6*(8), Article 8. https://doi.org/10.1038/s41564-021-00935-7

Lim, S., Jung, J.-H., Blanchard, L., & de Groot, A. (2018). Conservation and diversity of radiation and oxidative stress resistance mechanisms in *Deinococcus* species. *FEMS Microbiology Reviews*, *43*(1), 19–52. https://doi.org/10.1093/femsre/fuy037

Ludanyi, M., Blanchard, L., Dulermo, R., Brandelet, G., Bellanger, L., Pignol, D., Lemaire, D., & de Groot, A. (2014). Radiation response in *Deinococcus* deserti:

IrrE is a metalloprotease that cleaves repressor protein DdrO. *Molecular Microbiology*, *94*(2), 434–449. https://doi.org/10.1111/mmi.12774

Makarova, K. S., Aravind, L., Wolf, Y. I., Tatusov, R. L., Minton, K. W., Koonin, E. V., & Daly, M. J. (2001). Genome of the Extremely Radiation-Resistant Bacterium *Deinococcus radiodurans* Viewed from the Perspective of Comparative Genomics. *Microbiology and Molecular Biology Reviews*, *65*(1), 44–79. https://doi.org/10.1128/mmbr.65.1.44-79.2001

Makarova, K. S., & Daly, M. J. (2010). Comparative Genomics of Stress Response Systems in *Deinococcus* Bacteria. In *Bacterial Stress Responses* (pp. 445–457). John Wiley & Sons, Ltd. https://doi.org/10.1128/9781555816841.ch27

Makarova, K. S., Omelchenko, M. V., Gaidamakova, E. K., Matrosova, V. Y., Vasilenko, A., Zhai, M., Lapidus, A., Copeland, A., Kim, E., Land, M., Mavromatis, K., Pitluck, S., Richardson, P. M., Detter, C., Brettin, T., Saunders, E., Lai, B., Ravel, B., Kemner, K. M., … Daly, M. J. (2007). *Deinococcus* geothermalis: The Pool of Extreme Radiation Resistance Genes Shrinks. *PLoS ONE*, *2*(9), e955. https://doi.org/10.1371/journal.pone.0000955

Makarova, K. S., Wolf, Y. I., White, O., Minton, K., & Daly, M. J. (1999). Short repeats and IS elements in the extremely radiation-resistant bacterium *Deinococcus radiodurans* and comparison to other bacterial species. *Research in Microbiology*, *150*(9–10), 711–724. https://doi.org/10.1016/S0923-2508(99)00121-7

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, *37*(5), 1530–1534. https://doi.org/10.1093/molbev/msaa015

Mohseni, M., Abbaszadeh, J., & Nasrollahi Omran, A. (2014). Radiation resistant of native *Deinococcus* spp. Isolated from the Lout desert of Iran "the hottest place on Earth". *International Journal of Environmental Science and Technology*, *11*(7), 1939–1946. https://doi.org/10.1007/s13762-014-0643-7

Monfared, A. S., Jalali, F., Mozdarani, H., Hajiahmadi, M., & Samavat, H. (2005). Living in high natural background radiation areas in Ramsar, Iran. Is it dangerous for health? *International Congress Series*, *1276*, 438–439. https://doi.org/10.1016/j.ics.2004.12.007

Mu, D.-S., Wang, S., Liang, Q.-Y., Du, Z.-Z., Tian, R., Ouyang, Y., Wang, X.-P., Zhou, A., Gong, Y., Chen, G.-J., Van Nostrand, J., Yang, Y., Zhou, J., & Du, Z.-J. (2020). Bradymonabacteria, a novel bacterial predator group with versatile survival strategies in saline environments. *Microbiome*, *8*(1), 126. https://doi.org/10.1186/s40168-020-00902-0

Murray, R. G. E. (1992). The Family Deinococcaceae. In *The Prokaryotes* (pp. 3732–3744). Springer New York. https://doi.org/10.1007/978-1-4757-2191-1_42

Narasimha, A., & Basu, B. (2021). New insights into the activation of Radiation Desiccation Response regulon in *Deinococcus radiodurans*. *Journal of Biosciences*, *46*(1), 10. https://doi.org/10.1007/s12038-020-00123-5

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood

Phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274. https://doi.org/10.1093/molbev/msu300

Omelchenko, M. V., Wolf, Y. I., Gaidamakova, E. K., Matrosova, V. Y., Vasilenko, A., Zhai, M., Daly, M. J., Koonin, E. V., & Makarova, K. S. (2005). Comparative genomics of Thermus thermophilus and *Deinococcus* radiodurans: Divergent routes of adaptation to thermophily and radiation resistance. *BMC Evolutionary Biology*, *5*, 57. https://doi.org/10.1186/1471-2148-5-57

Panitz, C., Frösler, J., Wingender, J., Flemming, H.-C., & Rettberg, P. (2019). Tolerances of *Deinococcus* geothermalis Biofilms and Planktonic Cells Exposed to Space and Simulated Martian Conditions in Low Earth Orbit for Almost Two Years. *Astrobiology*, *19*(8), 979–994. https://doi.org/10.1089/ast.2018.1913

Parks, D. H., Chuvochina, M., Rinke, C., Mussig, A. J., Chaumeil, P.-A., & Hugenholtz, P. (2022). GTDB: An ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Research*, *50*(D1), D785–D794. https://doi.org/10.1093/nar/gkab776

Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P. A., & Hugenholtz, P. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nature Biotechnology*. https://doi.org/10.1038/nbt.4229

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research*, *25*(7), 1043–1055. https://doi.org/10.1101/gr.186072.114

Pasternak, C., Ton-Hoang, B., Coste, G., Bailone, A., Chandler, M., & Sommer, S. (2010). Irradiation-Induced *Deinococcus radiodurans* Genome Fragmentation Triggers Transposition of a Single Resident Insertion Sequence. *PLOS Genetics*, *6*(1), e1000799. https://doi.org/10.1371/journal.pgen.1000799

Pavlov, A. K., Kalinin, V. L., Konstantinov, A. N., Shelegedin, V. N., & Pavlov, A. A. (2006). Hypothesis Paper Was Earth Ever Infected by Martian Biota? Clues from Radioresistant Bacteria. In *ASTROBIOLOGY* (6; Vol. 6). www.liebertpub.com

Raposo, P., Viver, T., Albuquerque, L., Froufe, H., Barroso, C., Egas, C., Rosselló-Móra, R., & da Costa, M. S. (2019). Transfer of Meiothermus chliarophilus (Tenreiro et al.1995) Nobre et al. 1996, Meiothermus roseus Ming et al. 2016, Meiothermus terrae Yu et al. 2014 and Meiothermus timidus Pires et al. 2005, to Calidithermus gen. Nov., as Calidithermus chliarophilus comb. Nov., Calidithermus roseus comb. Nov., Calidithermus terrae comb. Nov. And Calidithermus timidus comb. Nov., respectively, and emended description of the genus Meiothermus. *International Journal of Systematic and Evolutionary Microbiology*, *69*(4), 1060–1069. https://doi.org/10.1099/ijsem.0.003270

Reback, J., McKinney, W., jbrockmendel, Bossche, J. V. den, Augspurger, T., Cloud, P., gfyoung, Sinhrks, Klein, A., Roeschke, M., Tratner, J., She, C., Ayd, W., Hawkins, S., Petersen, T., Schendel, J., Hayden, A., Garcia, M., Jancauskas, V., … Kluyver, T. (2020). *pandas-dev/pandas: Pandas 1.0.0* (v1.0.0) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.3630805

Rosenberg, E. (2014). The family deinococcaceae. In *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea* (pp. 613–615). Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-642-38954-2_127

Sako, Y., Nakagawa, S., Takai, K., & Horikoshi, K. (2003). Marinithermus hydrothermalis gen. Nov., sp. Nov., a strictly aerobic, thermophilic bacterium from a deep-sea hydrothermal vent chimney. *International Journal of Systematic and Evolutionary Microbiology*, *53*(Pt 1), 59–65. https://doi.org/10.1099/ijs.0.02364-0

Scornavacca, C., Mayol, J. C. P., & Cardona, G. (2017). Fast algorithm for the reconciliation of gene trees and LGT networks. *Journal of Theoretical Biology*, *418*, 129–137. https://doi.org/10.1016/j.jtbi.2017.01.024

Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*, *30*(14), 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Selengut, J. D., Haft, D. H., Davidsen, T., Ganapathy, A., Gwinn-Giglio, M., Nelson, W. C., Richter, A. R., & White, O. (2007). TIGRFAMs and Genome Properties: Tools for the assignment of molecular function and biological process in prokaryotic genomes. *Nucleic Acids Research*, *35*(Database issue), D260-264. https://doi.org/10.1093/nar/gkl1043

Sharma, A., Gaidamakova, E. K., Grichenko, O., Matrosova, V. Y., Hoeke, V., Klimenkova, P., Conze, I. H., Volpe, R. P., Tkavc, R., Gostinčar, C., Gunde-Cimerman, N., Diruggiero, J., Shuryak, I., Ozarowski, A., Hoffman, B. M., Daly, M. J., Designed, M. J. D., & Performed, A. O. (2017). *Across the tree of life, radiation resistance is governed by antioxidant Mn 2+ , gauged by paramagnetic resonance*. https://doi.org/10.1073/pnas.1713608114

Shashidhar, R., & Bandekar, J. R. (2009). *Deinococcus* piscis sp. Nov., a radiation-resistant bacterium isolated from a marine fish. *International Journal of Systematic and Evolutionary Microbiology*, *59*(11), 2714–2717. https://doi.org/10.1099/ijs.0.003046-0

Shashidhar, R., Kumar, S. A., Misra, H. S., & Bandekar, J. R. (2010). Evaluation of the role of enzymatic and nonenzymatic antioxidant systems in the radiation resistance of *Deinococcus* . *Canadian Journal of Microbiology*, *56*(3), 195–201. https://doi.org/10.1139/w09-118

Sheppard, S. K., Guttman, D. S., & Fitzgerald, J. R. (2018). Population genomics of bacterial host adaptation. *Nature Reviews Genetics*, *19*(9), Article 9. https://doi.org/10.1038/s41576-018-0032-z

Sheridan, P. O., Raguideau, S., Quince, C., Holden, J., Zhang, L., Williams, T. A., & Gubry-Rangin, C. (2020). Gene duplication drives genome expansion in a major lineage of Thaumarchaeota. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-19132-x

Shimoyama, Y. (2021, October 21). *Moshi4/FastDTLmapper*. GitHub. https://github.com/moshi4/FastDTLmapper/blob/main/CITATION.cff

Slade, D., Lindner, A. B., Paul, G., & Radman, M. (2009). Recombination and Replication in DNA Repair of Heavily Irradiated *Deinococcus radiodurans. Cell*, *136*(6), 1044–1055. https://doi.org/10.1016/j.cell.2009.01.018

Slade, D., & Radman, M. (2011). Oxidative Stress Resistance in *Deinococcus radiodurans. Microbiology and Molecular Biology Reviews*, *75*(1), 133–191. https://doi.org/10.1128/mmbr.00015-10

Snipen, L., & Liland, K. H. (2015). micropan: An R-package for microbial pan-genomics. *BMC Bioinformatics*, *16*(1), 79. https://doi.org/10.1186/s12859-015-0517-0

Sturmbauer, C. (2013). Paraphyly. In S. Maloy & K. Hughes (Eds.), *Brenner's Encyclopedia of Genetics (Second Edition)* (pp. 225–226). Academic Press. https://doi.org/10.1016/B978-0-12-374984-0.01119-0

Suh, A. (2019). Genome Size Evolution: Small Transposons with Large Consequences. *Current Biology*, *29*(7), R241–R243. https://doi.org/10.1016/j.cub.2019.02.032

Suresh, K., Reddy, G. S. N., Sengupta, S., & Shivaji, S. (2004). *Deinococcus* indicus sp. Nov., an arsenic-resistant bacterium from an aquifer in West Bengal, India. *International Journal of Systematic and Evolutionary Microbiology*, *54*(2), 457–461. https://doi.org/10.1099/ijs.0.02758-0

Szöllősi, G. J., Rosikiewicz, W., Boussau, B., Tannier, E., & Daubin, V. (2013). Efficient Exploration of the Space of Reconciled Gene Trees. *Systematic Biology*, *62*(6), 901–912. https://doi.org/10.1093/sysbio/syt054

Tettelin, H., Riley, D., Cattuto, C., & Medini, D. (2008). Comparative genomics: The bacterial pan-genome. *Current Opinion in Microbiology*, *11*(5), 472–477. https://doi.org/10.1016/j.mib.2008.09.006

Tobin, G. J., Tobin, J. K., Gaidamakova, E. K., Wiggins, T. J., Bushnell, R. V., Lee, W.-M., Matrosova, V. Y., Dollery, S. J., Meeks, H. N., Kouiavskaia, D., Chumakov, K., & Daly, M. J. (2020). A novel gamma radiation-inactivated sabin-based polio vaccine. *PLOS ONE*, *15*(1), e0228006. https://doi.org/10.1371/journal.pone.0228006

Touchon, M., & Rocha, E. P. C. (2007). Causes of Insertion Sequences Abundance in Prokaryotic Genomes. *Molecular Biology and Evolution*, *24*(4), 969–981. https://doi.org/10.1093/molbev/msm014

Van Rossum, T., Ferretti, P., Maistrenko, O. M., & Bork, P. (2020). Diversity within species: Interpreting strains in microbiomes. *Nature Reviews Microbiology*, *18*(9), Article 9. https://doi.org/10.1038/s41579-020-0368-1

Vigil-Stenman, T., Ininbergs, K., Bergman, B., & Ekman, M. (2017). High abundance and expression of transposases in bacteria from the Baltic Sea. *The ISME Journal*, *11*(11), Article 11. https://doi.org/10.1038/ismej.2017.114

von Meijenfeldt, F. A. B., Hogeweg, P., & Dutilh, B. E. (2023). A social niche breadth score reveals niche range strategies of generalists and specialists. *Nature Ecology & Evolution*, *7*(5), Article 5. https://doi.org/10.1038/s41559-023-02027-7

Wang, L.-G., Lam, T. T.-Y., Xu, S., Dai, Z., Zhou, L., Feng, T., Guo, P., Dunn, C. W., Jones, B. R., Bradley, T., Zhu, H., Guan, Y., Jiang, Y., & Yu, G. (2020). Treeio: An R Package for Phylogenetic Tree Input and Output with Richly Annotated and Associated Data. *Molecular Biology and Evolution*, *37*(2), 599–603. https://doi.org/10.1093/molbev/msz240

Wang, X.-P., Li, C.-M., Yu, Y., Li, H.-R., Du, Z.-J., & Mu, D. 2019. (n.d.). *Deinococcus* arcticus sp. Nov., isolated from Silene acaulis rhizosphere soil of the Arctic

tundra. *International Journal of Systematic and Evolutionary Microbiology*, *69*(11), 3437–3442. https://doi.org/10.1099/ijsem.0.003636

White, O., Eisen, J. A., Heidelberg, J. F., Hickey, E. K., Peterson, J. D., Dodson, R. J., Haft, D. H., Gwinn, M. L., Nelson, W. C., Richardson, D. L., Moffat, K. S., Qin, H., Jiang, L., Pamphile, W., Crosby, M., Shen, M., Vamathevan, J. J., Lam, P., McDonald, L., … Fraser, C. M. (1999). Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. *Science (New York, N.Y.)*, *286*(5444), 1571–1577. https://doi.org/10.1126/science.286.5444.1571

Wick, R. R., & Holt, K. E. (2022). Polypolish: Short-read polishing of long-read bacterial genome assemblies. *PLOS Computational Biology*, *18*(1), e1009802. https://doi.org/10.1371/journal.pcbi.1009802

Wick, R. R., Judd, L. M., Cerdeira, L. T., Hawkey, J., Méric, G., Vezina, B., Wyres, K. L., & Holt, K. E. (2021). Trycycler: Consensus long-read assemblies for bacterial genomes. *Genome Biology*, *22*(1), 266. https://doi.org/10.1186/s13059-021-02483-z

Wick, R. R., Judd, L. M., & Holt, K. E. (2019). Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biology*, *20*(1), 129. https://doi.org/10.1186/s13059-019-1727-y

Williams, T. A., Davin, A. A., Morel, B., Szánthó, L. L., Spang, A., Stamatakis, A., Hugenholtz, P., & Szöllősi, G. J. (2023). *The power and limitations of species tree-aware phylogenetics* [Preprint]. Evolutionary Biology. https://doi.org/10.1101/2023.03.17.533068

Yamagishi, A., Kawaguchi, Y., Hashimoto, H., Yano, H., Imai, E., Kodaira, S., Uchihori, Y., & Nakagawa, K. (2018). Environmental Data and Survival Data of *Deinococcus* aetherius from the Exposure Facility of the Japan Experimental Module of the International Space Station Obtained by the Tanpopo Mission. *Astrobiology*, *18*(11), 1369–1374. https://doi.org/10.1089/ast.2017.1751

Yang, Y., Itoh, T., Yokobori, S., Itahashi, S., Shimada, H., Satoh, K., Ohba, H., Narumi, I., & Yamagishi, A. (2009). *Deinococcus* aerius sp. Nov., isolated from the high atmosphere. *International Journal of Systematic and Evolutionary Microbiology*, *59*(8), 1862–1866. https://doi.org/10.1099/ijs.0.007963-0

Yin, L.-Z., Li, J.-L., Liu, Z.-T., Fang, B.-Z., Wang, P., Luo, X.-Q., Dong, L., Duan, L., Li, S.-H., & Li, W.-J. (2022). *Deinococcus* aestuarii sp. Nov. And *Deinococcus* aquaedulcis sp. Nov., two novel resistant bacteria isolated from pearl river estuary. *Antonie van Leeuwenhoek*, *115*(1), 59–68. https://doi.org/10.1007/s10482-021-01680-x

Yu, G. (2020). Using ggtree to Visualize Data on Tree-Like Structures. *Current Protocols in Bioinformatics*, *69*(1), e96. https://doi.org/10.1002/cpbi.96

Zhang, K., Zhu, J., Li, S., Rao, M. P. N., Li, N.-M., Guo, A.-Y., Jiang, Z., Tang, Q.-Y., Wan, Y., Zhang, Z.-D., & Li, W.-J. (2020a). *Deinococcus* detaillensis sp. Nov., isolated from humus soil in Antarctica. *Archives of Microbiology*, *202*(9), 2493–2498. https://doi.org/10.1007/s00203-020-01920-0

Zhang, K., Zhu, J., Li, S., Rao, M. P. N., Li, N.-M., Guo, A.-Y., Jiang, Z., Tang, Q.-Y., Wan, Y., Zhang, Z.-D., & Li, W.-J. (2020b). *Deinococcus* detaillensis sp. Nov.,

isolated from humus soil in Antarctica. *Archives of Microbiology*, *202*(9), 2493–2498. https://doi.org/10.1007/s00203-020-01920-0

Zhou, Z., Tran, P. Q., Breister, A. M., Liu, Y., Kieft, K., Cowley, E. S., Karaoz, U., & Anantharaman, K. (2022). METABOLIC: High-throughput profiling of microbial genomes for functional traits, metabolism, biogeochemistry, and community-scale functional networks. *Microbiome*, *10*(1), 33. https://doi.org/10.1186/s40168-021-01213-8

3 Chapter III - Genomic Basis for Radiation Resistance Diversity in *Deinococcus*

Question:

1- What drives the diversity of radiation resistance in the *Deinococcus* genus?

2- Is IR resistance predictable using comparative genomics methods?

Hypothesis: Analysis of diverse genomic samples in the *Deinococcus* genus can reveal a correlation between radiation resistance phenotype and genomic content, suggesting a genomic basis for this phenotype.

3.1 Abstract:

Since *Deinococcus radiodurans* was isolated from irradiated meat cans in 1956, it has fascinated microbiologists because of its exceptional resilience to the lethal impacts of ionising radiation. Several mechanisms have been proposed to explain radiation resistance in *Deinococcus,* but scientists have failed to define a genomic basis for this peculiar trait. Radiation resistance is usually measured in D_{10}, which is the dose of gamma radiation required to kill 90% of the population. This value is highly variable and ranges in the order of magnitude for *Deinococcus* species. The inconsistency in radiation resistance capacities further complicated the lack of genomic explanation for this phenotype. As a result, prominent researchers have posited that radiation resistance in *Deinococcus* is a phenotype without genotype and that this trait cannot be predicted from genomic content. Radiation resistance mechanisms are usually defined by the upregulation of DNA repair mechanisms and antioxidant protection of proteomes against oxidative stress.

However, previous studies have always focused on genes related to radiation resistance found in *D. radiodurans,* and other genomic components of the *Deinococcus* members have remained unexplored. This chapter aims to address the gap in our knowledge by establishing a robust phylogeny between members of the *Deinococcus* genus and then searching for drivers of radiation resistance variability in the accessory genome, which may contribute to the phenotypic diversity.

I assembled a dataset including genomes of 44 *Deinococcus* species with their radiation resistance metadata and found a phylogenomic pattern in this trait. Except for one species, all radio-sensitive species fell into two clades, and this taxonomic pattern was used for statistical analyses of their accessory genome. My data showed a taxonomic pattern in radiation resistance levels where sensitive species formed two cades. Statistical analyses revealed a negative correlation between 188 gene families and radiation resistance, meaning gaining those genes could have led to radiation sensitivity. Functional annotation of those genes showed that most gene families belong to co-factors and oxidative agents, which increase sensitivity to oxidative stressors. In

conclusion, the study suggests that specific gene acquisition can induce radiation sensitivity in *Deinococcus* species. Given the genomic flexibility of *Deinococcus* members and the absence of natural evolutionary pressure for radiation resistance on Earth, the gain of genes aiding adaptation to new environments may inadvertently lead to toxicity under extreme oxidative stress conditions.

3.2 Introduction

Deinococcus radiodurans, the type species of the *Deinococcus* genus, was isolated from irradiated meat can (Anderson et al., 1956) and has become a well-studied model bacterium for extreme resistance to lethal doses of ionising radiation (IR)(Cox & Battista, 2005). Even low doses of ionising radiation are highly lethal to the majority of life forms. For instance, only five to ten Gy of gamma radiation can kill a human, and 800 Gy is enough to eradicate 90 % of the populations of *Escherichia coli* (Slade & Radman, 2011). However, the dose to decrease 90% of the population for *D. radiodurans* is 15 kGy (12000 Gy) at room temperature. It has been reported that *D. radiodurans* could survive under a staggering dose of 140 kGy when the cells are desiccated and frozen (Horne et al., 2022). As an extremophile, *D. radiodurans* gained popularity over the past five decades and was among the first sequenced bacteria due to its extreme resistance to ionising radiation and its potential in bioremediation and biotechnology (White et al., 1999).

As of February 2024, the *Deinococcus* genus has 94 species, of which 91 are validly published (https://www.bacterio.net/genus/*Deinococcus*) and more than 300 reports on the cultivation and identification of isolates at the genus level. The radiation resistance level is usually reported in the D_{10} value and is highly diverse among the *Deinococcaceae* family and can range from 1.5 kGy for *D. persicinus* (Jeon et al., 2016) to 15 kGy for *D. geothermalis* and *D. radiodurans* (Ferreira et al., 1997; Jeon et al., 2016; Makarova & Daly, 2010). This example and some other radiosensitive *Deinococcus* species suggest that extreme radiation resistance is not conserved among all *Deinococcus* species. However, the radiation-resistance phenotype in highly resistant *Deinococcus* species has attracted the majority of the research interest because of its role in bioremediation and biotechnology.

Ionising radiation and UV rays pose substantial threats to organisms in two ways. Directly by disrupting cell walls, inducing DNA lesions and double-strand breaks, and indirectly by generating reactive oxygen species (ROS) that lead to irreversible damage in DNA and proteins. Over five decades, scientists have explored the complex trait of

radiation resistance. The known mechanisms can be categorised into three general areas (Makarova & Daly, 2010): (i) polyploid genomic structure and efficient DNA repair mechanisms, (ii) oxidative protection against ROS, and (iii) regulatory pathways controlling gene transcription. Classical DNA repair mechanisms, including base excision repair, nucleotide excision repair, and mismatch repair, are present in *D. radiodurans* (Lim et al., 2018; Makarova et al., 2001). Furthermore, a specific recombination type called extended synthesis-dependent strand annealing, known as ESDSA, enables *D. radiodurans* to reassemble genome fragments and repair DNA shortly after radiation exposure (Bentchikou et al., 2010; Zahradka et al., 2006). On the other hand, transcriptomic studies highlighted regulatory pathways that tightly control the expression of radiation response genes in *D. radiodurans* after exposure to radiation or reactive oxygen species (ROS) (S. Im et al., 2013; Narasimha & Basu, 2021; Tanaka et al., 2004). However, the knowledge about DNA repair genes and their regulatory pathways could not fully explain extreme radiation resistance traits (Lim et al., 2018; Slade & Radman, 2011).

In the early years of radiobiology, the classical view was that DNA damage determines cytotoxicity (Hutchinson, 1966). However, it was later shown that the double-strand DNA damage in highly resistant organisms was the same as in radiation-sensitive organisms (Daly et al., 2004; Makarova & Daly, 2010). M.J. Daly and his colleagues introduced a new theory in the field of radiation resistance and challenged this classical view. They proposed that proteins, instead of DNA, are the main target of oxidative damage, and if an organism can protect its proteome and has the other factors (genomic polyploidy and efficient DNA repair machinery), it can reassemble the fragmented DNA and survive the lethal effects of oxidative stress (Daly, 2009). This theory was proposed from the study that showed that *D. radiodurans*, as well as other radiation-resistant organisms, protected its proteome against ROS by small complexes of Mn^{2+} molecules with peptides (Daly, 2009; Daly et al., 2004, 2010). These findings highlighted the critical role of Mn^{2+} ions in the antioxidant system of radiation-resistant organisms (Bruch et al., 2015; Gaidamakova et al., 2022; Gupta et al., 2016).

Another important finding regarding the proteome protection in the *D. radiations* was the key role of high ratios of Mn/Fe in a cell. In contrast to highly radioresistant cells like *D. radiodurans*, the metal-reducing bacterium *Shewanella oneidensis* accumulates Fe but not Mn and is extremely sensitive to radiation (Ghosal et al., 2005). In contrast, *D. radiodurans* reduced its dependence on iron ions by eliminating heme-like proteins and decreasing proteins that require [Fe_2-S_2] and [Fe_4-S_4] clusters as co-factors (Slade & Radman, 2011). Iron is essential for all living organisms but generates highly reactive oxygen radicals through the Fenton reaction (Chandrangsu et al., 2017). It has been shown that localisation and transport of the metals may be more effective than only high ratios of Mn/Fe for reducing oxidative damage in *Deinococcus* (Santos et al., 2019).

In addition to Fe, there are other sources of endogenous ROS in the cell. Co-factors, such as metal ions and small organic molecules, are required in more than 30% of proteins and are involved in respiration, enzyme production, transcription, and signal transduction (Cao & Li, 2011). The balance in using such molecules is very important in cellular survival under oxidative stress conditions (Slade & Radman, 2011). Moreover, organic compounds like Acetyl coenzyme A (acetyl-CoA), flavin adenine dinucleotide (FAD), nicotinamide adenine dinucleotide (NAD) and vitamins like thiamin are common co-factors are involved enzymes such as transferases, oxidoreductases, lyases, ligases, isomerase, and hydrolases (Sun et al., 2023). These co-factors have redox properties and can become a threat in the presence of ROS. For instance, accumulation of the free form of divalent metals, especially iron, catalyses the Fenton reaction and generates hydroxyl radicals, which is the most reactive ROS, and denaturing proteins such as enzymes and DNA but also phospholipid structures such as cell membranes (Szabo et al., 2021).

Previous studies on *D. radiodurans* have shown that the evolution of this organism has been in the direction of reducing endogenous ROS production (Ghosal et al., 2005). It is well established that the auto-oxidation of respiratory chain enzymes and, consequently, the transfer of electrons from NADH and (FADH2) to oxygen yields ROS (Imlay, 2003). It has been shown that *D. radiodurans* evolved a lifestyle with fewer

cytochromes flavoproteins and enzymes with iron-sulfur clusters, leading to minimum ROS generation (Slade & Radman, 2011). For instance, one study shows that *D. radiolarians* encode 65% fewer proteins with iron-sulfur clusters ([2Fe-2S] and [4Fe-4S]) and 56% fewer cytochromes and flavoproteins compared to the radiation-sensitive bacterium *Shewanella oneidensis* (Ghosal et al., 2005). Moreover, *D. radiodurans* limits the release of NADH and FADH2 by inducing the glyoxylate bypass of the TCA cycle when grown in a defined minimal medium and in response to radiation, which bypasses the formation of 2-oxoglutarate and succinyl coenzyme A (CoA) and the release of NADH and FADH2 (Ghosal et al., 2005; Liu et al., 2003).

Comparative genomics studies have been carried out to compare IR resistance in *D. radiodurans* with sensitive species but have failed to find a set of genes that define this phenotype (Lim et al., 2018; Makarova et al., 2001; Makarova & Daly, 2010), because the genes that are consistently present in radioresistant but absent from radiosensitive organisms across the living organisms has shrunk to zero (Daly, 2023). Despite decades of research on radiation resistance, we are facing many open questions regarding the radiation resistance mechanism in the *Deinococcaceae* family, especially members of the *Deinococcus* genus. An enigma in the case of the *Deinococcus* genus is the highly variable radiation resistance capabilities, which are not only highly variable among different species but also among strains of the same species (de Groot et al., 2005). This enigma is partly due to the focus on the extreme radiation resistance of *D. radiodurans* and some other species within the family. This focus on radiation resistance has led to the conventional wisdom that the radiation resistance phenotype is the defining characteristic of the *Deinococcaceae* family (Daly, 2023; Murray, 1992). As an implication of such a view, microbiologists have used gamma radiation as an enrichment method to isolate *Deinococcus* species from environmental samples (Ferreira et al., 1997; D. U. Kim, Jang, et al., 2018; J. H. Lee et al., 2022; J. J. Lee et al., 2013; Rainey et al., 2005). This method is effective for the isolation of radiation-resistant *Deinococcus* strains but potentially leads to a biased population of highly resistant isolates and a distorted understanding of their phenotypes and ecology.

Currently, it is widely accepted that radiation resistance is only predictable by detecting high concentrations of Mn^{2+} complexes, which govern resistance to high doses of gamma radiation across all domains of life (Daly, 2023; Daly et al., 2004, 2010; Gaidamakova et al., 2022; Gupta et al., 2016; Sharma et al., 2017). In this line, it has been argued that "while the number of *Deinococcus* genomes is now sufficiently large to determine the core genome and pangenome of the family *Deinococcaceae*, extreme resistance persistently remains a phenotype without a genotype." (Daly, 2023), highlighting this persisting gap in linking the genomic content of *Deinococcus* species to the diverse radiation resistance phenotype. However, as of January 2024, there has not been any published analysis of the *Deinococcaceae* family pangenome, and the comparative genomic analysis in the second chapter of this thesis as the first construction of *Deinococcus* pangenome showed that this genus has an open pangenome, meaning the available genomes do not encapsulate the entire diversity of the pangenome.

In recent years, comparative genomics analyses with the pangenome approach have successfully answered questions about bacterial complex phenotypes like their virulence (Vornhagen et al., 2022), production of the capsule (Bentley et al. 2006), synthesis of the extracellular polymeric substance EPS (Harris et al. 2017), and modification of the cell wall (Gerlach et al. 2018). In this context, a pangenome is defined as all genes shared between all strains of a species and categorised into the core genome (genes conserved across all strains in the species) and the accessory genome (or distributed genomes). In general, core genomes are enriched for housekeeping functions of basic biology and confer the determining characteristics of the species, while accessory genomes provide phenotypic variability in a species (Innamorati et al., 2020).

In this study, we build on the pangenome of the *Deinococcus* genus and detect a phylogenomic pattern in the variability of radiation resistance among *Deinococcus* species. Then, we use the accessory genome, which explains the phenotype variability (Tettelin et al., 2008) combined with a statistical approach and gain insight into the

genomic basis for radiation resistance. This approach is novel because it uses all genes constituting the diverse ecological traits in the *Deinococcus* genus rather than focusing on radiation resistance genes known from *D. radiodurans*, which have been extensively studied. With this analysis, we try to answer two main questions (i) Is radiation resistance the defining characteristic of *Deinococcaceae*? (ii) Can radiation resistance levels in the *Deinococcus* genus be predicted using genomic data? We identified a phylogenomic pattern in the levels of radiation resistance and used this pattern to compare members of two radiosensitive clades with radioresistant *Deinococcus* species. Then, we applied a statistical approach to calculate the correlation between the presence/absence and copy number of all genes in the accessory genome of the *Deinococcus* genus with the level of radiation resistance. Finally, we used a linear regression model to assess the predictability of radiation resistance levels using genomic contents of known *Deinococcus* species.

3.3 Methods:

3.3.1 Datasets compilation

Representative genomes of the *Deinococcus* genus were downloaded from NCBI and quality-controlled using checkM v.1.2.2 (Parks et al., 2015). Genome quality score was calculated based on the following formula (completeness - contamination * 5), and high-quality genomes with quality scores > 90 and completeness > 94 were selected. Then, we extracted the available D_{10} values reported for those *Deinococcus* species from the literature, which resulted in 42 species with the required metadata for this analysis. Furthermore, we included the unpublished genomes of *D. radiomollis* and *D. altitudinis* and compiled a dataset comprising the whole genome and metadata of 44 *Deinococcus* species, including D_{10} values and genomic features, including genome size and GC content. We categorised *Deinococcus* species into three classes based on their D_{10} values: from 1- 4.9 kGy are categorised as a low resistance group, from 5-9.9 kGy moderate resistance group, and more than 10 kGy high resistance groups (Table 3-1, Figure 3-1). We used this categorisation to detect a general pattern in radiation resistance. We note that D_{10} values might not always be accurate, and some studies might have errors in calculating the D_{10} values. Therefore, the thresholds defined here must be treated as a general trend rather than a definitive level of resistance for each species (see Table 3-1). We used this dataset to calculate the correlation between radiation resistance phenotype and the genome content of *Deinococcus* members. Open reading frames (ORFs) were predicted from the genomic DNA sequences using the Prokka package v 1.14.5 (Seemann, 2014), and the output proteome sequences were used for Orthology and phylogenomic analyses.

3.3.2 Pangenome construction:

To access the genomic repertoire of the *Deinococcus* species in our dataset, we constructed the pangenome of 44 selected genomes. OrthoFinder v.2.5.2 with default parameters was used to infer orthogroups, which are genes resulting from speciation events (Emms & Kelly, 2019). A pangenome can be divided into core genome, accessory, and unique genes. A core genome is defined as genes shared between all

species members. Accessory genome is defined as genes shared between more than one but not all species. Unique genes are genes only found in one species. The single-copy core genome provides a conserved set of genes and can be used as a marker for phylogenomic analysis, while the accessory genome is known for defining differences in organisms derived from a common ancestor (Tettelin et al., 2008). Unassigned genes, also known as singletons or unique genes, were discarded because they are present in only one species and cannot provide statistical power. Accessory genomes of the 44 *Deinococcus* species were used to investigate the variability of radiation resistance, and single copy orthogroups from the core genome were used for phylogenomic analyses.

3.3.3 Phylogenomic analysis:

To construct a robust phylogenomic tree, we used 604 single-copy orthogroups shared between 44 genomes (identified from the OrthoFinder). Then, we concatenated those genes and generated a Multiple Sequence Alignment (MSA) file for species tree reconstruction. The MSA file was used as input to the IQ-TREE v.2.2.2.2 using the model finder option with the following option**:** iqtree2 -m MFP -madd LG+C20, LG+C60 with 1000 bootstraps. The best substitution model was chosen as LG+F+R8 based on the Bayesian Information Criterion (BIC). *Deinobacterium chartae, Truepera radiovictrix* and *Thermus thermophilus* were used as outgroup species. We also reconstructed a species tree using 34 gene markers used in CheckM1 (Parks et al., 2015) to re-evaluate the topology of the tree reconstructed using single-copy orthogroups (data not shown).

To explore whether the IR-resistance phenotype has a phylogenomic basis within the *Deinococcus* genus, we used our dataset to map D_{10} values on the phylogenomic tree. The Treeio package in R was used for tree modification (L.-G. Wang et al., 2020), and the Ggtree package v 3.6.2 was used for tree visualisation and annotation (Yu, 2020; Yu et al., 2017). The phylogenomic tree was annotated with D_{10} values as bar plots and revealed a phylogenomic pattern in radiation resistance (Figure 3-1)

3.3.4 Statistical analysis:

We performed a combination of phylogenomic analysis with statistical tests to explore the correlation between the phylogeny and the presence absence of accessory genes and different levels of ionising radiation resistance. We used the phylogenomic pattern in radiation resistance observed in the phylogenomic tree (Figure 3-1). We used the two distinct clades to categorise sensitive species into two distinct clades: Clade N6 and N31 (Figure 3-1). The accessory genes are defined as genes present in more than one but not all species.

We calculated the correlation of accessory genome between all radioresistant species with identified clades (N6 and N31) once at a time. For the N6 clade, we filtered accessory genes for species belonging to N6 against all other radioresistant species. Then, we calculated the correlation of the gene copy number of each orthogroup with the radiation resistance level designated with the D_{10} value. This process was carried out for the N31 clade, allowing us to identify orthogroups demonstrating a significant correlation with radiation resistance levels in each of those two sensitive clades separately to allow for the detection of distinct evolutionary events. *D. budaensis,* which formed a single branch, was excluded from our analysis.

For the statistical analysis, we used the SciPy 1.0 package in Python (Virtanen et al., 2020). The Pearson correlation coefficient (r) was calculated with a 95% confidence interval (CI). This analysis aimed to discover the correlations between the copy number of each orthogroup and the D_{10} values of the corresponding genome. False Discovery Rate (FDR) adjustment was performed using the Benjamini–Hochberg method (Benjamini & Hochberg, 1995) with an adjusted p-value cut-off 0.05. We also sought to look for genes that solely their presence or absence might have a correlation with radiation resistance. Wilcoxon rank-sum test on categorical variables was performed by converting the dataset to a binary matrix, where one represents the presence, and zero represents the absence of orthogroups. We used the Benjamini–Hochberg method for FDR adjustment.

3.3.5 Functional annotation of correlated orthogroups:

To understand the functional potential of orthogroups with significant correlation with radiation resistance levels, we used Interproscan and EggNOG mapper to annotate all sequences within the correlated orthogroups.

3.3.6 Linear regression model

We created a linear regression model to predict the radiation resistance levels of *Deinococcus* species. Orthogroups identified as significant in the statistical analysis were used to create the model as a predictor function of D_{10} values. In short, we used DIAMOND v2.1.6 (Buchfink et al., 2015) to create a database from identified orthogroup sequences in both clades. Our database included 188 gene families, 93 orthogroups for N6 and 95 orthogroups for N31. Then, we used the sensitive argument in DIAMOND to detect the exact number of hits in each genome (as the query sequence) against the identified orthogroups database. To avoid repeated counting, orthogroups with more than one hit were counted as one. Next, we made a model based on the total number of hits for orthogroups found in both clades for each genome and related radiation resistance level (D_{10} value) using the lm() function in R lm (gamma~og_sum, data = correlation_data_predict). Then, we used this linear model to predict the D_{10} values for each clade using the built-in function predict() in R.

3.4 Results:

3.4.1 Ecological implications of radiation resistance variability

A comparison of radiation resistance between *Deinococcus* species in our dataset revealed a significant diversity in radiation resistance levels previously reported in the literature. For instance, *D. grandis* exhibited a low D_{10} value of 1.7 kGy (See Appendix 1), while *D. geothermalis* demonstrated high resistance with a D_{10} value of 15 kGy. For visual representation, we categorised the *Deinococcus* species into three main categories based on their reported radiation resistance levels (D10 values). Those groups are low resistance groups (LR), moderate resistance (MR), and high resistance (HR) groups (Table 1). Species within the low resistance group were mostly isolated from cold environments such as Antarctica, the Arctic, alpine regions, and freshwater habitats (Callegan et al., 2008; W.-T. Im et al., 2008; E. B. Kim et al., 2017; Oyaizu et al., 1987). An exception to this pattern was *D. budaensis*, a mesophilic species isolated from a biofilm sample in a hydrothermal spring cave (Makk et al., 2016). *D. budaensis* also formed a separate clade in the phylogenomic tree. Species in the moderate resistance group were isolated from diverse habitats, including hot pools, marine animals, metal surfaces, and air samples of a clean room from the Phoenix project in NASA. Species in the HR group are mostly isolated from the soil of deserts, hot springs, and animal faeces. However, the model organism *D. radiodurans* was isolated from irradiated meat can.

Table 3-1. List of *Deinococcus* species with known radiation resistance levels*

Category	Species	Isolation Source	D10 (kGy)	Reference
Low resistance	*Deinococcus grandis**	Water	1.7	(Oyaizu et al., 1987)
	Deinococcus aquaticus	Water	2	(W.-T. Im et al., 2008)
	Deinococcus radiomollis	Soil	2.2	(Callegan et al., 2008)
	Deinococcus sedimenti	Water	2.5	(J. J. Lee et al., 2016)
	Deinococcus budaensis	Biofilm	3	(Makk et al., 2016)
	Deinococcus ruber	Soil	2.5	(E. B. Kim et al., 2017)
	Deinococcus arcticus	host-plant	3	(Dong et al., 2015)
	Deinococcus taeanensis	coastal dune	3.1	(J. H. Lee et al., 2022)
	Deinococcus seoulensis	Water	3.5	(J.-J. Lee et al., 2016)
	Deinococcus altitudinis	Soil	3.8	(Callegan et al., 2008)
	Deinococcus alpinitundrae	Soil	4	(Callegan et al., 2008)
	Deinococcus indicus	Water	4.2	(Suresh et al., 2004)
	Deinococcus soli	Soil	4.5	(Cha et al., 2014b)
	Deinococcus detaillensis	Soil	4.8	(Zhang et al., 2020)
Moderate resistance	*Deinococcus aerius***	Air dust	5	(Yang et al., 2009)
	Deinococcus radiopugnans	Animal	5.5	(Brooks & Murray, 1981)
	Deinococcus metallilatus	Metal-surface	5.7	(D. U. Kim et al., 2015)
	Deinococcus multiflagellatus	Metal surface	6.1	(D. U. Kim, Lee, et al., 2018)
	Deinococcus malanensis	Soil	6.5	(Zhu et al., 2017)
	Deinococcus irradiatisoli	Soil	6.5	(D. U. Kim, Jang, et al., 2018)
	Deinococcus piscis	Animal	7.4	(Shashidhar & Bandekar, 2009)
	Deinococcus fonticola	Biofilm	7.5	(Makk et al., 2019)
	Deinococcus deserti VCD115	Soil	8	(de Groot et al., 2005)
	Deinococcus phoenicis	Air dust	8.5	(Vaishampayan et al., 2014)
	Deinococcus radiotolerans	Soil	8.5	(Cha et al., 2014a)
	Deinococcus murrayi	Water	9	(Ferreira et al., 1997)
	Deinococcus humi	Soil	9.5	(Srinivasan et al., 2012)
	Deinococcus actinosclerus	Soil	9.5	(Sun Joo et al., 2016)
	Deinococcus peraridilitoris	Soil	10	(Rainey et al., 2007)
High resistance	*Deinococcus hopiensis*	Soil	10	(Rainey et al., 2005)
	Deinococcus yavapaiensis	Soil	10	(Rainey et al., 2005)
	Deinococcus apachensis	Soil	10	(Rainey et al., 2005)
	Deinococcus maricopensis	Soil	10	(Rainey et al., 2005)
	*Truepera radiovictrix****	Water	10	(Albuquerque et al., 2005)
	Deinococcus proteolyticus	Animal	10.3	(Slade & Radman, 2011)
	Deinococcus koreensis	Water	10.6	(Baek et al., 2018)
	Deinococcus ficus	Plant	11	(Lai et al., 2006)
	Deinococcus puniceus	Soil	11.5	(J. J. Lee et al., 2015)
	*Deinobacterium chartae*****	Metal-surface	12	(Ekman et al., 2011)
	Deinococcus psychrotolerans	Soil	12	(Tian et al., 2019)

Deinococcus wulumuqiensis R12	Soil	12	(Xu et al., 2013)
Deinococcus terrestris	china	12	(J.-J. Wang et al., 2020)
Deinococcus radiodurans R1	Animal	12.7	(Daly et al., 2004)
Deinococcus gobiensis I-0	Soil	12.7	(Yuan et al., 2009)
Deinococcus geothermalis	Water	15	(Ferreira et al., 1997)

* This table provides the list of *Deinococcus* species compiled in this study.
** The measured D_{10} value for *Deinococcus aerius* may be incorrect because (Yang et al., 2009) used *D. radiodurans* as their control, and the measured D_{10} value was 6.7 kGy, but we know this value for wild-type *D. radiodurans* is around 12 – 15 kGy
***Truepera radiovictrix* is the only isolated and cultivated member of the *Trueperaceae* family
****Deinobacterium chartae* is the single species within the *Deinobacterium* genus

3.4.2 Pangenome and phylogenomic analyses:

We constructed a phylogenomic tree using the single copy orthogroups and annotated the tree with D_{10} species values. The phylogenomic analysis revealed a pattern in radiation resistance trends of two clades of *Deinococcus* species. All members of low resistance were clustered in two distinct clades, N6 and N31, and one branch without any sister taxa (*D. budaensis*) (Figure 3-1, Table 3-2. Radiosensitive Deinococcus *species* belonging to the two clades). Notably, an exception was observed with the presence of the highly radiation-resistant species *D. psychrotolerans*, isolated from Antarctic soil within the N6 clade. Additionally, a few species within the moderate resistance group were identified in both radiosensitive clades. One species in N6 and three in N31 belonged to the moderate resistant group, but two species were between 5 - 6 kGy, which still can be treated as low resistance considering the standard errors in the calculation of D_{10} values.

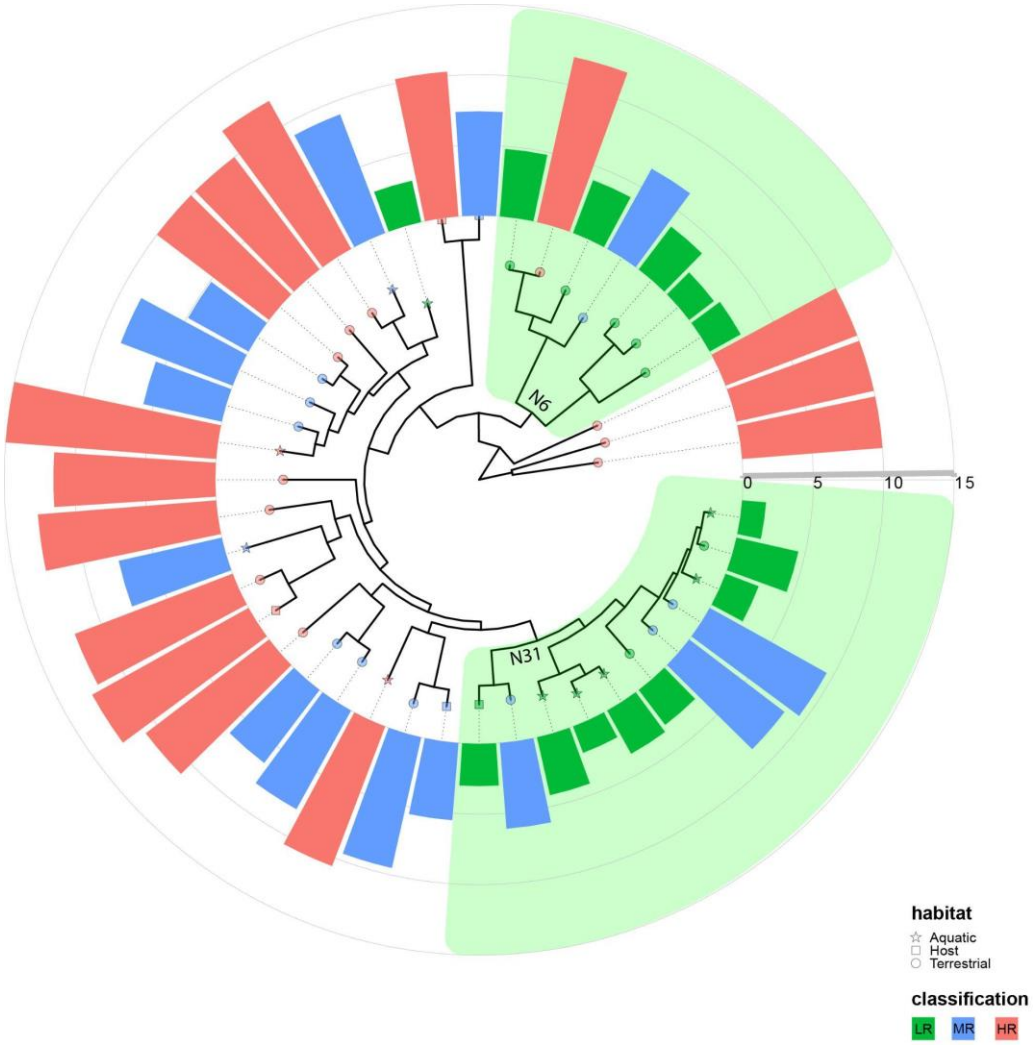

Figure 3-1. Phylogenomic tree of the studied *Deinococcus* species.

The maximum likelihood tree was constructed using the concatenation of 604 single-copy orthogroups as a marker. Bar plots represent the radiation resistance levels. Highlighted crowns: The child taxa of the N6 and N31 clades are mostly radiosensitive and show a radiation sensitivity pattern in two clades. *Deinobacterium chartae* and *Truepera radiovictrix* were used as outgroups to root the tree but were not shown in the figure.

Table 3-2. Radiosensitive *Deinococcus* species belonging to the two clades

Clades	Species	D_{10} (kGy)
Node N6	*Deinococcus ruber*	2.5
	Deinococcus radiomollis	2.2
	Deinococcus alpinitundrae	4
	Deinococcus irradiatisoli	6.5
	Deinococcus detaillensis	4.8
	Deinococcus psychrotolerans	12
Node N31	*Deinococcus grandis**	1.7
	Deinococcus soli	4.5
	Deinococcus actinosclerus	9.5
	Deinococcus sedimenti	2.5
	Deinococcus radiotolerans	8.5
	Deinococcus taeanensis	3.1
	Deinococcus aquaticus	2
	Deinococcus seoulensis	3.5
	Deinococcus indicus	4.2
	Deinococcus arcticus	3
	Deinococcus metallilatus	5.7

3.4.3 Accessory genome and radiation resistance diversity

To correlate the radiation resistance phenotype variability of 44 *Deinococcus* species with their genomic content, we use a pangenome approach focusing on the accessory genes, including genes shared among more than one but not all group members. This analysis was done separately for each radiosensitive clade.

For the N6 clade versus others, 6720 accessory orthogroups were included in the statistical analysis, and 93 orthogroups exhibited a significant correlation with D_{10} values (p-value < 0.05) after FDR correction. Similarly, the comparison of the genomes of the N31 clade with others revealed that 95 orthogroups out of 6864 accessory orthogroups displayed a significant correlation for presence, absence, or copy number against D_{10} values (

Figure 3-2, Supplementary Data S3-2, S3-3).

Almost all significant orthogroups showed a negative Pearson coefficient value. In other words, when these orthogroups are present in a group or have a higher copy number, the D_{10} values tend to be lower. This finding indicates an association between these genes and the radiation sensitivity phenotype observed in species belonging to the N6 and N31 clades. An exception to this trend was observed with the OG0002059 orthogroup within the N6, which was annotated as the PrsW-protease enzyme and showed a positive correlation with D_{10} values. Notably, only one orthogroup, out of 188, was shared between the two sensitive clades. This gene belongs to the HAD family hydrolase, which comprises a diverse group of enzymes that are involved in the hydrolysis of a wide range of substrates like carbon-phosphorus bonds, phosphoesters, and epoxide rings (Morais et al., 2004).

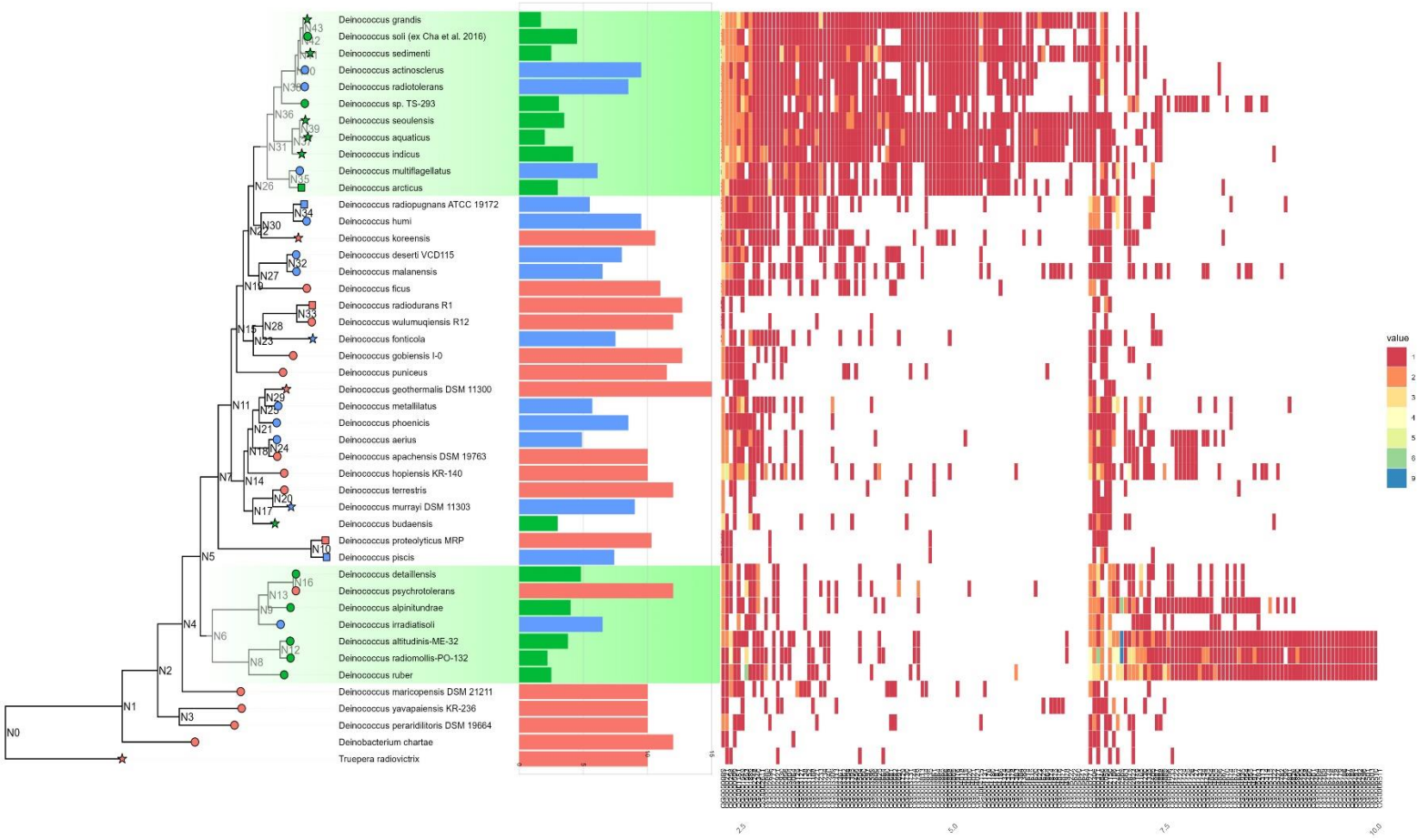

Figure 3-2. The presence-absence heatmap of significant OGs

The presence and absence of correlated orthogroups are mapped on the phylogenomic tree. The highlighted clades, N31 and N6, show the sensitive clades found in the phylogenomic analysis. The heatmap represents the presence and absence and copy number of orthogroups in each genome (each row) based on the orthology analysis. This heatmap highlights genes that might have a role in lower resistance among sensitive *Deinococcus* species.

3.4.4 Functional annotation of sensitivity orthogroups

Comparing COG functional categories between the two clades showed a completely different enrichment pattern in the function of identified genes in those clades. The function of most of the identified genes in both clades was assigned as unknown; N31 was significantly more, with 45% of genes annotated as unknown, compared to 20% in the N6 clade. In general, N6 had a higher abundance in the following categories: energy production (C), cell cycle control (D), coenzyme transport and metabolism (H), lipid transport and metabolism (I), replication, recombination and repair (L), and, cell wall and membrane (M). Clade N6 showed a higher abundance of genes involved in amino

acid transport and metabolism (E), Carbohydrate transport and metabolism (G), and posttranslational modification (O).

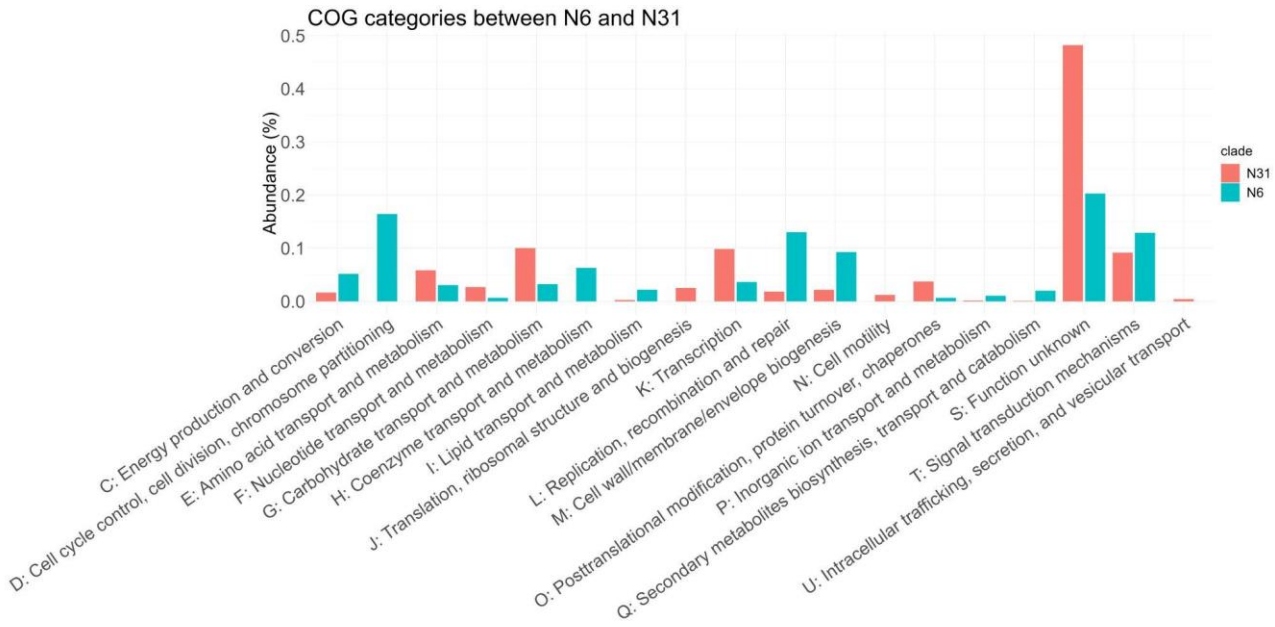

Figure 3-3. COG Functional annotation for correlated OGs.

This barplot shows the abundance of identified OGs across functional categories. OGs found in the N31 clade showed a high abundance of unknown gene families. Moreover, genes related to related transcription and post-translational modification were abundant in N31. Genes related to signal transduction were abundant in both clades. Significant OGs found in the N6 clade showed a high abundance in Cell cycle control, Coenzyme transport and metabolism, Replication recombination and repair, and cell wall biogenesis.

Within the N6 clade, the presence of genes belonging to the OG0004674 orthogroup (*atzB*) exhibited the strongest negative correlation with the D_{10} value, with a Pearson correlation coefficient of -0.71 and a p-value of 0.02 after FDR correction, based on a sample size of 6720 OGs (Supplementary Data S3-2). Superfamily and Gene3D databases identified a Metal-dependent hydrolase belonging to the Amidohydrolase family. The EggNOg database showed that this orthogroup includes sequences related to the repair enzyme 8-oxoguanine deaminase and has S-adenosyl-homocysteine (SAH) deaminase activity. The *atzB* was present in seven *Deinococcus* species, including four members of the N6 clade *D. ruber*, *D. radiomollis*, *D. altitudinis*, and *D. alpinitundrae* with D_{10} values ranging from 2.2 - 4 kGy, and three other species belonging to MR

group *D. metallilatus* (D_{10} = 5.7 kGy), *D.radiopugnans* (D_{10} = 5.5 kGy), *D. phoenicis* (D_{10} = 8.5).

The 93 orthogroups within N6 showed diverse functions, but there is a pattern in the enrichment of certain co-factors for these enzymes. Zinc Finger, NAD, FAD, Rieske [2Fe-2S] iron-sulphur domain superfamily, and Acyl-CoA were highly enriched among those orthogroups. Most of the orthogroups with stronger negative correlation with D_{10} values were present in four genomes: *D. ruber*, *D. radiomollis*, *D. altitudinis,* and *D. alpinitundrae.* The top ten orthogroups with the strongest negative correlation with radiation resistance (r = -0.71 to -0.66) were annotated as Metallo-dependent hydrolases (atzB), Acyl-CoA dehydrogenase/oxidase, NAD(P)-binding Rossmann-fold domains, MmgE/PrpD (Flavin-dependent oxidoreductase), NAD-dependent epimerase/dehydratase, Alpha/Beta hydrolase (Thiamine pyrophosphate enzyme), Acyl-CoA N-acyltransferase (Supplementary Data S3-4).

In the N31 clade, the five orthogroups with the strongest correlation (r = -75 to - 0.69) were hypothetical proteins and did not have any match in the above-mentioned protein databases. The other correlated genes are NAD(P)-binding domain superfamily (Rossmann-fold domains), FAD/NAD(P)-binding domain superfamily, and Metallo-dependent phosphatase-like, molybdopterin cofactor-binding domain, aldehyde oxidase/xanthine dehydrogenase, second molybdopterin binding domain. Many orthogroups in this clade belong to proteins of unknown function, and there was no match in multiple protein databases (Supplementary Data S3-5).

In the N6 clade, the presence and higher copy number of 93 orthogroups were significantly correlated with D_{10} values. Except for one gene family (zinc metallopeptidase (prsW - OG0002059), all orthogroups showed a negative Pearson correlation coefficient (Supplementary Data S3-4 and S3-5**)**. On the other hand, in the N31 clade, 92 orthogroups showed a negative correlation with higher D_{10} values. Only one orthogroup was shared between two sensitive clades (tadA) (OG0000092) - MafB19-like deaminase (PF00383 Cytidine and deoxycytidylate deaminase zinc-binding region), which is present in almost all *Deinococcus* genomes but has a higher

copy number in members of N6 and N31. No significant gene was detected from Spearman and Mann Whitney using only the gene presence-absence matrix.

3.4.5 Prediction of radiation resistance from genomic content

We sought to test whether the enrichment of identified orthogroups can predict the D_{10} values of *Deinococcaceae* members using their genomes. To do so, we searched for those orthogroups within every single *Deinococcus* genome using DIAMOND blast and counted the number of blast hits on each genome. Then, we created a linear model by using the number of blast hits on every single genome and calculated the regression for the sum of the numbers of orthogroups.

The adjusted R-squared value for cumulative numbers of OGs in two sensitive clades was r^2 =0.5682 with a p-value of 1.38e-09 using DIAMOND with default setting. The Pearson correlation between D_{10} values and the number of orthogroups was -0.76, and the p-value was 1.38e-09.

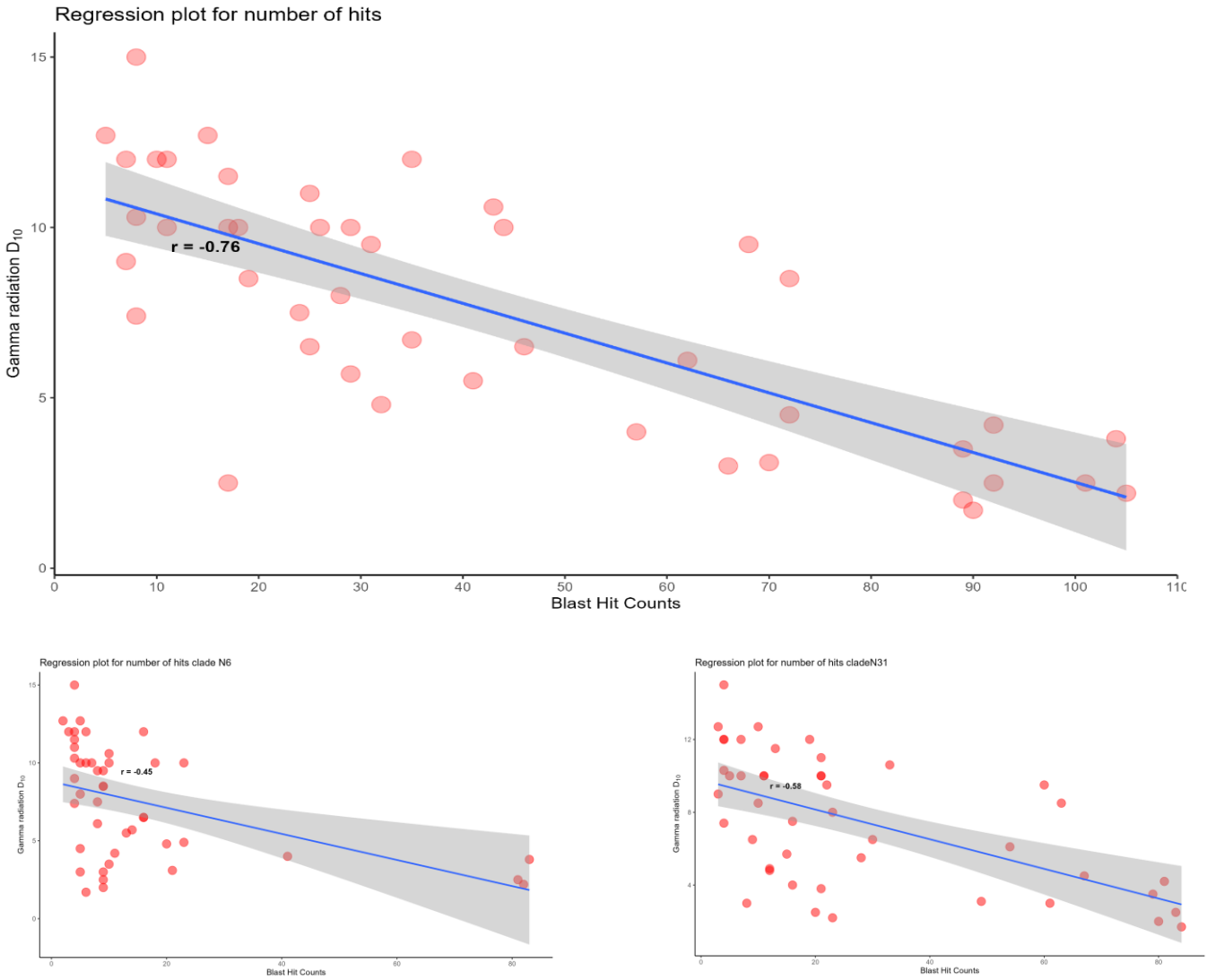

Figure 3-4. Regression plot for numbers of orthogroup hits.

The regression plot for the number of unique hits for each genome (x-axis) shows a negative correlation with radiation resistance levels. In other words, genomes with a higher number of those orthogroups show lower D_{10} values. The top panel shows a regression plot for a cumulative number of hits for both clades. The bottom left panels show each clade separately, N6 left and N31.

3.5 Discussion

Radiation resistance is often regarded as a defining characteristic of the *Deinococcaceae* family, mainly due to the extreme radiation resistance capabilities of *D. radiodurans* and some other members (Battista, 2016; Copeland et al., 2012; Daly, 2023; Murray, 1992; Rosenberg, 2014). This notion has led to the application of gamma radiation as a pre-cultivation enrichment method for environmental samples and the indiscriminate eradication of all radiosensitive cells in the search for *Deinococcus* species (Copeland et al., 2012; Rainey et al., 2005; Yin et al., 2022). Although this method has demonstrated its effectiveness in cultivating hundreds of *Deinococcus* isolates, it has likely introduced a bias in our knowledge about the entire genus's members. The evidence for this bias is that most radiosensitive *Deinococcus* species were isolated without gamma-radiation enrichment or very low doses (Callegan et al., 2008; Jeon et al., 2016). This evidence suggests the existence of *Deinococcus* species that may be highly radiosensitive, potentially as sensitive as *T. thermophilus* or *E. coli*. This finding challenges our current understanding of the ecology and evolution of the *Deinococcaceae* family, as it highlights the potential presence of radiosensitive *Deinococcus* species that may have been overlooked due to the predominant focus on radioresistant strains and isolation biases.

3.5.1 Phylogenomic patterns in radiation resistance

Single-copy orthogroups are commonly used to construct a phylogenomic tree because they minimise the potential for including gene duplication artefacts in the phylogenomic inference (Nakhleh, 2013). In this study, we used a dataset comprising 44 *Deinococcus* species with newly sequenced and publicly available genomes and their ecological and phenotypic data on radiation resistance in the literature to show the considerable diversity in radiation resistance levels in this genus. The construction of a species tree using single-copy core genes revealed an evident evolutionary pattern, grouping radiosensitive species into two distinct clades (N6, N31) and one single branch (*D. budaensis*)(Figure 3-1). Informed by this evolutionary pattern and by statistical evaluation of the accessory genome – a determinant of species or genus diversity – we identified 188 orthogroups exhibiting a significant correlation with radiation resistance

capacity, denoted by D_{10} values of related species. Except for one orthogroup, all identified gene families showed a strong negative correlation between the presence of these genes and D_{10} values, indicating that the acquisition of these genes led to radiation sensitivity.

Several comparative genomics studies have explored radiation resistance mechanisms in *Deinococcus* and other organisms (Lim et al., 2018; Makarova et al., 2007; Makarova & Daly, 2010). However, they fell short of identifying a pattern in the radiation resistance phenotype using the genomic content, and "radiation resistance consistently remained a phenotype without genotype" (Daly, 2023). Previous studies have focused on genes identified through transcriptomics and gene knockout experiments in *D. radiodurans* as model organisms and, to a limited extent, on a few other radioresistant species (Lim et al., 2018). Here, we extend the search for such a pattern in the diversity of radiation resistance to the pangenome of the *Deinococcus* genus with a focus on accessory genes which signify the diversity of phenotypes in a group of related organisms (Motyka-Pomagruk et al., 2020; Tettelin et al., 2008; Tettelin & Medini, 2020). To our knowledge, this is the first analysis of radiation resistance phenotype agnostic from known genes and using all genes that could be the result of HGT.

Previous studies have indicated that *D. radiodurans* enhances its resistance to oxidative stress by minimising endogenous ROS generation (Slade & Radman, 2011). These strategies include coding for fewer cytochromes and iron sulphur clusters and NADPH- and NADH-dependent enzymes (Ghosal et al., 2005; Liu et al., 2003). In line with these previous discoveries, our data revealed that the orthogroups identified in the two radiation-sensitive clades predominantly consist of proteins with co-factors such as zinc, iron-sulfur clusters, NAD, FAD2, and Acetyl-CoA-dependent enzymes.

3.5.2 The ecology of *Deinococcus* might impact its radiation resistance.

The ecological data of *Deinococcus* species in our dataset indicates that most members in the radiosensitive clades were isolated from cold environments, such as Antarctica and Alpine regions. Additionally, the genomic flexibility of *Deinococcus,* as discussed in the last chapter, can be seen as a strategy for adapting to new environments by

acquiring genes that maximise energy production in the absence of readily available organic substances. Therefore, the acquisition of specific genes that increase endogenous ROS production could be a significant driver of radiation resistance diversity in the *Deinococcus* genus.

It is a long-standing theory that the radiation resistance in *Deinococcus* cannot be an adaptation to ionising radiation because there is no selective pressure in the natural world (Battista, 1997; Mattimore & Battista, 1996). Instead, radiation resistance is considered a byproduct of other reactive oxygen species (ROS)-inducing factors, such as desiccation, which is prevalent on Earth, especially in desert soils (Fredrickson et al., 2008). This study supports this hypothesis by demonstrating that when *Deinococcus* species are not under oxidative stress, they can acquire genes that enhance their energy production efficiency and employ other mechanisms for survival in their environment. However, acquiring these genes can become deleterious in the presence of oxidative agents like ionising radiation and desiccation.

3.5.3 Radiation resistance is most likely an ancestral trait

The phylogenomic tree and presence-absence heatmap of significantly correlated genes in

Figure 3-2 indicate that radiation resistance is an ancestral trait in the *Deinococcus* genus. The evidence for this claim is that radiation resistance phenotype is also observed in *Truepera radiovictrix,* which is the only cultivated member of the *Trueperaceae* family and a distant relative of *Deinococcaceae* (Battista, 2016). A D_{10} value of 10 kGy was reported for *Truepera radiovictrix* (Albuquerque et al., 2005). Moreover, it seems that radiation resistance was conserved after the divergence of *Trueperaceae* and *Deinococcaceae* since it is also observed in *Deinobacterium*, the other genus in the *Deinococcaceae* family, which shares a direct common ancestor with the *Deinococcus* genus. *Deinobacterium chartae,* the only member of this genus was isolated from a Finnish paper mill and has a D_{10} value of approximately 10 kGy (Ekman et al., 2011). Due to the unavailability of any other cultivated species in the *Deinobacterium* and *Truepera* genera, it is difficult to interpret this with high certainty, but it is logical to

consider that radiation resistance is an ancestral trait because it was transferred vertically.

3.5.4 Oxidative factors may cause radiation sensitivity

The statistical evaluation of accessory genes and radiation resistance in this study revealed an enrichment of oxidative co-factors only present in the the two radiosensitive clades of *Deinococcus* species, which are predominantly psychrophilic. The observation that radiation sensitivity occurred in multiple events confirms the genetic flexibility of the *Deinococcus* species. Even though the identified genes in the N6 and N31 clades are not overlapping (Figure 3-2), they predominantly encode redox-related co-enzymes, which have been shown to be reduced in the genome of highly radioresistant *D. radiodurans* (Ghosal et al., 2005).

This discovery challenges the conventional notion that radiation resistance is the defining characteristic of *Deinococcaceae* and suggests that even though resistance to extreme radiation is an ancestral genome, it can be lost in favour of acquiring new traits that aid *Deinococcus* species in adapting to environments where resistance to oxidative damage is not essential. We suggest extreme radiation resistance should be regarded as a phenotype only in some *Deinococcus* species, and alternative cultivation methods should be developed instead of gamma radiation as a pre-culture treatment. Previous studies have explored designing specific culture media to accommodate easier cultivation of *Deinococcus* (Bornot et al., 2015; He, 2009). Focusing on the basic physiological needs of *Deinococcus* can solve this problem.

3.5.5 Prediction of radiation resistance

It has been repeatedly reiterated that radiation resistance cannot be predicted using genomic data (Daly, 2023; Daly et al., 2010; Sharma et al., 2017). Our linear model using identified orthogroups in both clades showed an adjusted R-squared value of 0.56, which suggests it could predict radiation resistance about 56% of the time. We note that the model found in this study has some limitations. First, there is no standardised method for precise measurement of radiation resistance among *Deinococcus* species, and a standard error should be considered for the reported D_{10}. However, the majority

of collected D_{10} values from the literature were measured in comparison with *D. radiodurans* as positive and *E.coli* as negative controls, which indicates their measurements are close to reality. Another limitation of this study is that some of these genes can be phylogenetic signals that are not related to radiation resistance. These genes are hypothetic targets for further study, and experimental studies like inserting those sequences into radiation-resistant species like *D. radiodurans* can narrow down those genes to a definitive set of genes for radiation sensitivity.

The two clades showed a similar pattern in the presence of redox-related co-factors like metal ions, NAD, FAD, thiamin, and Acyl-CoA. Subsequently, we counted the number of those genes in each *Deinococcus* genus and made a linear regression model to model the correlation of the higher number of those genes with lower D10 values. Finally, we used the model to predict the radiation resistance values in the genome of unknown *Deinococcus* species.

3.6 Conclusion

Radiation resistance in *Deinococcus* has been studied for 68 years since Arthur Anderson cultivated it for the first time (Anderson et al., 1956), but radiation resistance diversity remained an open question. Our study fills key gaps in our knowledge of this phenotype and shows that gene gain in specific environments can select for radiation sensitivity in the lack of oxidative stress. Our study proposes exploring novel physiological methods for the enrichment of *Deinococcus* species as a complementary method instead of only using ionising radiation, which potentially created a bias in our knowledge about this enigmatic group of bacteria.

In conclusion, addressing the misconception that the *Deinococcaceae* are radioresistant is important because changing our approach towards this group of bacteria can help us to find new *Deinococcus* isolates and provide an opportunity to discover a more diverse population of *Deinococcaceae*, some of which might be more sensitive than the most sensitive known species, *D. persicinus,* with a D_{10} value of 1.5 kGy (Jeon et al., 2016). This higher diversity can extend the pangenome of *Deinococcus* and help us better

understand the true ecology of the *Deinococcus* genus independent from radiation resistance and focus on the strategies it uses for adaptation to diverse environments.

3.7 References

Albuquerque, L., Simões, C., Nobre, M. F., Pino, N. M., Battista, J. R., Silva, M. T., Rainey, F. A., & de Costa, M. S. (2005). Truepera radiovictrix gen. Nov., sp. Nov., a new radiation resistant species and the proposal of Trueperaceae fam. Nov. *FEMS Microbiology Letters*, *247*(2), 161–169. https://doi.org/10.1016/j.femsle.2005.05.002

Anderson, A., Nordon, H., Cain, R. F., Parrish, G., Duggan, D., Nordan, H., Parish, G., & Cullum-Dugan, D. (1956). Studies on a radio-resistant micrococcus. I. Isolation, morphology, cultural characteristics, and resistance to gamma radiation. *Food Technology*, *10*, 575–578.

Baek, K., Chung, E. J., Choi, G.-G., Kim, M.-K., Lim, S., & Choi, A. (2018). *Deinococcus koreensis* sp. Nov., a gamma radiation-resistant bacterium isolated from river water. *International Journal of Systematic and Evolutionary Microbiology*, *68*(8), 2545–2550. https://doi.org/10.1099/ijsem.0.002872

Battista, J. R. (1997). AGAINST ALL ODDS: The Survival Strategies of *Deinococcus radiodurans*. In *Annu. Rev. Microbiol* (Vol. 51, pp. 203–227). www.annualreviews.org

Battista, J. R. (2016). *Deinococcus* – Thermus Group. In *Encyclopedia of Life Sciences* (pp. 1–12). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470015902.a0021151

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Bentchikou, E., Servant, P., Coste, G., & Sommer, S. (2010). A Major Role of the RecFOR Pathway in DNA Double-Strand-Break Repair through ESDSA in *Deinococcus radiodurans*. *PLOS Genetics*, *6*(1), e1000774. https://doi.org/10.1371/journal.pgen.1000774

Bornot, J., Molina-Jouve, C., Uribelarrea, J.-L., & Gorret, N. (2015). Quantitative Characterisation of the Growth of *Deinococcus geothermalis* DSM-11302: Effect of Inoculum Size, Growth Medium and Culture Conditions. *Microorganisms*, *3*(3), 441–463. https://doi.org/10.3390/microorganisms3030441

Brooks, B. W., & Murray, R. G. E. (1981). Nomenclature for 'Micrococcus *radiodurans*' and other radiation-resistant cocci: *Deinococcaceae* fam. Nov. And *Deinococcus* gen. Nov., including five species. *International Journal of Systematic Bacteriology*, *31*(3), 353–360. https://doi.org/10.1099/00207713-31-3-353

Bruch, E. M., de Groot, A., Un, S., & Tabares, L. C. (2015). The effect of gamma-ray irradiation on the Mn(II) speciation in *Deinococcus radiodurans* and the potential role of Mn(II)-orthophosphates. *Metallomics : Integrated Biometal Science*, *7*(5), 908–916. https://doi.org/10.1039/c5mt00009b

Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, *12*(1), Article 1. https://doi.org/10.1038/nmeth.3176

Callegan, R. P., Noble, M. F., McTernan, P. M., Battista, J. R., Navarro-González, R., McKay, C. P., da Costa, M. S., & Rainey, F. A. (2008). Description of four novel psychrophilic, ionising radiation-sensitive *Deinococcus* species from alpine environments. *International Journal of Systematic and Evolutionary Microbiology*, *58*(5), 1252–1258. https://doi.org/10.1099/ijs.0.65405-0

Cao, Y., & Li, H. (2011). Dynamics of Protein Folding and Cofactor Binding Monitored by Single-Molecule Force Spectroscopy. *Biophysical Journal*, *101*(8), 2009–2017. https://doi.org/10.1016/j.bpj.2011.08.051

Cha, S., Srinivasan, S., Seo, T., & Kim, M. K. (2014a). *Deinococcus radiotolerans* sp. Nov., a gamma-radiation-resistant bacterium isolated from gamma ray-irradiated soil. *Antonie van Leeuwenhoek*, *105*(1), 229–235. https://doi.org/10.1007/s10482-013-0069-0

Cha, S., Srinivasan, S., Seo, T., & Kim, M. K. (2014b). *Deinococcus soli* sp. Nov., a gamma-radiation-resistant bacterium isolated from rice field soil. *Current Microbiology*, *68*(6), 777–783. https://doi.org/10.1007/s00284-014-0542-7

Chandrangsu, P., Rensing, C., & Helmann, J. D. (2017). Metal homeostasis and resistance in bacteria. In *Nature Reviews Microbiology*. https://doi.org/10.1038/nrmicro.2017.15

Copeland, A., Zeytun, A., Yassawong, M., Nolan, M., Lucas, S., Hammon, N., Deshpande, S., Cheng, J. F., Han, C., Tapia, R., Goodwin, L. A., Pitluck, S., Mavromatis, K., Liolios, K., Pagani, I., Ivanova, N., Mikhailova, N., Pati, A., Chen, A., … Lapidus, A. (2012). Complete genome sequence of the orange-red pigmented, radioresistant *Deinococcus* proteolyticus type species (MRPT). *Standards in Genomic Sciences*, *6*(2), 240–250. https://doi.org/10.4056/sigs.2756060

Cox, M. M., & Battista, J. R. (2005). *Deinococcus radiodurans*—The consummate survivor. *Nature Reviews Microbiology*, *3*(11), 882–892. https://doi.org/10.1038/nrmicro1264

Daly, M. J. (2009). A new perspective on radiation resistance based on *Deinococcus radiodurans*. *Nature Reviews Microbiology*, *7*(3), 237–245. https://doi.org/10.1038/nrmicro2073

Daly, M. J. (2023). The scientific revolution that unraveled the astonishing DNA repair capacity of the *Deinococcaceae*: 40 years on. *Canadian Journal of Microbiology*. https://doi.org/10.1139/cjm-2023-0059

Daly, M. J., Gaidamakova, E. K., Matrosova, V. Y., Kiang, J. G., Fukumoto, R., Lee, D. Y., Wehr, N. B., Viteri, G. A., Berlett, B. S., & Levine, R. L. (2010). Small-molecule antioxidant proteome-shields in *Deinococcus radiodurans*. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0012570

Daly, M. J., Gaidamakova, E. K., Matrosova, V. Y., Vasilenko, A., Zhai, M., Venkateswaran, A., Hess, M., Omelchenko, M. V., Kostandarithes, H. M., Makarova, K. S., Wackett, L. P., Fredrickson, J. K., & Ghosal, D. (2004). Accumulation of Mn(II) in *Deinococcus radiodurans* Facilitates Gamma-Radiation Resistance. *Science*, *306*(5698), 1025–1028. https://doi.org/10.1126/science.1103185

de Groot, A., Chapon, V., Servant, P., Christen, R., Fischer-Le Saux, M., Sommer, S., & Heulin, T. (2005). *Deinococcus deserti* sp. Nov., a gamma-radiation-tolerant bacterium isolated from the Sahara Desert. *International Journal of Systematic and Evolutionary Microbiology*, *55*(6), 2441–2446. https://doi.org/10.1099/ijs.0.63717-0

Dong, N., Li, H.-R., Yuan, M., Zhang, X.-H., & Yu, Y. (2015). *Deinococcus* antarcticus sp. Nov., isolated from soil. *INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY*, *65*(Pt 2), 331–335. https://doi.org/10.1099/ijs.0.066324-0

Ekman, J. V., Raulio, M., Busse, H.-J., Fewer, D. P., & Salkinoja-Salonen, M. (2011). Deinobacterium chartae gen. Nov., sp. Nov., an extremely radiation-resistant, biofilm-forming bacterium isolated from a Finnish paper mill. *International Journal of Systematic and Evolutionary Microbiology*, *61*(3), 540–548. https://doi.org/10.1099/ijs.0.017970-0

Ferreira, A. C., Nobre, M. F., Rainey, F. A., Silva, M. T., Wait, R., Burghardt, J., Chung, A. P., & da Costa, M. S. (1997). *Deinococcus geothermalis* sp. Nov. And *Deinococcus* murrayi sp. Nov., two extremely radiation-resistant and slightly thermophilic species from hot springs. *International Journal of Systematic Bacteriology*, *47*(4), 939–947. https://doi.org/10.1099/00207713-47-4-939

Fredrickson, J. K., Li, S. W., Gaidamakova, E. K., Matrosova, V. Y., Zhai, M., Sulloway, H. M., Scholten, J. C., Brown, M. G., Balkwill, D. L., & Daly, M. J. (2008). Protein oxidation: Key to bacterial desiccation resistance? *The ISME Journal*, *2*(4), Article 4. https://doi.org/10.1038/ismej.2007.116

Gaidamakova, E. K., Sharma, A., Matrosova, V. Y., Grichenko, O., Volpe, R. P., Tkavc, R., Conze, I. H., Klimenkova, P., Balygina, I., Horne, W. H., Gostinčar, C., Chen, X., Makarova, K. S., Shuryak, I., Srinivasan, C., Jackson-Thompson, B., Hoffman, B. M., & Daly, M. J. (2022). Small-Molecule Mn Antioxidants in Caenorhabditis elegans and *Deinococcus radiodurans* Supplant MnSOD Enzymes during Aging and Irradiation. *mBio*, *13*(1), e0339421. https://doi.org/10.1128/mbio.03394-21

Ghosal, D., Omelchenko, M. V., Gaidamakova, E. K., Matrosova, V. Y., Vasilenko, A., Venkateswaran, A., Zhai, M., Kostandarithes, H. M., Brim, H., Makarova, K. S., Wackett, L. P., Fredrickson, J. K., & Daly, M. J. (2005). How radiation kills cells: Survival of *Deinococcus radiodurans* and Shewanella oneidensis under oxidative stress⋆. *FEMS Microbiology Reviews*, *29*(2), 361–375. https://doi.org/10.1016/j.fmrre.2004.12.007

Gupta, P., Gayen, M., Smith, J. T., Gaidamakova, E. K., Matrosova, V. Y., Grichenko, O., Knollmann-Ritschel, B., Daly, M. J., Kiang, J. G., & Maheshwari, R. K. (2016). MDP: A *Deinococcus* Mn2+-Decapeptide Complex Protects Mice from Ionizing Radiation. *PloS One*, *11*(8), e0160575. https://doi.org/10.1371/journal.pone.0160575

He, Y. (2009). High cell density production of *Deinococcus radiodurans* under optimised conditions. *Journal of Industrial Microbiology and Biotechnology*, *36*(4), 539–546. https://doi.org/10.1007/s10295-008-0524-5

Horne, W. H., Volpe, R. P., Korza, G., DePratti, S., Conze, I. H., Shuryak, I., Grebenc, T., Matrosova, V. Y., Gaidamakova, E. K., Tkavc, R., Sharma, A., Gostinčar, C., Gunde-Cimerman, N., Hoffman, B. M., Setlow, P., & Daly, M. J. (2022). Effects of Desiccation and Freezing on Microbial Ionizing Radiation Survivability: Considerations for Mars Sample Return. *Astrobiology*, *22*(11), 1337–1350. https://doi.org/10.1089/ast.2022.0065

Hutchinson, F. (1966). The molecular basis for radiation effects on cells. *Cancer Research*, *26*(9), 2045–2052.

Im, S., Joe, M., Kim, D., Park, D.-H., & Lim, S. (2013). Transcriptome analysis of salt-stressed *Deinococcus radiodurans* and characterisation of salt-sensitive mutants. *Research in Microbiology*, *164*(9), 923–932. https://doi.org/10.1016/j.resmic.2013.07.005

Im, W.-T., Jung, H.-M., Ten, L. N., Kim, M. K., Bora, N., Goodfellow, M., Lim, S., Jung, J., & Lee, S.-T. (2008). *Deinococcus* aquaticus sp. Nov., isolated from fresh water, and *Deinococcus* caeni sp. Nov., isolated from activated sludge. *International Journal of Systematic and Evolutionary Microbiology*, *58*(10), 2348–2353. https://doi.org/10.1099/ijs.0.64082-0

Imlay, J. A. (2003). Pathways of Oxidative Damage. *Annual Review of Microbiology*, *57*(1), 395–418. https://doi.org/10.1146/annurev.micro.57.030502.090938

Innamorati, K. A., Earl, J. P., Aggarwal, S. D., Ehrlich, G. D., & Hiller, N. L. (2020). The Bacterial Guide to Designing a Diversified Gene Portfolio. In H. Tettelin & D. Medini (Eds.), *The* Pan-genome*: Diversity, Dynamics and Evolution of Genomes* (pp. 51–87). Springer International Publishing. https://doi.org/10.1007/978-3-030-38281-0_3

Jeon, S. H., Kang, M. S., Joo, E. S., Kim, E. B., Lim, S., Jeong, S. W., Jung, H. Y., Srinivasan, S., & Kim, M. K. (2016). *Deinococcus* persicinus sp. Nov., a radiationresistant bacterium from soil. *International Journal of Systematic and Evolutionary Microbiology*, *66*(12), 5077–5082. https://doi.org/10.1099/ijsem.0.001473

Kim, D. U., Jang, J. H., Kang, M.-S., Kim, J.-Y., Zhang, J., Lim, S., & Kim, M. K. (2018). *Deinococcus* irradiati*soli* sp. Nov., isolated from gamma ray-irradiated soil. *International Journal of Systematic and Evolutionary Microbiology*, *68*(10), 3232–3236. https://doi.org/10.1099/ijsem.0.002968

Kim, D. U., Lee, H., Lee, J. H., Ahn, J. H., Lim, S., Jeong, S., Park, S. Y., Seong, C. N., & Ka, J. O. (2015). *Deinococcus* metallilatus sp. Nov. And *Deinococcus* carri sp. Nov., isolated from a car air-conditioning system. *International Journal of Systematic and Evolutionary Microbiology*, *65*(9), 3175–3182. https://doi.org/10.1099/ijsem.0.000396

Kim, D. U., Lee, H., Lee, S., Park, S., Yoon, J. H., Zhao, L., Kim, M.-K., Ahn, J.-H., & Ka, J.-O. (2018). *Deinococcus* multiflagellatus sp. Nov., isolated from a car air-conditioning system. *Antonie van Leeuwenhoek*, *111*(4), 619–627. https://doi.org/10.1007/s10482-017-0982-8

Kim, E. B., Kang, M. S., Joo, E. S., Jeon, S. H., Jeong, S. W., Lim, S. Y., Jung, H. Y., Srinivasan, S., & Kim, M. K. (2017). *Deinococcus* ruber sp. Nov., a radiation-resistant bacterium isolated from soil. *International Journal of Systematic and*

Evolutionary Microbiology, *67*(1), 72–76. https://doi.org/10.1099/ijsem.0.001567

Lai, W. A., Kämpfer, P., Arun, A. B., Shen, F. T., Huber, B., Rekha, P. D., & Young, C. C. (2006). *Deinococcus* ficus sp. Nov., isolated from the rhizosphere of Ficus religiosa L. *International Journal of Systematic and Evolutionary Microbiology*, *56*(4), 787–791. https://doi.org/10.1099/ijs.0.64007-0

Lee, J. H., Jung, J.-H., Kim, M.-K., & Lim, S. (2022). *Deinococcus taeanensis* sp. Nov., a Radiation-Resistant Bacterium Isolated from a Coastal Dune. *Current Microbiology*, *79*(11), 334. https://doi.org/10.1007/s00284-022-03044-8

Lee, J. J., Lee, H. J., Jang, G. S., Yu, J. M., Cha, J. Y., Kim, S. J., Lee, E. B., & Kim, M. K. (2013). *Deinococcus* swuensis sp. Nov., a gamma-radiation-resistant bacterium isolated from soil. *Journal of Microbiology*, *51*(3), 305–311. https://doi.org/10.1007/s12275-013-3023-y

Lee, J. J., Lee, Y. H., Park, S. J., Lim, S., Jeong, S. W., Lee, S.-Y., Park, S., Choi, H.-W., Kim, M. K., & Jung, H.-Y. (2016). *Deinococcus sedimenti* sp. Nov. Isolated from river sediment. *Journal of Microbiology*, *54*(12), 802–808. https://doi.org/10.1007/s12275-016-6361-8

Lee, J. J., Srinivasan, S., Lim, S., Joe, M., Im, S., & Kim, M. K. (2015). *Deinococcus* puniceus sp. Nov., a Bacterium Isolated from Soil-Irradiated Gamma Radiation. *Current Microbiology*, *70*(4), 464–469. https://doi.org/10.1007/s00284-014-0748-8

Lee, J.-J., Lee, Y.-H., Park, S.-J., Lim, S., Jeong, S.-W., Lee, S.-Y., Cho, Y.-J., Kim, M. K., & Jung, H.-Y. (2016). *Deinococcus* seoulensis sp. Nov., a bacterium isolated from sediment at Han River in Seoul, Republic of Korea. *Journal of Microbiology*, *54*(8), 537–542. https://doi.org/10.1007/s12275-016-6253-y

Lim, S., Jung, J.-H., Blanchard, L., & de Groot, A. (2018). Conservation and diversity of radiation and oxidative stress resistance mechanisms in *Deinococcus* species. *FEMS Microbiology Reviews*, *43*(1), 19–52. https://doi.org/10.1093/femsre/fuy037

Liu, Y., Zhou, J., Omelchenko, M. V., Beliaev, A. S., Venkateswaran, A., Stair, J., Wu, L., Thompson, D. K., Xu, D., Rogozin, I. B., Gaidamakova, E. K., Zhai, M., Makarova, K. S., Koonin, E. V., & Daly, M. J. (2003). Transcriptome dynamics of *Deinococcus radiodurans* recovering from ionising radiation. *Proceedings of the National Academy of Sciences*, *100*(7), 4191–4196. https://doi.org/10.1073/pnas.0630387100

Makarova, K. S., Aravind, L., Wolf, Y. I., Tatusov, R. L., Minton, K. W., Koonin, E. V., & Daly, M. J. (2001). Genome of the Extremely Radiation-Resistant Bacterium *Deinococcus radiodurans* Viewed from the Perspective of Comparative Genomics. *Microbiology and Molecular Biology Reviews*, *65*(1), 44–79. https://doi.org/10.1128/mmbr.65.1.44-79.2001

Makarova, K. S., & Daly, M. J. (2010). Comparative Genomics of Stress Response Systems in *Deinococcus* Bacteria. In *Bacterial Stress Responses* (pp. 445–457). John Wiley & Sons, Ltd. https://doi.org/10.1128/9781555816841.ch27

Makarova, K. S., Omelchenko, M. V., Gaidamakova, E. K., Matrosova, V. Y., Vasilenko, A., Zhai, M., Lapidus, A., Copeland, A., Kim, E., Land, M.,

Mavromatis, K., Pitluck, S., Richardson, P. M., Detter, C., Brettin, T., Saunders, E., Lai, B., Ravel, B., Kemner, K. M., … Daly, M. J. (2007). *Deinococcus geothermalis*: The Pool of Extreme Radiation Resistance Genes Shrinks. *PLoS ONE*, *2*(9), e955. https://doi.org/10.1371/journal.pone.0000955

Makk, J., Enyedi, N. T., Tóth, E., Anda, D., Szabó, A., Felföldi, T., Schumann, P., Mádl-Szőnyi, J., & Borsodi, A. K. (2019). *Deinococcus* fonticola sp. Nov., isolated from a radioactive thermal spring in Hungary. *International Journal of Systematic and Evolutionary Microbiology*, *69*(6), 1724–1730. https://doi.org/10.1099/ijsem.0.003383

Makk, J., Tóth, E. M., Anda, D., Pál, S., Schumann, P., Kovács, A. L., Mádl-Szőnyi, J., Márialigeti, K., & Borsodi, A. K. (2016). *Deinococcus* budaensis sp. Nov., a mesophilic species isolated from a biofilm sample of a hydrothermal spring cave. *International Journal of Systematic and Evolutionary Microbiology*, *66*(12), 5345–5351. https://doi.org/10.1099/ijsem.0.001519

Mattimore, V., & Battista, J. R. (1996). Radiation resistance of *Deinococcus radiodurans*: Functions Necessary To Survive Ionising Radiation Are Also Necessary To Survive Prolonged Desiccation. In *JOURNAL OF BACTERIOLOGY* (3; Vol. 178, pp. 633–637). http://jb.asm.org/

Motyka-Pomagruk, A., Zoledowska, S., Misztak, A. E., Sledz, W., Mengoni, A., & Lojkowska, E. (2020). Comparative genomics and pangenome-oriented studies reveal high homogeneity of the agronomically relevant enterobacterial plant pathogen Dickeya solani. *BMC Genomics*, *21*(1), 449. https://doi.org/10.1186/s12864-020-06863-w

Murray, R. G. E. (1992). The Family *Deinococcaceae*. In *The Prokaryotes* (pp. 3732–3744). Springer New York. https://doi.org/10.1007/978-1-4757-2191-1_42

Nakhleh, L. (2013). Computational approaches to species phylogeny inference and gene tree reconciliation. *Trends in Ecology & Evolution*, *28*(12), 719–728. https://doi.org/10.1016/j.tree.2013.09.004

Narasimha, A., & Basu, B. (2021). New insights into the activation of Radiation Desiccation Response regulon in *Deinococcus radiodurans*. *Journal of Biosciences*, *46*(1), 10. https://doi.org/10.1007/s12038-020-00123-5

Oyaizu, H., Stackebrandt, E., & Schleifer, K. H. (1987). A radiation-resistant rod-shaped bacterium, Deinobacter grandis gen. Nov., sp. Nov., with peptidoglycan containing ornithine. *International Journal of Systematic Bacteriology*, *37*(1), 62–67. https://doi.org/10.1099/00207713-37-1-62

Rainey, F. A., Ferreira, M., Nobre, M. F., Ray, K., Bagaley, D., Earl, A. M., Battista, J. R., Gómez-Silva, B., McKay, C. P., & da Costa, M. S. (2007). *Deinococcus* peraridilitoris sp. Nov., isolated from a coastal desert. *International Journal of Systematic and Evolutionary Microbiology*, *57*(7), 1408–1412. https://doi.org/10.1099/ijs.0.64956-0

Rainey, F. A., Ray, K., Ferreira, M., Gatz, B. Z., Nobre, M. F., Bagaley, D., Rash, B. A., Park, M. J., Earl, A. M., Shank, N. C., Small, A. M., Henk, M. C., Battista, J. R., Kämpfer, P., & Da Costa, M. S. (2005). Extensive diversity of ionising-radiation-resistant bacteria recovered from Sonoran Desert soil and description of nine new species of the genus *Deinococcus* obtained from a single soil sample.

Applied and Environmental Microbiology, *71*(9), 5225–5235. https://doi.org/10.1128/AEM.71.9.5225-5235.2005

Rosenberg, E. (2014). The family deinococcaceae. In *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea* (pp. 613–615). Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-642-38954-2_127

Santos, S. P., Yang, Y., Rosa, M. T. G., Rodrigues, M. A. A., De La Tour, C. B., Sommer, S., Teixeira, M., Carrondo, M. A., Cloetens, P., Abreu, I. A., & Romão, C. V. (2019). The interplay between Mn and Fe in *Deinococcus radiodurans* triggers cellular protection during paraquat-induced oxidative stress. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-53140-2

Sharma, A., Gaidamakova, E. K., Grichenko, O., Matrosova, V. Y., Hoeke, V., Klimenkova, P., Conze, I. H., Volpe, R. P., Tkavc, R., Gostinčar, C., Gunde-Cimerman, N., Diruggiero, J., Shuryak, I., Ozarowski, A., Hoffman, B. M., Daly, M. J., Designed, M. J. D., & Performed, A. O. (2017). *Across the tree of life, radiation resistance is governed by antioxidant Mn 2+, gauged by paramagnetic resonance*. https://doi.org/10.1073/pnas.1713608114

Shashidhar, R., & Bandekar, J. R. (2009). *Deinococcus* piscis sp. Nov., a radiation-resistant bacterium isolated from a marine fish. *International Journal of Systematic and Evolutionary Microbiology*, *59*(11), 2714–2717. https://doi.org/10.1099/ijs.0.003046-0

Slade, D., & Radman, M. (2011). Oxidative Stress Resistance in *Deinococcus radiodurans*. *Microbiology and Molecular Biology Reviews*, *75*(1), 133–191. https://doi.org/10.1128/mmbr.00015-10

Srinivasan, S., Lee, J.-J., Lim, S., Joe, M., & Kim, M. K. (2012). *Deinococcus* humi sp. Nov., isolated from soil. *International Journal of Systematic and Evolutionary Microbiology*, *62*, 2844–2850. https://doi.org/10.1099/ijs.0.037234-0

Sun Joo, E., Jin Lee, J., Kang, M.-S., Lim, S., Jeong, S., Bit Kim, E., Hwa Jeon, S., Srinivasan, S., & Kyum Kim, M. (2016). *Deinococcus* actinosclerus sp. Nov., a novel bacterium isolated from soil of a rocky hillside. *International Journal of Systematic and Evolutionary Microbiology*, *66*(2), 1003–1008. https://doi.org/10.1099/ijsem.0.000825

Sun, Y., Zhang, T., Lu, B., Li, X., & Jiang, L. (2023). Application of cofactors in the regulation of microbial metabolism: A state of the art review. *Frontiers in Microbiology*, *14*. https://www.frontiersin.org/articles/10.3389/fmicb.2023.1145784

Suresh, K., Reddy, G. S. N., Sengupta, S., & Shivaji, S. (2004). *Deinococcus* indicus sp. Nov., an arsenic-resistant bacterium from an aquifer in West Bengal, India. *International Journal of Systematic and Evolutionary Microbiology*, *54*(2), 457–461. https://doi.org/10.1099/ijs.0.02758-0

Szabo, R., Bodolea, C., & Mocan, T. (2021). Iron, Copper, and Zinc Homeostasis: Physiology, Physiopathology, and Nanomediated Applications. *Nanomaterials*, *11*(11), 2958. https://doi.org/10.3390/nano11112958

Tanaka, M., Earl, A. M., Howell, H. A., Park, M. J., Eisen, J. A., Peterson, S. N., & Battista, J. R. (2004). Analysis of *Deinococcus radiodurans*'s transcriptional response to ionising radiation and desiccation reveals novel proteins that

contribute to extreme radioresistance. *Genetics*, *168*(1), 21–33. https://doi.org/10.1534/genetics.104.029249

Tettelin, H., & Medini, D. (Eds.). (2020). *The* Pan-genome*: Diversity, Dynamics and Evolution of Genomes*. Springer International Publishing. https://doi.org/10.1007/978-3-030-38281-0

Tettelin, H., Riley, D., Cattuto, C., & Medini, D. (2008). Comparative genomics: The bacterial pangenome. *Current Opinion in Microbiology*, *11*(5), 472–477. https://doi.org/10.1016/j.mib.2008.09.006

Tian, J., Wang, L., Liu, P., Geng, Y., Zhu, G., Zheng, R., Liu, Z., Zhao, Y., Yang, J., & Peng, F. (2019). *Deinococcus* psychrotolerans sp. Nov., isolated from soil on the South Shetland Islands, Antarctica. *International Journal of Systematic and Evolutionary Microbiology*, *69*(12), 3696–3701. https://doi.org/10.1099/ijsem.0.003484

Vaishampayan, P., Roberts, A. H. ayden, Augustus, A., Pukall, R., Schumann, P., Schwendner, P., Mayilraj, S., Salmassi, T., & Venkateswaran, K. (2014). *Deinococcus* phoenicis sp. Nov., an extreme ionising-radiation-resistant bacterium isolated from the Phoenix Lander assembly facility. *International Journal of Systematic and Evolutionary Microbiology*, *64*(Pt 10), 3441–3446. https://doi.org/10.1099/ijs.0.063107-0

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … van Mulbregt, P. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), Article 3. https://doi.org/10.1038/s41592-019-0686-2

Vornhagen, J., Roberts, E. K., Unverdorben, L., Mason, S., Patel, A., Crawford, R., Holmes, C. L., Sun, Y., Teodorescu, A., Snitkin, E. S., Zhao, L., Simner, P. J., Tamma, P. D., Rao, K., Kaye, K. S., & Bachman, M. A. (2022). Combined comparative genomics and clinical modeling reveals plasmid-encoded genes are independently associated with Klebsiella infection. *Nature Communications*, *13*(1), Article 1. https://doi.org/10.1038/s41467-022-31990-1

Wang, J.-J., Wu, S.-G., Chen, Q., Sheng, D.-H., Du, Z.-J., & Li, Y.-Z. (2020). *Deinococcus* terrestris sp. Nov., a gamma ray- and ultraviolet-resistant bacterium isolated from soil. *International Journal of Systematic and Evolutionary Microbiology*, *70*(9), 4993–5000. https://doi.org/10.1099/ijsem.0.004369

Wang, L.-G., Lam, T. T.-Y., Xu, S., Dai, Z., Zhou, L., Feng, T., Guo, P., Dunn, C. W., Jones, B. R., Bradley, T., Zhu, H., Guan, Y., Jiang, Y., & Yu, G. (2020). Treeio: An R Package for Phylogenetic Tree Input and Output with Richly Annotated and Associated Data. *Molecular Biology and Evolution*, *37*(2), 599–603. https://doi.org/10.1093/molbev/msz240

White, O., Eisen, J. A., Heidelberg, J. F., Hickey, E. K., Peterson, J. D., Dodson, R. J., Haft, D. H., Gwinn, M. L., Nelson, W. C., Richardson, D. L., Moffat, K. S., Qin, H., Jiang, L., Pamphile, W., Crosby, M., Shen, M., Vamathevan, J. J., Lam, P., McDonald, L., … Fraser, C. M. (1999). Genome sequence of the radioresistant

bacterium *Deinococcus radiodurans* R1. *Science (New York, N.Y.)*, *286*(5444), 1571–1577. https://doi.org/10.1126/science.286.5444.1571

Xu, X., Jiang, L., Zhang, Z., Shi, Y., & Huang, H. (2013). Genome Sequence of a Gamma- and UV-Ray-Resistant Strain, *Deinococcus* wulumuqiensis R12. *Genome Announcements*, *1*(3), 10.1128/genomea.00206-13. https://doi.org/10.1128/genomea.00206-13

Yang, Y., Itoh, T., Yokobori, S., Itahashi, S., Shimada, H., Satoh, K., Ohba, H., Narumi, I., & Yamagishi, A. (2009). *Deinococcus* aerius sp. Nov., isolated from the high atmosphere. *International Journal of Systematic and Evolutionary Microbiology*, *59*(8), 1862–1866. https://doi.org/10.1099/ijs.0.007963-0

Yin, L.-Z., Li, J.-L., Liu, Z.-T., Fang, B.-Z., Wang, P., Luo, X.-Q., Dong, L., Duan, L., Li, S.-H., & Li, W.-J. (2022). *Deinococcus* aestuarii sp. Nov. And *Deinococcus* aquaedulcis sp. Nov., two novel resistant bacteria isolated from pearl river estuary. *Antonie van Leeuwenhoek*, *115*(1), 59–68. https://doi.org/10.1007/s10482-021-01680-x

Yu, G. (2020). Using ggtree to Visualise Data on Tree-Like Structures. *Current Protocols in Bioinformatics*, *69*(1), e96. https://doi.org/10.1002/cpbi.96

Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T.-Y. (2017). ggtree: An r package for visualisation and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, *8*(1), 28–36. https://doi.org/10.1111/2041-210X.12628

Yuan, M., Zhang, W., Dai, S., Wu, J., Wang, Y., Tao, T., Chen, M., & Lin, M. (2009). *Deinococcus* gobiensis sp. Nov., an extremely radiation-resistant bacterium. *INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY*, *59*(6), 1513–1517. https://doi.org/10.1099/ijs.0.004523-0

Zahradka, K., Slade, D., Bailone, A., Sommer, S., Averbeck, D., Petranovic, M., Lindner, A. B., & Radman, M. (2006). Reassembly of shattered chromosomes in *Deinococcus radiodurans*. *Nature*, *443*(7111), 569–573. https://doi.org/10.1038/nature05160

Zhang, K., Zhu, J., Li, S., Rao, M. P. N., Li, N.-M., Guo, A.-Y., Jiang, Z., Tang, Q.-Y., Wan, Y., Zhang, Z.-D., & Li, W.-J. (2020). *Deinococcus* detaillensis sp. Nov., isolated from humus soil in Antarctica. *Archives of Microbiology*, *202*(9), 2493–2498. https://doi.org/10.1007/s00203-020-01920-0

Zhu, J., Li, S.-H., Tang, Q.-Y., Chu, M., Wang, W., Salam, N., Li, L., Hozzein, W. N., Zhang, Z.-D., & Li, W.-J. (2017). *Deinococcus* malanensis sp. Nov., isolated from radiation-polluted soil. *Archives of Microbiology*, *199*(4), 621–626. https://doi.org/10.1007/s00203-016-1335-0

4 Chapter IV: Evolutionary History of the *Deinococcota* phylum and the Emergence of the *Deinococcaceae* family

Question:

What are the key evolutionary events that gave rise to the common ancestor of *Deinococcaceae*, and how did these factors shape the distinctive characteristics observed in this family?

Hypothesis:

The emergence of the *Deinococcaceae* family is associated with a notable genome expansion facilitated by genetic flexibility. This genomic plasticity enabled adaptation to diverse habitats and played a pivotal role in the family's evolutionary success and the development of distinct traits.

4.1 Abstract

Deinococcota is an ancient bacterial phylum comprising members with strikingly divergent evolution. *Deinococcaceae* and *Thermaceae* are the two major families in this phylum. The evolution of *Deinococcaceae* members was towards global dispersal and adaptation to diverse habitats, while the evolution of *Thermaceae* made it a generalist that only occupies thermophilic environments. The first comparative genomic study of the *Deinococcota* phylum compared the genome of *Deinococcus radiodurans* and *Thermus thermophilus* and provided insight into their evolution but did not achieve its goals in addressing the radiation resistance phenotype. Later, *D. geothermalis* was added to comparative genomic analysis and addressed more detailed evolutionary events among these two species and *T. thermophilus* but again failed to address the radiation resistance trait. Therefore, it was concluded that comparative genomics could not explain radiation resistance, and afterwards, the few comparative genomic studies focused on genes related to radiation resistance. Consequently, no comprehensive comparative genomics study has been conducted since 2007 that exclusively studied the evolution of *Deinococcaceae* as a family. In this chapter, we aimed to use genomes of the *Deinococcota* phylum to infer evolutionary events that gave rise to the common ancestor of the *Deinococcaceae* family and gain insight into the potential properties of those evolutionary events. We constructed a robust phylogenomic tree to establish the evolutionary relationships between *Deinococcota* members and compared the genomic features of the *Deinococcaceae* and *Thermaceae* families. Then, we used the gene tree-species tree reconciliation method on a subset of genomes to infer evolutionary events that resulted in the emergence of the common ancestor of *Deinococcaceae*. Our comparative analyses demonstrated a significant difference in the diversity of genome size and protein count between the two major families. *Deinococcaceae* members showed a more diverse and larger genome size, while *Thermaceae* members had a homogeneous and smaller genome size. The ancestral reconstruction results demonstrated that the node representing the common ancestor of *Deinococcaceae* had the most notable gene gain event with 1096 genes. The data suggests that this gene gain event enabled genomic diversity of the *Deinococcaceae* crown group and its generalist

lifestyle. On the other hand, *Thermaceae* had few gene gains and remained a specialist taxon. From a functional perspective, transcriptional regulators with roles in oxidative stress resistance response constituted the most gained genes. Aminoacid and carbohydrate transporters followed by replication, recombination, and repair categories also showed high diversity among gained genes. Overall, we conclude that the common ancestor of *Deinococcaceae* went through a significant gene gain event, potentially enabling it to survive under oxidative stressors using transcriptional regulators and amino acid transporters, which can provide proteome protection. Survival of this proto-*Deinococcus*, under oxidative stress, provided an opportunity for it to expand its genome and adapt to new habitats.

4.2 Introduction

Deinococcota, previously known as *Deinococcus-Thermus,* is a deep-branching phylum with some of the best-studied species and is highly significant among extremophilic bacteria. *Deinococcus radiodurans* is the model organism for radiation-resistant phenotype, and its relative *Thermus thermophilus* has been intensively studied for its thermophilic properties. Based on the Genome Taxonomy Database (GTDB) (Parks et al., 2018), this phylum includes one order, *Deinococcales*, classified into five families, four of which have cultivated members: *Deinococcaceae, Trueperaceae, Thermaceae*, and *Marinithermaceae*. One family named JAFFIG01 is only known from environmental sequences and is closely related to the family *Deinococcaceae*. Two of these families, *Trueperaceae* and *Marinithermaceae,* comprise only one cultivated and chemotaxonomically classified member, and our understanding of their members is currently limited (Albuquerque et al., 2005; Battista, 2016; Sako et al., 2003). However, *Deinococcaceae* and *Thermaceae* have been copiously isolated from environmental samples (Battista, 2016; Rosenberg, 2014), with an important role in understanding different extremophilic phenotypes. Some of its members have been the subject of extensive research, and many mechanisms have been discovered that are involved in diverse environmental stressors.

Despite the phylogenetic relationship between *Deinococcaceae* and *Thermaceae*, their evolution was highly divergent, with a thermophilic lifestyle for *Thermaceae* and resistance to ionising radiation in *Deinococcaceae* (Omelchenko et al., 2005). On the one hand, the *Thermaceae* family includes thermophilic genera *Thermus*, *Meiothermus,* and *Calidithermus. Thermus chliarophilus* was reclassified as a type species of a new genus, *Calidithermus.* (Raposo et al., 2019). On the other hand, the *Deinococcaceae* family includes seven genera: two genera have validly published names, three genera delineated based on former members of *Deinococcus* and two represented by environmental sequences with placeholder names. *Deinococcus* and *Deinobacterium* are the two genera with validly published names and include species with radiation

resistance levels (Rosenberg, 2014). Unlike *Thermaceae* members, which only live in high-temperature habitats and are sensitive to ionising radiation (Omelchenko et al., 2005), *Deinococcaceae* members are ubiquitous and found in diverse habitats, including geothermal pools. For instance, *D. geothermalis*, *D. murrayi* (Ferreira et al., 1997), and *Deinobacterium chartae* (Ekman et al., 2011) demonstrate thermophilic and radiation resistance phenotypes. Family *Trueperaceae* has one genus and species with a validly published name, but 17 other genera are known only from the environmental sequences. This is a great potential source for evolutionary studies. However, it was excluded from our dataset due to low completeness. *Truepera radiovictrix*, the only cultivated and named member of *Trueperaceae* (Albuquerque et al., 2005), has the same phenotype as those *Deinococcus* species. This ecological adaptability underscores the fascinating diversity within *Deinococcaceae* and highlights the divergent evolution of this family with *Thermaceae.*

D. radiodurans was among the first sequenced organisms because of its extraordinary phenotype in radiation resistance and potential applications in biotechnology and bioremediation (White et al., 1999). The availability of the whole genome sequence of *D. radiodurans* initiated comparative genomics studies by comparing the *D. radiodurans* genome with other microorganisms to understand the genomic basis of radiation resistance (Makarova et al., 1999). Later, *D. radiodurans* and *Thermus thermophilus* were compared with a focus on their divergent adaptation. It was found that the mega-plasmids of *T. thermophilus* and *D. radiodurans* are homologous and likely inherited from a common ancestor (Omelchenko et al., 2005). When the thermophilic *D. geothermalis* was sequenced, Makarova and colleagues compared its genome to *D. radiodurans* and *T. thermophilus* but failed to identify unique DNA repair mechanisms (Makarova et al., 2007). The result of these studies weakened the case for the role of specific universal genes in radiation resistance but strengthened the case for the role of manganese and iron homeostasis as a scavenging strategy for reactive oxygen species (ROS) and shifted the field's focus to the role of Mn antioxidants in radiation resistance phenotypes (Daly, 2023; Makarova et al., 2007). After these attempts, a few comparative genomic studies focused on the radiation resistance of a few *Deinococcus*

species based on known genes from *D. radiodurans* and in the form of review articles (Lim et al., 2018; Makarova & Daly, 2010).

Despite the availability of numerous genomes and more advanced tools, the dominance of radiation-resistance research on *D. radiodurans* and a few other resistant species overshadowed a deep understanding of the evolutionary history of *Deinococcaceae*. The only recent comprehensive evolutionary study in the *Deinococcota* phylum was conducted by Jiao and colleagues, who looked at the evolutionary history of 23 *Thermus* species and found an incomplete denitrification pathway that was vertically inherited (Jiao et al., 2022b). They suggested *Thermus* as a significant denitrifier in hydrothermal environments. Meanwhile, the evolutionary history of *Deinococcaceae* has remained vastly underexplored, and there is a gap in our knowledge about the evolutionary events that gave rise to the *Deinococcaceae* family.

With advancements in computational capabilities and novel bioinformatic methods, detecting evolutionary events, including gene duplication, transfer, and loss (DTL), has become more reliable. Over the past decade, species tree-aware phylogeny has become useful for detecting evolutionary events, including DTL, by modelling how gene trees are generated along the species tree by a series of evolutionary events (Williams et al., 2023). David and Alm initially proposed a parsimony framework algorithm to map the evolutionary history of thousands of gene families (David & Alm, 2011). They discovered that many prominent modern gene families evolved in a brief genetic innovation span with a rapid diversification of bacterial lineages during the Archaeozoic era. Recent studies extended the reconciliation method to the probabilistic framework (Scornavacca et al., 2017; Szöllősi et al., 2013) and tried to answer complex questions. Coleman and colleagues reconstructed the last bacterial common ancestor by rooting the tree of life without considering archaea as an outgroup (Coleman et al., 2021). This method was also used to discover that metabolic complexity could increase during the endosymbiotic lifestyle of *Chlamydiae* (Dharamshi et al., 2023). Sheridan and colleagues found novel mechanisms of genome expansion in ammonia-oxidising archaea of the phylum *Thaumarchaeota* (Sheridan et al., 2020). These promising results

have shown the versatility of the gene tree-species tree reconciliation approach and that this method is suitable for discovering complex evolutionary events.

In this study, we used state-of-the-art methods to address the abovementioned knowledge gap by answering the following questions: what are the key evolutionary events that gave rise to the common ancestor of *Deinococcaceae,* and how did these factors shape the distinctive characteristics observed in this family? We used the parsimony-based method AnGST tool (Analyser of Gene and Species Trees) that accounts for HGT by comparing individual gene phylogenies with the phylogeny of species and generates evolutionary histories for all gene families within the genomes under study (David & Alm, 2011). We also examined the functional roles of genes involved in these evolutionary events and identified insights into the evolution of the *Deinococcaceae* family.

4.3 Methods

4.3.1 Selection of representative genomes

A dataset comprising high-quality genomes of the *Deinococcota* phylum with species-level representatives was selected from the GTDB database r214 (Parks et al., 2022). Genomes were quality filtered based on the following criteria: CheckM completeness > 90, CheckM contamination < 3.5, and quality score > 85 (quality is defined as completeness - (5 * contamination) (Parks et al., 2018). The final dataset, after quality filtering, included 107 genomes from both named species and environmental sequences. The dataset includes 66 species-level representatives of *Deinococcaceae*, 37 representatives of *Thermaceae,* and one representative genome for each family of *Trueperaceae* and *Marinithermaceae*, and two members of *Actinomycetota*, *Thermoleophilum album* and *Rubrobacter radiotolerans*, as outgroups for the phylogenomic studies and rooting (Supplementary Data S4-1). Genomic features were obtained from the GTDB bac120_metadata_r214 database and used for statistical analyses (Figure 4-1 and Supplementary Data S4-1). The genomes were downloaded from the NCBI assembly database. Prokka v1.14.5 (Seemann, 2014) was used to

identify open reading frames (ORFs), and the proteome sequences were used for further analysis.

4.3.2 Phylogenomic reconstruction

4.3.2.1 Datasets

In this thesis, I used two datasets: a general dataset for phylum-level phylogenomic reconstruction, including 107 genomes, to demonstrate relationships between members of *Deinococcota*, and a sub-sample of those genomes includes 14 genomes for calculating evolutionary events. We used the sub-sample because of limitations in computational resources, but the genomes were selected to represent general trends of genomes to generate reliable results for the common ancestor of *Deinococcaceae*. Species trees for both datasets were inferred using the same method through phylogenomic data of concatenated single-copy marker genes (Supplementary Data S4-1-S, S4-2).

4.3.2.2 Ortholog inference

Single-copy orthologs were identified using OrthoFinder with the default setting (Emms & Kelly, 2019). Ortholog Group (OG) Sequences were aligned using MAFFT v7.471 (Katoh & Standley, 2013) With the -auto option, and spurious sequences and poorly aligned regions were filtered with trimAl v1.2 with the -automated1 option (Capella-Gutiérrez et al., 2009). The ortholog inference resulted in 175 marker genes in the general data species (including the outgroup species) and 545 single-copy OGs for the phylogenomic reconstruction. This difference in the number of markers is because two genomes used as outgroups in the general dataset belonged to the *Actinomycetota* phylum and reduced the number of genes present in all genomes for the general dataset. However, this difference in marker genes did not impact the tree topology.

4.3.2.3 Species tree construction

Phylogenomic reconstruction was performed on each dataset using a concatenated alignment of core OG alignments. IQ-TREE v2.2.2.6 (Nguyen et al., 2015) was used to

infer the unrooted species tree. The LG+R8+F model was chosen as the best-fitting model by the BIC criterion predicted in ModelFinder (Kalyaanamoorthy et al., 2017)**.** Branch supports were computed using the 1000 ultrafast bootstraps (UFBoot) replicates. The species tree was rooted with the *Actinomycetota* genomes as outgroups. The species tree for the general dataset was visualised and annotated using the genomic data by the Interactive Tree Of Life (iTOL) (Letunic & Bork, 2021).

4.3.3 Genome-wide DTL detection

The sub-sample dataset was used for detecting evolutionary events that gave rise to the common ancestor of *Deinococcaceae*. For this purpose, we used the AnGST tool (David & Alm, 2011) for the gene tree species tree reconciliation implemented by the FastDTLmapper pipeline automates estimating and mapping genome-wide gene gain/loss (Shimoyama, 2021). AnGST requires two inputs: a species tree, which was constructed as described and a gene tree for each OG.

4.3.3.1 Gene tree reconstruction

First, the sub-samples dataset was used for orthology analysis and OG sequences inference by OrthoFinder default format -I = 1.5 (Emms & Kelly, 2019). Then, the OG sequences were aligned using MAFFT (Katoh & Standley, 2013) and trimmed by the trimAl tool (Capella-Gutiérrez et al., 2009). In the next step, IQ-TREE with the option iqtree -s <OGX_aln_trim.fa> --prefix -m TEST -mset JTT,WAG,LG --seed 0 --ufboot 1000 --boot-trees --wbtl (Nguyen et al., 2015) was used to infer gene trees for amalgamation using the ultra-fast bootstrap and generation of *.ufboot files. Then, Treerecs (Comte et al., 2020) was used to correct gene tree multifurcation in each OG sequence.

4.3.3.2 DTL reconciliation

Then, DTL reconciliation of the species tree and each OG gene tree was performed using AnGST (David & Alm, 2011). The AnGST_wrapper.py function in FastDTLmapper was used to generate the evolutionary events using AnGST. Finally,

the reconciled DTL result was aggregated and mapped on the species tree, and the number of DTL events was calculated for each internal node.

4.3.4 Ancestral reconstruction and functional annotation:

FastDTLmapper sub-tool FastDTLgoea was used to extract gene OG id of sequences that were predicted as gain and loss at each node. We used three internal nodes, N002, N006, and N007, to investigate changes in evolutionary events that form the common ancestor of *Deinococcaceae* (N007), which we call proto-*Deinococcus* in this study. Protein family sequences in the three internal nodes were annotated using the EggNOG mapper (Cantalapiedra et al., 2021). The annotation was used for ancestral metabolic reconstruction at each node of the species tree.

4.4 Results:

4.4.1 *Deinococcaceae* have diverse genomic features

To gain a general understanding of the evolution of *Deinococcota* members, we compared the genomic features of all genomes in our dataset, which includes genomes from both cultivated and uncultivated species. *Trueperaceae* and *Marinithermaceae* were excluded from statistical analyses because they have only one available genome and lack sufficient data for meaningful comparative analysis. A comparison of genomic features at the family level showed a significant difference between *Deinococcaceae* and *Thermaceae* regarding genome size and protein counts. However, no significant difference was observed in GC content. *Deinococcaceae* members have a larger genome size and a higher number of proteins than *Thermaceae*. There was no statistically significant difference between the two families in their GC content. However, three genomes in the *Deinococcaceae,* including *D. cellulosilyticus*, *D. roseus*, and *D. misasensis*, showed a notably lower GC percentage and were outliers. Those species formed a distant clade in the *Deinococcaceae* family with a larger genome size (Figure 4-1A***Error! Reference source not found.***). At the species level, members of the *Thermus* genus showed a homogeneous genome size, but *Meiothermus* genomes showed higher diversity with a larger genome size and lower GC content. Interestingly, *Meiothermus* species exhibited significant differences from *Deinococcus* and *Thermus* members, but no significant difference was observed between *Deinococcus* and *Thermus* (Figure 4-1)

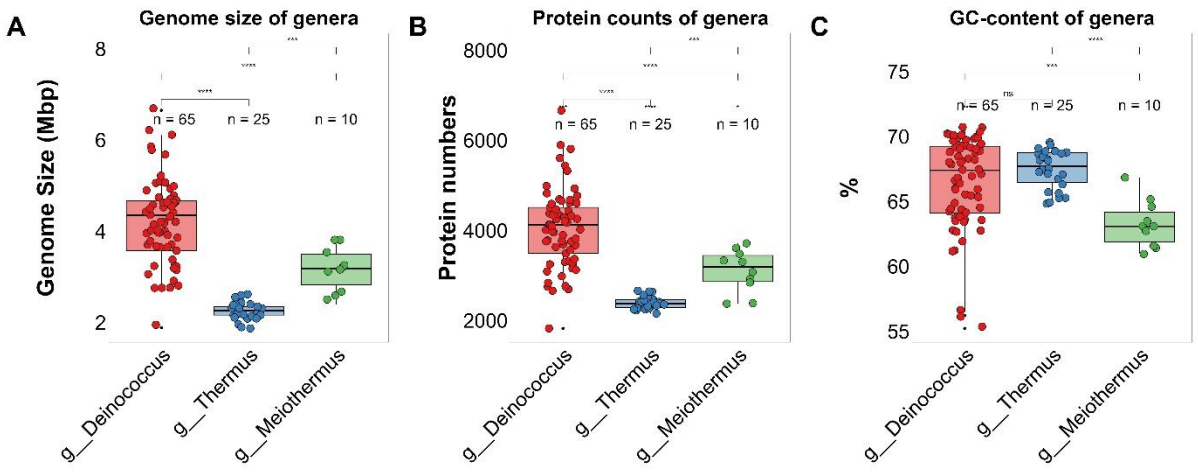

Figure 4-1. Genomic features of *Deinococcota*.

Boxplot demonstrating differences in the genomic features of *Deinococcota* A. *Deinococcus* genus (red) has a highly diverse genome size, but *Thermaceae* shows a different trend for the *Thermus* genus (blue) and *Meiothermus* (green), and there is a significant difference between each group. B. Protein counts show the same trend as genome size. C. GC content in *Deinococcaceae* is highly diverse, with three genomes being outliers

4.4.2 Establishing a robust phylogeny for *Deinococcota*

To gain deeper insights into the evolutionary history of *Deinococcota* and events that gave rise to the *Deinococcaceae* family, we reconstructed the phylogenomic tree of *Deinococcota* phylum using a concatenated multiple sequence alignment (MSA) of single-copy orthogroups shared between all genomes. Our tree comprised 108 high-quality genomes belonging to *Deinococcota* and *Actynomycetota* as outgroups (Supplementary Data S4-1 and Figure 4-2). Our phylogenomic tree provided a well-supported relationship between the four families within the *Deinococcota*, with most nodes with ultrafast bootstrap (UFBoot) values >95% (Figure 4-2).

Genomes used in this study had high-quality criteria values (average of 98.74%), with the lowest value being an uncultured *Deinococcus* sp. 91.9 %). The average CheckM contamination was 0.32%, with the highest value of 1.72% for the *Thermoleophilum album*. Members of *Deinococcaceae* showed a wide range of genome sizes and protein

counts, with an average genome size of 4.17 Mbp and ranging from 1.8-6.6 Mbp, while *Thermaceae*, especially the *Thermus* genus, showed a homogeneous genome size. The average genome size in the *Thermus* genus was 42.7% smaller than *Deinococcaceae*, with 2.39 Mbp and ranging from 2.1 to 2.6 Mbp. However, *Calidithermus* and *Meiothermus* within the *Thermaceae* family showed larger genomes with an average of 3.1 Mbp and higher variability (2.3-3.7 Mbp) (Figure 4-1). In most clades, GC content correlated with phylogeny, with the notable exception in the distant clade of *Deinococcaceae,* where *D. misasensis, D. roseus,* and *D. cellulosilyticus,* showed a very low GC content compared to other members of *Deinococcaceae* and *Thermaceae.* Moreover, these named *Deinococcus* species formed a clade with *Deinobacterium chartae,* which belongs to another genus, indicating the *Deinococcaceae* family is paraphyletic, meaning even though the common ancestor is present, all the descendants are not present. Members of the *Thermus* genus showed a homogeneous GC content, but *Deinococcus* and *Meiothermus* showed higher levels of heterogeneity, the same trend as their genome size.

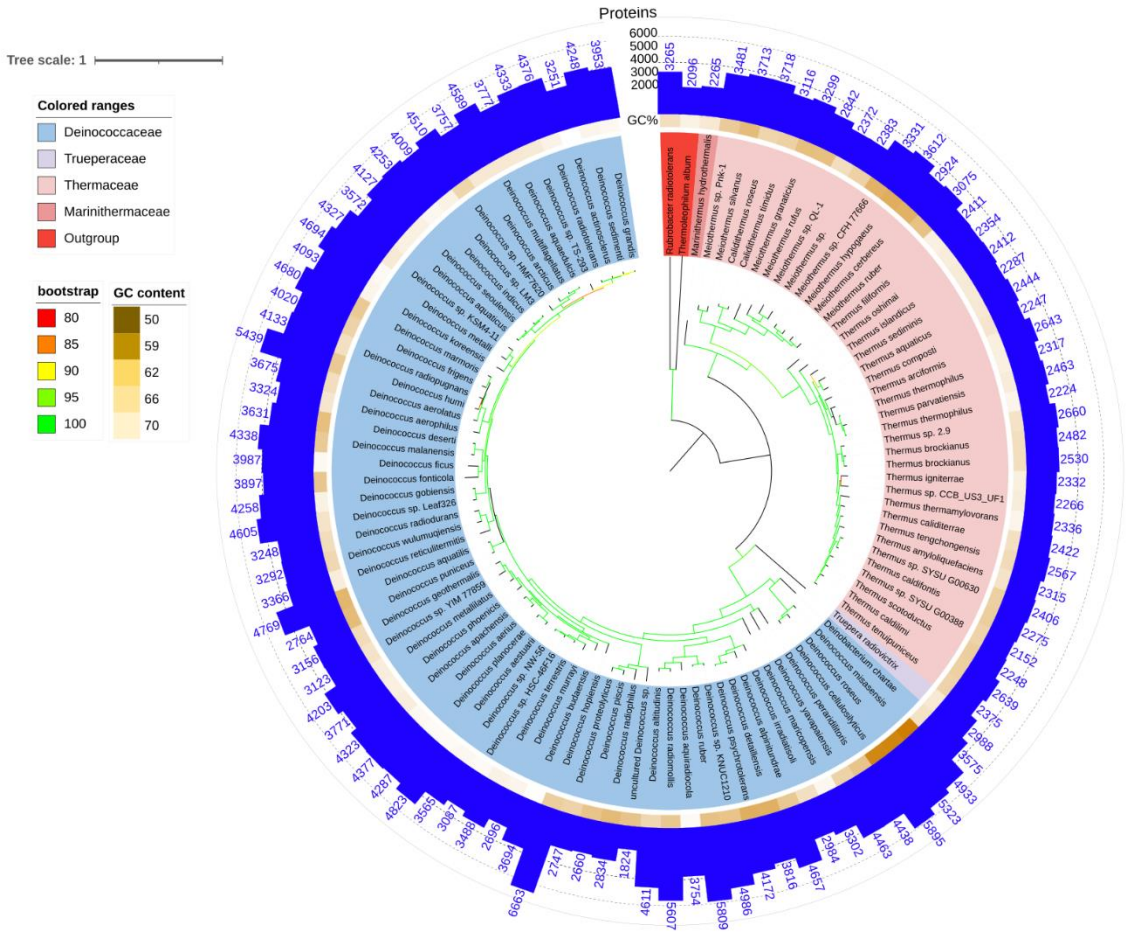

Figure 4-2. Phylogenomic tree of *Deinococcota* phylum.

This tree includes 108 genomes (66 *Deinococcaceae* genomes, 37 *Thermaceae*, including 25 *Thermus*,10 *Meiothermus*, and 2 *Calidithermus* genomes, and one genome for each *Truepera* and *Marinithermus*). Two genomes from Actinomycetota were used as outgroups. The tree was inferred by maximum likelihood from 175 concatenated phylogenetic markers, which were aligned separately and analysed using the best-fitting model for each alignment. Light blue clade colours indicate the *Deinococcaceae* family. *Thermaceae* are highlighted with pink. Dots indicate branches with >95% UFBoot. GC content values are demonstrated with a brown to yellow gradient strip with lighter yellow colours indicating higher GC content. Protein counts are demonstrated with bar plots and numbers in the outer layer. Detailed genome information is provided in Supplementary Data S4-1.

4.4.3 Evolutionary history of *Deinococcota*

To gain a deeper understanding of the evolution of *Deinococcaceae,* we explored the evolutionary events leading to the emergence of the common ancestor of *Deinococcus,* which we denote as proto-*Deinococcus*. We used a macroevolutionary model based on the parsimony framework gene tree-species tree reconciliation method previously described by David and Alm (David & Alm, 2011) to estimate gene birth, transfer, duplication, and loss events and map the evolutionary history of 4295 gene families across 14 representative genomes in our selection based on tree topology. Comparing the evolutionary events in each internal node, which represents an ancestor, can help us understand processes that shape the genomic features of current genomes at the tip of the tree.

Our results demonstrate that the common ancestor of *Deinococcaceae* and *Thermaceae* (node N002) gained 662 genes, with no gene loss and had a total of 2338 genes. However, in node N003, the divergence of *Thermus* from *Calidithermus* and *Meiothermus* caused a significant gene loss in *Thermus* but a moderate gene gain for the other branch. It seems that the N002 was more similar to the current-time *Thermus* in terms of genomic content and features and was probably a thermophilic organism. The next evolutionary event in the direction of the emergence of *Deinococcaceae* was the divergence of *Truepera* from *Deinococcaceae* at the N006 node. This node showed a relatively consistent gene gain trend with the N002 node, with 438 gene gains, 222 gene losses, and 2554 genes. *T. radiovictrix,* the only cultivated member of *Trueperaceae,* is both radiation-resistant and slightly thermophilic. We could not conclude whether the common ancestor of *Deinococcus* and *Truepera* was resistant to radiation. These observations show that *Thermus* members gained a few genes but lost more genes, and its proteome size remained consistent within a limited range.

4.4.4 The common ancestor of *Deinococcaceae*.

The most significant evolutionary event in the *Deinococcota* family was the emergence of *Deinococcaceae* at the N007 node. This proto-*Deinococcus* genome gained 1096 new genes and lost 199 genes, expanded its genome to 3450 genes and established a more extensive genomic backbone for its descendent taxa (Figure 4-3). The majority of these gene gains (79%, 865 genes) resulted from gene birth, 13.5% (149 genes) from duplication, and 7.5% (82 genes) from gene transfer. The proto-*Deinococcus* also lost 199 genes after diverging from its ancestor N006. Descendants of proto-*Deinococcus* were divided into three major deep-branching groups consisting of three *Deinococcus* species, *D. misasensis*, *D. roseus*, and *D. cellulosilyticus*, which are categorised as *Deinococcus_C* in GTDB and clustered with *Deinobacterium chartae*. *D. maricopensis* is the other deep-branching species, and it is designated *Deinococcus_B* in GTDB. And finally, *D. peraridilitoris* and *D. pimensis* are members of the other deep branching group, which are designated as *Deinococcus_A* in the GTDB.

The genome size of *D. chartae* remained unchanged, but the three *Deinococcus* species had significant gene gains. For instance, *D. cellulosilyticus* underwent a large gene gain event and gained 2809 genes (1413 born, 949 duplications, 447 transfer), lost 551 genes, and expanded its proteome size to 5857 genes (Figure 4-3**Error! Reference source not found.**). The data shows that genome diversification began after proto-*Deinococcus* diverged into two branches: a smaller branch with three genomes and a more extensive and diverse branch with 62 genomes (Figure 4-2).

The other branch of *Deinococcus* was highly diversified and formed multiple clades with various genome sizes. In those clades, we observed dramatic gene gain and loss events. For instance, *D. piscis*, isolated from a marine fish (Shashidhar & Bandekar, 2009), lost 1329 genes versus 668 gene gains and underwent genome reduction. In another clade, *D. ruber*, which is a radiation-sensitive *Deinococcus* species and was isolated from garden soil in South Korea (Kim et al., 2017), gained 2614 genes and lost 487 genes, expanding its proteome to 2614 genes.

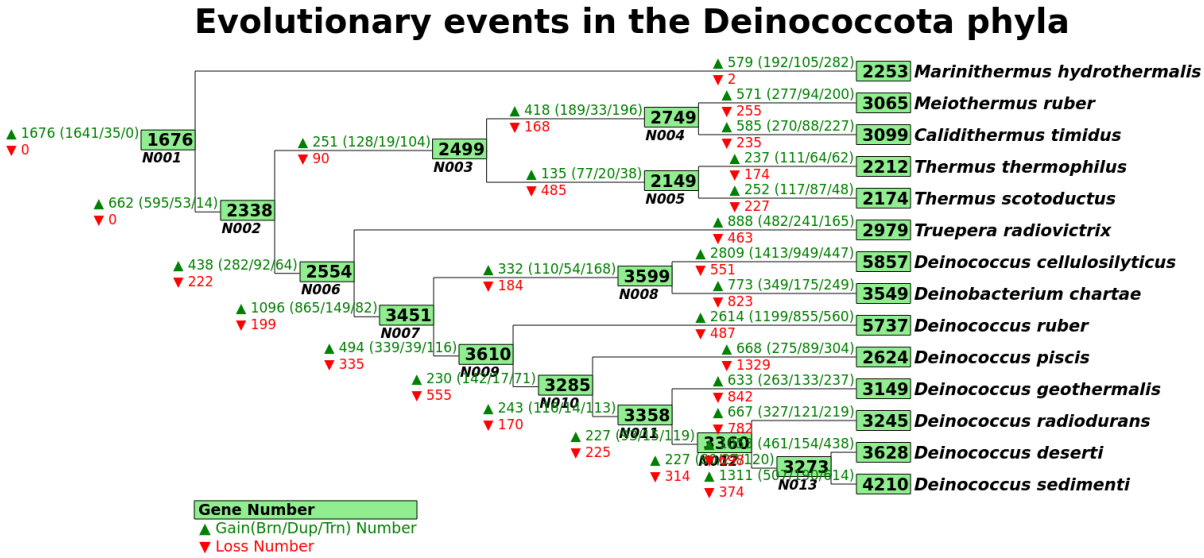

Figure 4-3. Evolutionary events in the *Deinococcota* phylum.

The predictions of the proteome changes at evolutionary events were estimated across representative genomes of the *Deinococcota* phylum and reported on the cladogram possessing the same topology of the ML tree presented in Figure 4-2. Internal nodes represent the ancestors of each clade, and N007 represents the common ancestor of *Deinococcaceae* with 1096 gene gain events. In addition to the changes in proteome size (green box at the internal nodes), the number of gene gains and detailed numbers of born, duplicated, and transferred genes were shown in parenthesis next to the total number in green colour next to the green arrow (▲) and the number of gene loss was demonstrated with red down arrow (▼). In the subsample tree, *Marinithermus hydrothermal* was used as an outgroup.

4.4.5 Functional annotation of proto-*Deinococcus*

To examine genes that constitute the discussed ancestral nodes N002, N006, and the proto-*Deinococcus* (N007), we annotated gene families comprising each node against COG databases. The comparison of evolutionary events in the three internal nodes leading to the emergence of the proto-*Deinococcus* showed significant gene gain events in specific functional categories. As demonstrated in Figure 4-4, the most noteworthy gene-gain event in the proto-*Deinococcus* was observed in the transcription category (K) with 106 gene families, constituting about 10% of all gained genes in this node and about 3.32 % of the proto-*Deinococcus* genome. N002 and N006 nodes, with 34 and 33 acquired genes, respectively, showed a lower number and abundance of transcription-related gene gains.

Among the 106 gene families gained in the proto-*Deinococcus* in the transcription category, 92 gene families were unique and only observed in proto-*Deinococcus* (Supplementary Data S4-4). The highest diversity of acquired genes in the transcription category was related to DNA-binding proteins such as PhnO (N-acetyltransferase, GNAT superfamily), AcrR (nucleoid occlusion protein SlmA), LacI/PurR, and MarR (environmental surveillance of aromatic compounds). Other genes with implications in environmental adaptation are PadR (regulating amino acid catabolism and cellular response to chemical stress agents and drugs) and ArsR families (transcriptional repressor of an arsenic resistance operon). These transcription factors are known for regulating broad mechanisms, including post-translational modifications, cell division, and oxidative stress response (Makarova et al., 2007).

The DdrO transcriptional regulator is one of the most essential regulatory genes in oxidative stress response (de Groot et al., 2019). This regulatory gene was only observed in proto-*Deinococcus* and not in the N002 and N006 nodes. However, the IrrE, a Zn metalloprotease that activates the radiation desiccation response (RDR) regulon by cleavage of DdrO (Ludanyi et al., 2014), was gained earlier in N002. Some other acquired genes in the transcription category are the Gntr family transcriptional regulator, winged helix DNA-binding domain, LysR family (involved in virulence, metabolism, quorum sensing and motility) regulation of single-species biofilm formation, and 2Fe-2S cluster binding regulators. *Proto-Deinococcus* also lost nine gene families in this category. Among them are the TetR family (widely associated with antibiotic resistance and the regulation of genes encoding small-molecule exporters) and the negative regulator of class I heat shock genes (grpE- dnaK-dnaJ and groELS operons).

The next category with a high gene gain ratio was amino acid transport and metabolism (E), with 82 gene families for proto-*Deinococcus*, constituting 2.57 % of this ancestral genome Figure 4-4. Among those gained genes, 75 gene families were unique for proto-*Deinococcus*. The most diverse genes acquired in this category are ProP (Proline/betaine transporter), GloA (lactoylglutathione lyase activity), RhaT (Permease

of the drug/metabolite transporter (DMT) superfamily), and DdpA (dipeptide ABC-transporters). In the amino acid transport category, the N002 node also had many gene gains with 73 genes, but the number of acquired genes in N006 was limited to 24 gene families.

The third category with the highest abundance of gene gain for proto-*Deinococcus* was the carbohydrate transport and metabolism category (G). In this category, genes like ProP and RhaT (with multiple functionalities and present in different categories), NagC (Transcriptional regulator sugar kinase), Glycosyl hydrolases family, and UgpB (ABC-type sugar transport system periplasmic component) showed higher abundance. Other genes, like different families of epimerase, YliI pyrroloquinoline quinone binding protein, and alpha-galactosidase, were also gained in proto-*Deinococcus*.

The proto-*Deinococcus* also had a high ratio of gene gain in the replication, recombination, and repair category (L), comprising 2.06 % of its genome. This category's highest diversity of gained genes was observed in transposase families and CRISPR-Cas endonuclease/helicase Cas3. Furthermore, various genes with roles in different oxidative repair mechanisms were acquired in proto-*Deinococcus* . Genes like RecD (DNA-dependent ATPase and ATP-dependent 5'-3' DNA helicase), the UvrABC ATPase subunit, DdrB protein (which is an alternative ssDNA binding protein and essential for protecting fragmented DNA strands), AlkA (endonuclease III), SbcDC DNA repair exonuclease SbcCD ATPase subunit and many other enzymes involved in oxidative stress repair. The common ancestor of *Deinococcus* and *Thermus* (N002) showed a different pattern in this category, and most of the gained genes were related to transposases and CRISPER-Cas enzymes (Supplementary Data S4-3 and Figure 4-4).

In the signal transduction category (T), the proto-*Deinococcus* demonstrated the highest percentage of gene gain compared to other nodes, with 52 gene families, which includes about 5% of all gained genes and 2.06% of its genome. The following functions showed the highest abundance in this category. OmpR (Response regulators consisting of a CheY-like receiver domain), BaeS and ComP (two types of Histidine kinase-, DNA gyrase B-, and HSP90-like ATPase), diguanylate cyclase (GGDEF) domain, SpoIIAA

and ChrR (antisigma factor related genes), TerZ (cAMP binding), and the universal stress protein *(*UspA).

The top five abundant functional categories discussed above, transcription, amino acid transport, carbohydrate transport, replication, recombination, and signal transduction, constituted 36 % of all gained genes in the proto-*Deinococcus*. Except for the inorganic ion transport and metabolism (P) category, where N002 showed the highest gene gain rate, the three ancestors showed similar trends in gene gain for most of the remaining categories. N002 did not show any gene loss in any of the categories, and that could be because it was near the root and gene flow could not be detected from the outgroup. In two COG categories, including cytoskeleton and extracellular structure, proto-*Deinococcus* gained only one gene, and other nodes showed no gene gain events. Immunoglobulin-like and fibronectin type III were acquired in the cytoskeleton category, and FecR was in the extracellular structure category.

The common ancestor of *Deinococcaceae* and *Trueperaceae*, which is depicted as the N006 node, showed a gene loss of 222 genes (Figure 3-3). The proto-*Deinococcus* also lost 199 genes after divergence from N006. The majority of those lost genes were related to (E) Amino acid transport and metabolism, (M) Cell wall/membrane/envelope biogenesis, (C) Energy production and conversion, and (P) Inorganic ion transport and metabolism. Coenzyme, transport and mechanism (H) was a category in which both N007 and the proto-*Deinococcus* lost several genes, but N002 did not lose any genes.

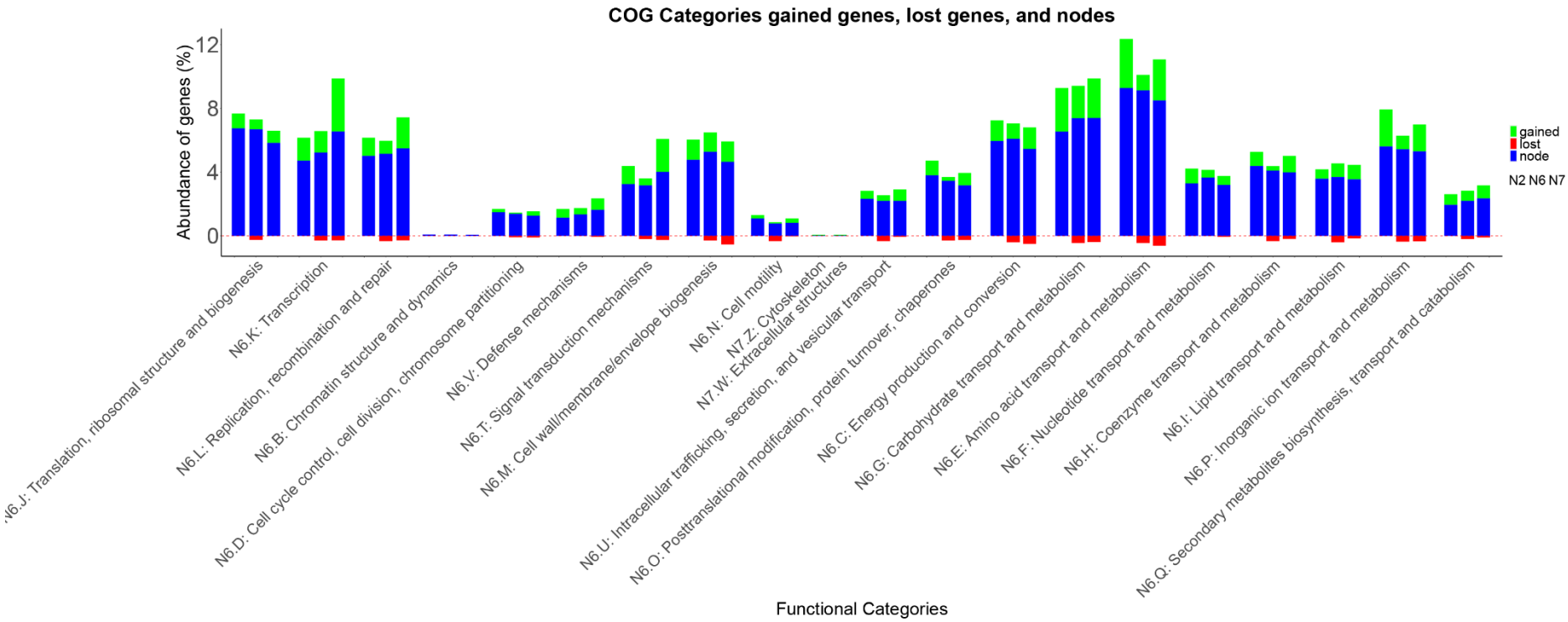

Figure 4-4. Functional annotation of three ancestors.

Three bars represent the three evolutionary events and internal nodes. Respective, the first bar denotes the N002 node, the middle bar represents the N006 node, and the third bar denotes N007 or the proto-*Deinococcaceae*. Internal colours show gene gain (green), loss(red), and an abundance of annotated gene numbers at each node.

4.5 Discussion

This study aims to provide a general view of the evolutionary events that gave rise to the *Deinococcaceae* family. In doing so, we conducted a comparative genomics analysis to provide a high-resolution evolutionary relationship between all members of the *Deinococcota* phylum and demonstrate similarities and differences between families in this phylum. Then, we used the gene tree-species tree reconciliation method to determine evolutionary events, including horizontal gene transfer, duplication, and loss, to explain topological incongruities between the two trees and infer evolutionary events. Our results indicate a significant gene gain event at the internal node belonging to the common ancestor of *Deinococcaceae* (which we denote as 'proto-*Deinococcus'* in this study). We argue that this gene-gain event enabled higher genomic flexibility and diversity among the *Deinococcaceae* members, especially the *Deinococcus* genus.

4.5.1 Divergent evolution of *Deinococcaceae*

Our data demonstrated a striking difference between the genomic features of *Deinococcus* species and other members of the *Deinococcota* phylum. *Deinococcus* species have diverse genome sizes and protein numbers ranging from 1824 to 6663 and an average of 4051 proteins**.** However, *Thermus* members have homogeneous genome sizes and protein counts of 2125 to 2660 genes and an average of 2391 genes and seem to be more similar to the common ancestor of *Deinococcus* and *Thermus*, node N002 (Figure 4-3). Omelchenko and colleagues conducted a comparative genomic using *D. radiodurans* and *T. thermophilus*. They concluded that *Thermus* underwent extensive gene loss but acquired numerous genes from thermophiles, contributing to its thermophilic adaptation. Our data adds to this study and reveals that members of *Thermaceae* maintained the same genomic features, but evolutionary events that enabled genome diversification in *Deinococcaceae* members led to this evident divergence. The ecological implications of this divergent evolution are reflected in the cosmopolitan lifestyle of *Deinococcaceae*. As a result, this group can survive in various habitats and adapt to diverse environments by acquiring essential functions (Krisko &

Radman, 2013). In contrast, *Thermaceae* retain its specialist lifestyle by thriving exclusively in thermophilic environments (Jiao et al., 2022b).

Our results on this divergent evolution align with the findings of von Meijenfeldt and colleagues, which demonstrated that generalist species are genomically more heterogeneous than species of specialist genera, which have more similar genome sizes and less variation in functions (von Meijenfeldt et al., 2023). They also hypothesised that genomic flexibility allows members of generalist genera to rapidly acquire essential genes needed to thrive in a new environment, and higher growth rate potential allows them to outgrow specialists (von Meijenfeldt et al., 2023). Our study suggests that it is true that *Deinococcus* members can gain new genes and survive in new environments, but contradicts the hypothesis that generalists can always outgrow specialists in a given environment. *Deinococcus* members do not meet the description of generalist organisms because they are often among the rare populations and can not outgrow specialists in new environments and dominate their environment. Even though they are ubiquitous and can survive in a broad range of habitats, *Deinococcus* cannot thrive and grow under oxidative stress but can proliferate when optimum growth conditions are provided (Krisko & Radman, 2013). The body of literature illuminates the ubiquity of *Deinococcaceae* in light of its heterogeneous genomic features and generalist lifestyle. At the same time, *Thermaceae* members appear to experience minimal diversification after divergence from their common ancestor with *Deinococcaceae*.

Several studies have shown that *Deinococcus* species can survive prolonged cytotoxic stressors such as desiccation and extraterrestrial environments in an effort to test the panspermia hypothesis, which proposes interplanetary transfer of life (Kawaguchi et al., 2013, 2020; Panitz et al., 2019; Yamagishi et al., 2018). Extreme resistance of *Deinococcus* to these stressors raised whimsical theories that radiation resistance in terrestrial bacteria has a Martian origin and that life had transferred to Earth through Martian meteorites (Pavlov et al., 2006), which is far from scientific evidence on the emergence of life on earth (Krisko & Radman, 2013). Nevertheless, there are legitimate concerns regarding planetary protection and contamination of Mars through terrestrial

microorganisms by programs sending payloads to Mars's surface. Horne and colleagues tried to demonstrate the impact of the Martian surface, which is frozen, dry, and bombarded by solar radiation and galactic cosmic radiation. They showed that *Deinococcus* cells can survive a staggering dose of 140 kGy when desiccated and frozen and indicated that any contamination of the Martian subsurface with these microorganisms would essentially be permanent for 1000s of years, which could complicate scientific efforts to look for martian life (Horne et al., 2022).

Our phylogenomic analysis demonstrated a robust evolutionary relationship among members of the *Deinococcota* phylum, with most bootstrap values being >95. Previous classifications of *Deinococcota* have mainly been based on the single marker 16S rRNA gene (Akita et al., 2020; J. H. Lee et al., 2022; Yin et al., 2022; Zhang et al., 2020b). Since 2018, GTDB has provided phylogenomic classification based on the concatenated phylogeny of 120 protein markers for this phylum, along with all other bacterial and archeal phyla (Parks et al., 2018). Except for one comparative genomic study, which used 120 bacterial marker genes from GTDB and focused on the *Thermus* genus (Jiao et al., 2022b). Our study provides a robust phylogenetic relationship between members of the *Deinococcota* phylum and provides insights into the evolution of *Deinococcaceae*.

The GTDB algorithm categorised the *Deinococcaceae* family into seven genera, including *Deinococcus*, *Deinococcus_A, Deinococcus_B, Deinococcus_C, Deinobacterium*, *JACMOA01,* and *JAJZIR01* (Parks et al., 2018). The last two members only include environmental sequences, and no cultivated species have been reported in these genera. Our findings align with GTDB classification and show that three deep-branching *Deinococcus* species, *D. misasensis*, *D. roseus*, and *D. cellulosilyticus*, are clustered with *Deinobacterium chartae*, which belong to a different genus. This observation indicates that the *Deinococcus* genus is a paraphyletic group. A paraphyletic taxon is a form of a monophyletic taxon that includes the common ancestor of the group of organisms but doesn’t capture all members of the group. Unlike the polyphyletic group, the paraphyletic can still be informative about evolutionary events.

The absence of all descendants of the *Deinococcaceae* family could be due to the disproportionate representation of *Deinococcus* species regarding radiation resistance levels or the extinction of some species. In any case, this observation indicates that incomplete genome sampling can be a limiting factor in studying the *Deinococcota* phylum. While the commonly used nomenclature in *Deinococcota* is congruent with the taxonomic categorisation, the distant taxon, which is categorised with Deinobacterium and is designated as *Deinococcus_C*, may require reconsideration in classification and nomenclature. Therefore, based on our phylogenomic tree and the GTDB classification, we propose a new name for this genus, *Alleodeinobacterium* (meaning the other *Deinobacterium*).

Genome expansion increased genomic diversity in *Deinococcaceae.*

Our data showed that genes involved in replication, recombination, and repair category (L) constituted a high proportion of gained genes in the proto-*Deinococcus,* with 5 % of total acquired genes. Most diverse and abundant genes in this category belonged to different families of transposase enzymes and CRISPR-Cas endonuclease/helicase Cas3. However, more genes related to various DNA repair mechanisms were found in the proto-*Deinococcus* compared to the common ancestor of *Deinococcus* and *Thermus* (Figure 4-4 and Supplementary Data S4, S6). DNA repair enzymes such as RecD (ATP-dependent 5'-3' DNA helicase), RtcB (tRNA-splicing ligase RtcB), SpoIVCA (site-specific recombinase), and YjhB NUDIX domain were exclusive to the proto-*Deinococcus*. However, it seems that main DNA repair mechanisms like the RecFOR complex, which have an essential role in the main homologous recombination repair mechanism in the *Deinococcus,* extended synthesis-dependent strand annealing (ESDSA mechanisms) (Bentchikou et al., 2010), were present in the common ancestor of *Deinococcus* and *Thermus* N002 (data not shown). This observation indicates that homologous recombination and efficient protein repair in the form of ESDSA had evolved before the common ancestor of *Deinococcaceae* and possibly in response to DNA damage from high temperatures. We conclude that the proto-*Deinococcus* amassed a combination of genes related to transposition and transformation competence

genes such as Rec2 (ComEC) and DNA repair enzymes. Moreover, the presence of genes related to double-strand break repairs like RecFOR provided the underlying cause of large-scale genome expansions observed in the *Deinococcaceae* members.

We note that ancestral reconstruction analyses are not precise and are subject to certain limitations. Ancestral reconstruction analyses must always be approached with caution, and all data presented on the common ancestors must be treated as hypothetical and speculative. Moreover, the parsimony framework in the context of gene tree-species tree reconciliation is to infer the most likely history of gene evolution by minimizing the number of evolutionary events. However, this method has its drawbacks and can be a source of certain biases. We are also aware that using a sub-sample of our data could impact the detection of some evolutionary events. However, we aim to address this problem in future publications.

4.5.2 Gene gain in proto-*Deinococcus* enhanced ubiquity

Ancestral reconstruction analysis revealed a significant gene gain event in the common ancestor of *Deinococcaceae* with 1096 new genes. This remarkable gene gain event in proto-*Deinococcus* is arguably the first evolutionary event that accelerated genome diversification and enabled the generalist lifestyle of its descendent taxa, mainly the *Deinococcus* genus. Our functional annotation analysis revealed that many genes gained in proto-*Deinococcus* were related to transcription, replication, recombination and repair, and signal transduction, which could facilitate adaptation to diverse environments through two main steps: (i) surviving under environmental stressors and (ii) further expanding its genome to gain new functions and survive in a broader range of habitats.

We found a high proportion and diversity of acquired genes related to transcription, such as the Yfit/DinB family of proteins, acetyltransferases of the GNAT family, Nudix hydrolases, α/β superfamily hydrolases This finding is in line with previous ancestral reconstruction studies, which showed a plethora of gained genes in the common ancestor of *Deinococcus* were involved in cell-cleaning functions, which facilitated chemical stress-resistance early in the evolution of proto-*Deinococcus*. The independent genes acquired by the crown group enabled secondary adaptations to diverse stress environments through HGT from various sources (Makarova et al., 2007).

Our analysis shows that the proto-*Deinococcus* gained many amino acid transporter genes like DdpA (dipeptide transporters) and a variety of peptidase enzymes like PepE, PepF, and PepP (belonging to metalloendopeptidase families) (Supplementary Data S4-5). This gene-gain event could potentially enable one of the most efficient and well-studied mechanisms in radiation resistance of all organisms: non-enzymatic protection of proteome through Mn^{2+}- oligopeptide complexes (Daly, 2023). We note that these gene gain events are hypothetical, and ancestral reconstruction analyses can not be treated with certainty. However, this finding aligns with the previous studies that showed *D. radiodurans* accumulates peptides that form small molecules of Mn^{2+} complexes and have efficient antioxidant properties that protect enzymes from extreme

oxidative stress (Daly et al., 2004, 2010). Daly and colleagues used electron paramagnetic resonance spectroscopy to show that the *D. radiodurans* ultrafiltrated cell extract is enriched with Mn^{2+}, phosphate, nucleosides and bases, and peptides. They synthesised a decapeptide solution (H-Asp-Glu-His-GlyThr-Ala-Val-Met-Leu-Lys-OH) based on the composition of the most abundant amino acids in the ultrafiltrate of *D. radiodurans*. This synthetic complex supplied with 1mM Mn^{2+} buffer mixture protected the BamHI enzyme against a high dose of 22.5 kGy. They also showed that the cell extract of *D. radiodurans* alone could significantly increase radiation resistance in *E. coli*, Human Jurak T cell ex vivo. (Daly et al., 2010). The antioxidant properties of *Deinococcus* cell extract was so significant that it was used in biotechnology and applied microbiology.

This finding was the basis for developing new generations of vaccines that are resistant to gamma radiation and easily sterilised (Gayen et al., 2017; Tobin et al., 2020). The complex of small antioxidant molecules composed of Mn^{2+} and oligo peptides has been extensively studied in different organisms. However, the metabolic route that causes this metal homeostasis in *D. radiodurans* has remained largely unknown. (Bruch et al., 2015; Daly et al., 2010; Dewar et al., 2021; Gaidamakova et al., 2022; Gupta et al., 2016). More research on these peptide-regulating genes, peptidases, and peptide transporter genes combined with metal binding molecules and metal transporters can be the key to understanding these complex metabolic routes.

4.6 Conclusion

The *Deinococcota* phylum includes ancient groups of bacteria that have undergone divergent evolutionary paths. In this chapter, I conducted a phylogenomic analysis and demonstrated that the two major families within this phylum, *Deinococcaceae* and *Thermaceae*, followed distinct evolutionary trajectories. *Thermaceae* retained a relatively homogeneous genome size, while *Deinococcaceae* exhibited significant diversification in their genomic content.

By analysing 14 representative genomes, we estimated the evolutionary events and found that a substantial genome expansion occurred in the common ancestor of the *Deinococcaceae*, which we refer to as proto-*Deinococcus*. This genome expansion was marked by a significant increase in genes related to replication, recombination, and repair. Among these new genes were a diverse group of transposases and CRISPR-Cas families, suggesting that genomic flexibility and variability in genome size are characteristic of the *Deinococcaceae* descendants.

Additionally, we observed gene loss events, particularly in the categories of cofactor and coenzyme transport and metabolism, in the common ancestors of *Deinococcaceae* and Trueperaceae (node N006) as well as in proto-*Deinococcus*. This loss of genes involved in cofactor transport and metabolism may have contributed to the development of radiation resistance in the common ancestor of *Deinococcaceae* and *Trueperaceae*.

4.7 References

Akita, H., Itoiri, Y., Ihara, S., Takeda, N., Matsushika, A., & Kimura, Z. (2020). *Deinococcus* kurensis sp. Nov., isolated from pond water collected in Japan. *Archives of Microbiology*, *202*(7), 1757–1762. https://doi.org/10.1007/s00203-020-01845-8

Albuquerque, L., Simões, C., Nobre, M. F., Pino, N. M., Battista, J. R., Silva, M. T., Rainey, F. A., & de Costa, M. S. (2005). Truepera radiovictrix gen. Nov., sp. Nov., a new radiation resistant species and the proposal of Trueperaceae fam. Nov. *FEMS Microbiology Letters*, *247*(2), 161–169. https://doi.org/10.1016/j.femsle.2005.05.002

Asgarani, E., Soudi, M. R., Borzooee, F., & Dabbagh, R. (2012). Radio-resistance in psychrotrophic *Kocuria* sp. ASB 107 isolated from Ab-e-Siah radioactive spring. *Journal of Environmental Radioactivity*, *113*, 171–176. https://doi.org/10.1016/j.jenvrad.2012.04.009

Battista, J. R. (2016). *Deinococcus* – Thermus Group. In *Encyclopedia of Life Sciences* (pp. 1–12). https://doi.org/10.1002/9780470015902.a0021151

Baudet, M., Ortet, P., Gaillard, J.-C., Fernandez, B., Guérin, P., Enjalbal, C., Subra, G., de Groot, A., Barakat, M., Dedieu, A., & Armengaud, J. (2010). Proteomics-based Refinement of *Deinococcus* deserti Genome Annotation Reveals an Unwonted Use of Non-canonical Translation Initiation Codons*. *Molecular & Cellular Proteomics*, *9*(2), 415–426. https://doi.org/10.1074/mcp.M900359-MCP200

Bentchikou, E., Servant, P., Coste, G., & Sommer, S. (2010). A Major Role of the RecFOR Pathway in DNA Double-Strand-Break Repair through ESDSA in *Deinococcus radiodurans*. *PLOS Genetics*, *6*(1), e1000774. https://doi.org/10.1371/journal.pgen.1000774

Blanchard, L., & de Groot, A. (2021). Coexistence of SOS-Dependent and SOS-Independent Regulation of DNA Repair Genes in Radiation-Resistant *Deinococcus* Bacteria. *Cells*, *10*(4), Article 4. https://doi.org/10.3390/cells10040924

Bobay, L.-M., & Ochman, H. (2018). Factors driving effective population size and pan-genome evolution in bacteria. *BMC Evolutionary Biology*, *18*(1), 153. https://doi.org/10.1186/s12862-018-1272-4

Brooks, B. W., & Murray, R. G. E. (1981). Nomenclature for 'Micrococcus radiodurans' and other radiation-resistant cocci: *Deinococcaceae* fam. Nov. And *Deinococcus* gen. Nov., including five species. *International Journal of Systematic Bacteriology*, *31*(3), 353–360. https://doi.org/10.1099/00207713-31-3-353

Bruch, E. M., de Groot, A., Un, S., & Tabares, L. C. (2015). The effect of gamma-ray irradiation on the Mn(II) speciation in *Deinococcus radiodurans* and the potential role of Mn(II)-orthophosphates. *Metallomics : Integrated Biometal Science*, *7*(5), 908–916. https://doi.org/10.1039/c5mt00009b

Callegan, R. P., Noble, M. F., McTernan, P. M., Battista, J. R., Navarro-González, R., McKay, C. P., da Costa, M. S., & Rainey, F. A. (2008). Description of four novel psychrophilic, ionizing radiation-sensitive *Deinococcus* species from alpine

environments. *International Journal of Systematic and Evolutionary Microbiology*, *58*(5), 1252–1258. https://doi.org/10.1099/ijs.0.65405-0

Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P., & Huerta-Cepas, J. (2021). eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Molecular Biology and Evolution*, *38*(12), 5825–5829. https://doi.org/10.1093/molbev/msab293

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, *25*(15), 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Chen, M.-Y., Teng, W.-K., Zhao, L., Hu, C.-X., Zhou, Y.-K., Han, B.-P., Song, L.-R., & Shu, W.-S. (2021). Comparative genomics reveals insights into cyanobacterial evolution and habitat adaptation. *The ISME Journal*, *15*(1), 211–227. https://doi.org/10.1038/s41396-020-00775-z

Chen, W., Wang, B., Hong, H., Yang, H., & Liu, S. J. (2011). *Deinococcus* reticulitermitis sp. Nov., isolated from a termite gut. *International Journal of Systematic and Evolutionary Microbiology*, *62*(1), 78–83. https://doi.org/10.1099/ijs.0.026567-0

Coleman, G. A., Davín, A. A., Mahendrarajah, T. A., Szánthó, L. L., Spang, A., Hugenholtz, P., Szöllsi, G. J., & Williams, T. A. (2021). A rooted phylogeny resolves early bacterial evolution. *Science*, *372*(6542). https://doi.org/10.1126/science.abe0511

Comte, N., Morel, B., Hasić, D., Guéguen, L., Boussau, B., Daubin, V., Penel, S., Scornavacca, C., Gouy, M., Stamatakis, A., Tannier, E., & Parsons, D. P. (2020). Treerecs: An integrated phylogenetic tool, from sequences to reconciliations. *Bioinformatics*, *36*(18), 4822–4824. https://doi.org/10.1093/bioinformatics/btaa615

Contreras-Moreira, B., & Vinuesa, P. (2013). GET_HOMOLOGUES, a Versatile Software Package for Scalable and Robust Microbial Pangenome Analysis. *Applied and Environmental Microbiology*, *79*(24), 7696–7701. https://doi.org/10.1128/AEM.02411-13

Daly, M. J. (2012). Death by protein damage in irradiated cells. *DNA Repair*, *11*(1), 12–21. https://doi.org/10.1016/j.dnarep.2011.10.024

Daly, M. J. (2023). The scientific revolution that unraveled the astonishing DNA repair capacity of the Deinococcaceae: 40 years on. *Canadian Journal of Microbiology*. https://doi.org/10.1139/cjm-2023-0059

Daly, M. J., Gaidamakova, E. K., Matrosova, V. Y., Kiang, J. G., Fukumoto, R., Lee, D. Y., Wehr, N. B., Viteri, G. A., Berlett, B. S., & Levine, R. L. (2010). Small-molecule antioxidant proteome-shields in *Deinococcus radiodurans*. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0012570

Daly, M. J., Gaidamakova, E. K., Matrosova, V. Y., Vasilenko, A., Zhai, M., Venkateswaran, A., Hess, M., Omelchenko, M. V., Kostandarithes, H. M., Makarova, K. S., Wackett, L. P., Fredrickson, J. K., & Ghosal, D. (2004). Accumulation of Mn(II) in *Deinococcus radiodurans* Facilitates Gamma-

Radiation Resistance. *Science*, *306*(5698), 1025–1028. https://doi.org/10.1126/science.1103185

Daly, M. J., Ouyang, L., Fuchs, P., & Minton, K. W. (1994). In vivo damage and recA-dependent repair of plasmid and chromosomal DNA in the radiation-resistant bacterium *Deinococcus radiodurans*. *Journal of Bacteriology*, *176*(12), 3508–3517. https://doi.org/10.1128/jb.176.12.3508-3517.1994

David, L. A., & Alm, E. J. (2011). Rapid evolutionary innovation during an Archaean genetic expansion. *Nature*, *469*(7328), 93–96. https://doi.org/10.1038/nature09649

de Groot, A., Chapon, V., Servant, P., Christen, R., Fischer-Le Saux, M., Sommer, S., & Heulin, T. (2005). *Deinococcus* deserti sp. Nov., a gamma-radiation-tolerant bacterium isolated from the Sahara Desert. *International Journal of Systematic and Evolutionary Microbiology*, *55*(6), 2441–2446. https://doi.org/10.1099/ijs.0.63717-0

de Groot, A., Siponen, M. I., Magerand, R., Eugénie, N., Martin-Arevalillo, R., Doloy, J., Lemaire, D., Brandelet, G., Parcy, F., Dumas, R., Roche, P., Servant, P., Confalonieri, F., Arnoux, P., Pignol, D., & Blanchard, L. (2019). Crystal structure of the transcriptional repressor DdrO: Insight into the metalloprotease/repressor-controlled radiation response in *Deinococcus* . *Nucleic Acids Research*, *47*(21), 11403–11417. https://doi.org/10.1093/nar/gkz883

Dewar, A. E., Thomas, J. L., Scott, T. W., Wild, G., Griffin, A. S., West, S. A., & Ghoul, M. (2021). Plasmids do not consistently stabilize cooperation across bacteria but may promote broad pathogen host-range. *Nature Ecology & Evolution*, 1–13. https://doi.org/10.1038/s41559-021-01573-2

Dharamshi, J. E., Köstlbacher, S., Schön, M. E., Collingro, A., Ettema, T. J. G., & Horn, M. (2023). Gene gain facilitated endosymbiotic evolution of Chlamydiae. *Nature Microbiology*, *8*(1), Article 1. https://doi.org/10.1038/s41564-022-01284-9

Ekman, J. V., Raulio, M., Busse, H.-J., Fewer, D. P., & Salkinoja-Salonen, M. (2011). Deinobacterium chartae gen. Nov., sp. Nov., an extremely radiation-resistant, biofilm-forming bacterium isolated from a Finnish paper mill. *International Journal of Systematic and Evolutionary Microbiology*, *61*(3), 540–548. https://doi.org/10.1099/ijs.0.017970-0

Emms, D. M., & Kelly, S. (2019). OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biology*, *20*(1), 238. https://doi.org/10.1186/s13059-019-1832-y

Emms, D. M., & Kelly, S. (2020). Benchmarking Orthogroup Inference Accuracy: Revisiting Orthobench. *Genome Biology and Evolution*, *12*(12), 2258–2266. https://doi.org/10.1093/gbe/evaa211

Eugénie, N., Zivanovic, Y., Lelandais, G., Coste, G., Bouthier de la Tour, C., Bentchikou, E., Servant, P., & Confalonieri, F. (2021). Characterization of the Radiation Desiccation Response Regulon of the Radioresistant Bacterium *Deinococcus radiodurans* by Integrative Genomic Analyses. *Cells*, *10*(10), 2536. https://doi.org/10.3390/cells10102536

Fang, G.-Y., Chai, L.-J., Zhong, X.-Z., Lu, Z.-M., Zhang, X.-J., Wu, L.-H., Wang, S.-T., Shen, C.-H., Shi, J.-S., & Xu, Z.-H. (2022). Comparative Genomics Unveils the Habitat Adaptation and Metabolic Profiles of Clostridium in an Artificial Ecosystem for Liquor Production. *mSystems*, *7*(3), e00297-22. https://doi.org/10.1128/msystems.00297-22

Ferreira, A. C., Nobre, M. F., Rainey, F. A., Silva, M. T., Wait, R., Burghardt, J., Chung, A. P., & da Costa, M. S. (1997). *Deinococcus* geothermalis sp. Nov. And *Deinococcus* murrayi sp. Nov., two extremely radiation-resistant and slightly thermophilic species from hot springs. *International Journal of Systematic Bacteriology*, *47*(4), 939–947. https://doi.org/10.1099/00207713-47-4-939

Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E. L. L., Tate, J., & Punta, M. (2014). Pfam: The protein families database. *Nucleic Acids Research*, *42*(D1), D222–D230. https://doi.org/10.1093/nar/gkt1223

Gaidamakova, E. K., Sharma, A., Matrosova, V. Y., Grichenko, O., Volpe, R. P., Tkavc, R., Conze, I. H., Klimenkova, P., Balygina, I., Horne, W. H., Gostinčar, C., Chen, X., Makarova, K. S., Shuryak, I., Srinivasan, C., Jackson-Thompson, B., Hoffman, B. M., & Daly, M. J. (2022). Small-Molecule Mn Antioxidants in Caenorhabditis elegans and *Deinococcus radiodurans* Supplant MnSOD Enzymes during Aging and Irradiation. *mBio*, *13*(1), e0339421. https://doi.org/10.1128/mbio.03394-21

Gayen, M., Gupta, P., Morazzani, E. M., Gaidamakova, E. K., Knollmann-Ritschel, B., Daly, M. J., Glass, P. J., & Maheshwari, R. K. (2017). *Deinococcus* Mn2+-peptide complex: A novel approach to alphavirus vaccine development. *Vaccine*, *35*(29), 3672–3681. https://doi.org/10.1016/j.vaccine.2017.05.016

Ghiassi-nejad, M., Mortazavi, S. M. J., Cameron, J. R., Niroomand-rad, A., & Karam, P. A. (2002). Very high background radiation areas of Ramsar, Iran: Preliminary biological studies. *Health Physics*, *82*(1), 87–93. https://doi.org/10.1097/00004032-200201000-00011

Ghosal, D., Omelchenko, M. V., Gaidamakova, E. K., Matrosova, V. Y., Vasilenko, A., Venkateswaran, A., Zhai, M., Kostandarithes, H. M., Brim, H., Makarova, K. S., Wackett, L. P., Fredrickson, J. K., & Daly, M. J. (2005). How radiation kills cells: Survival of *Deinococcus radiodurans* and Shewanella oneidensis under oxidative stress⋆. *FEMS Microbiology Reviews*, *29*(2), 361–375. https://doi.org/10.1016/j.fmrre.2004.12.007

Gupta, P., Gayen, M., Smith, J. T., Gaidamakova, E. K., Matrosova, V. Y., Grichenko, O., Knollmann-Ritschel, B., Daly, M. J., Kiang, J. G., & Maheshwari, R. K. (2016). MDP: A *Deinococcus* Mn2+-Decapeptide Complex Protects Mice from Ionizing Radiation. *PloS One*, *11*(8), e0160575. https://doi.org/10.1371/journal.pone.0160575

Hirsch, P., Gallikowski, C. A., Siebert, J., Peissl, K., Kroppenstedt, R., Schumann, P., Stackebrandt, E., & Anderson, R. (2004). *Deinococcus* frigens sp. Nov., *Deinococcus* saxicola sp. Nov., and *Deinococcus* marmoris sp. Nov., Low Temperature and Draught-tolerating, UV-resistant Bacteria from Continental

Antarctica. *Systematic and Applied Microbiology*, *27*(6), 636–645. https://doi.org/10.1078/0723202042370008

Horne, W. H., Volpe, R. P., Korza, G., DePratti, S., Conze, I. H., Shuryak, I., Grebenc, T., Matrosova, V. Y., Gaidamakova, E. K., Tkavc, R., Sharma, A., Gostinčar, C., Gunde-Cimerman, N., Hoffman, B. M., Setlow, P., & Daly, M. J. (2022). Effects of Desiccation and Freezing on Microbial Ionizing Radiation Survivability: Considerations for Mars Sample Return. *Astrobiology*, *22*(11), 1337–1350. https://doi.org/10.1089/ast.2022.0065

Jeon, S. H., Kang, M. S., Joo, E. S., Kim, E. B., Lim, S., Jeong, S. W., Jung, H. Y., Srinivasan, S., & Kim, M. K. (2016). *Deinococcus* persicinus sp. Nov., a radiationresistant bacterium from soil. *International Journal of Systematic and Evolutionary Microbiology*, *66*(12), 5077–5082. https://doi.org/10.1099/ijsem.0.001473

Jiao, J.-Y., Lian, Z.-H., Li, M.-M., Salam, N., Zhou, E.-M., Liu, L., Ming, H., Nie, G., Shu, W., Zhao, G., Hedlund, B. P., & Li, W.-J. (2022a). Comparative genomic analysis of Thermus provides insights into the evolutionary history of an incomplete denitrification pathway. *mLife*, *1*(2), 198–209. https://doi.org/10.1002/mlf2.12009

Jiao, J.-Y., Lian, Z.-H., Li, M.-M., Salam, N., Zhou, E.-M., Liu, L., Ming, H., Nie, G., Shu, W., Zhao, G., Hedlund, B. P., & Li, W.-J. (2022b). Comparative genomic analysis of Thermus provides insights into the evolutionary history of an incomplete denitrification pathway. *mLife*, *1*(2), 198–209. https://doi.org/10.1002/mlf2.12009

Jin, M., Xiao, A., Zhu, L., Zhang, Z., Huang, H., & Jiang, L. (2019). The diversity and commonalities of the radiation-resistance mechanisms of *Deinococcus* and its up-to-date applications. *AMB Express*, *9*(1), 138. https://doi.org/10.1186/s13568-019-0862-x

Jonkheer, E. M., Brankovics, B., Houwers, I. M., van der Wolf, J. M., Bonants, P. J. M., Vreeburg, R. A. M., Bollema, R., de Haan, J. R., Berke, L., Smit, S., de Ridder, D., & van der Lee, T. A. J. (2021). The Pectobacterium pangenome, with a focus on Pectobacterium brasiliense, shows a robust core and extensive exchange of genes from a shared gene pool. *BMC Genomics*, *22*(1), 265. https://doi.org/10.1186/s12864-021-07583-5

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), Article 6. https://doi.org/10.1038/nmeth.4285

Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, *28*(1), 27–30.

Katoh, K., & Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kawaguchi, Y., Shibuya, M., Kinoshita, I., Yatabe, J., Narumi, I., Shibata, H., Hayashi, R., Fujiwara, D., Murano, Y., Hashimoto, H., Imai, E., Kodaira, S., Uchihori, Y., Nakagawa, K., Mita, H., Yokobori, S., & Yamagishi, A. (2020). DNA Damage and Survival Time Course of Deinococcal Cell Pellets During 3 Years of

Exposure to Outer Space. *Frontiers in Microbiology*, *11*. https://www.frontiersin.org/articles/10.3389/fmicb.2020.02050

Kawaguchi, Y., Yang, Y., Kawashiri, N., Shiraishi, K., Takasu, M., Narumi, I., Satoh, K., Hashimoto, H., Nakagawa, K., Tanigawa, Y., Momoki, Y., Tanabe, M., Sugino, T., Takahashi, Y., Shimizu, Y., Yoshida, S., Kobayashi, K., Yokobori, S., & Yamagishi, A. (2013). 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. *Origins of Life and Evolution of Biospheres*, *43*(4), 411–428. https://doi.org/10.1007/s11084-013-9346-1

Kim, E. B., Kang, M. S., Joo, E. S., Jeon, S. H., Jeong, S. W., Lim, S. Y., Jung, H. Y., Srinivasan, S., & Kim, M. K. (2017). *Deinococcus* ruber sp. Nov., a radiation-resistant bacterium isolated from soil. *International Journal of Systematic and Evolutionary Microbiology*, *67*(1), 72–76. https://doi.org/10.1099/ijsem.0.001567

Krisko, A., & Radman, M. (2013). Biology of Extreme Radiation Resistance: The Way of *Deinococcus radiodurans*. *Cold Spring Harbor Perspectives in Biology*, *5*(7), a012765. https://doi.org/10.1101/cshperspect.a012765

Lai, W. A., Kämpfer, P., Arun, A. B., Shen, F. T., Huber, B., Rekha, P. D., & Young, C. C. (2006). *Deinococcus* ficus sp. Nov., isolated from the rhizosphere of Ficus religiosa L. *International Journal of Systematic and Evolutionary Microbiology*, *56*(4), 787–791. https://doi.org/10.1099/ijs.0.64007-0

Lee, C., Choo, K., & Lee, S.-J. (2020). Active Transposition of Insertion Sequences by Oxidative Stress in *Deinococcus* geothermalis. *Frontiers in Microbiology*, *0*. https://doi.org/10.3389/fmicb.2020.558747

Lee, J. H., Jung, J.-H., Kim, M.-K., & Lim, S. (2022). *Deinococcus* taeanensis sp. Nov., a Radiation-Resistant Bacterium Isolated from a Coastal Dune. *Current Microbiology*, *79*(11), 334. https://doi.org/10.1007/s00284-022-03044-8

Lee, J., Lee, Y. H., Park, S. J., Lim, S., Jeong, S. W., Lee, S.-Y., Park, S., Choi, H.-W., Kim, M. K., & Jung, H.-Y. (2016). *Deinococcus* sedimenti sp. Nov. Isolated from river sediment. *Journal of Microbiology*, *54*(12), 802–808. https://doi.org/10.1007/s12275-016-6361-8

Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, *49*(W1), W293–W296. https://doi.org/10.1093/nar/gkab301

Liao, J., Guo, X., Weller, D. L., Pollak, S., Buckley, D. H., Wiedmann, M., & Cordero, O. X. (2021). Nationwide genomic atlas of soil-dwelling Listeria reveals effects of selection and population ecology on pangenome evolution. *Nature Microbiology*, *6*(8), Article 8. https://doi.org/10.1038/s41564-021-00935-7

Lim, S., Jung, J.-H., Blanchard, L., & de Groot, A. (2018). Conservation and diversity of radiation and oxidative stress resistance mechanisms in *Deinococcus* species. *FEMS Microbiology Reviews*, *43*(1), 19–52. https://doi.org/10.1093/femsre/fuy037

Ludanyi, M., Blanchard, L., Dulermo, R., Brandelet, G., Bellanger, L., Pignol, D., Lemaire, D., & de Groot, A. (2014). Radiation response in *Deinococcus* deserti:

IrrE is a metalloprotease that cleaves repressor protein DdrO. *Molecular Microbiology*, *94*(2), 434–449. https://doi.org/10.1111/mmi.12774

Makarova, K. S., Aravind, L., Wolf, Y. I., Tatusov, R. L., Minton, K. W., Koonin, E. V., & Daly, M. J. (2001). Genome of the Extremely Radiation-Resistant Bacterium *Deinococcus radiodurans* Viewed from the Perspective of Comparative Genomics. *Microbiology and Molecular Biology Reviews*, *65*(1), 44–79. https://doi.org/10.1128/mmbr.65.1.44-79.2001

Makarova, K. S., & Daly, M. J. (2010). Comparative Genomics of Stress Response Systems in *Deinococcus* Bacteria. In *Bacterial Stress Responses* (pp. 445–457). John Wiley & Sons, Ltd. https://doi.org/10.1128/9781555816841.ch27

Makarova, K. S., Omelchenko, M. V., Gaidamakova, E. K., Matrosova, V. Y., Vasilenko, A., Zhai, M., Lapidus, A., Copeland, A., Kim, E., Land, M., Mavromatis, K., Pitluck, S., Richardson, P. M., Detter, C., Brettin, T., Saunders, E., Lai, B., Ravel, B., Kemner, K. M., … Daly, M. J. (2007). *Deinococcus* geothermalis: The Pool of Extreme Radiation Resistance Genes Shrinks. *PLoS ONE*, *2*(9), e955. https://doi.org/10.1371/journal.pone.0000955

Makarova, K. S., Wolf, Y. I., White, O., Minton, K., & Daly, M. J. (1999). Short repeats and IS elements in the extremely radiation-resistant bacterium *Deinococcus radiodurans* and comparison to other bacterial species. *Research in Microbiology*, *150*(9–10), 711–724. https://doi.org/10.1016/S0923-2508(99)00121-7

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, *37*(5), 1530–1534. https://doi.org/10.1093/molbev/msaa015

Mohseni, M., Abbaszadeh, J., & Nasrollahi Omran, A. (2014). Radiation resistant of native *Deinococcus* spp. Isolated from the Lout desert of Iran "the hottest place on Earth". *International Journal of Environmental Science and Technology*, *11*(7), 1939–1946. https://doi.org/10.1007/s13762-014-0643-7

Monfared, A. S., Jalali, F., Mozdarani, H., Hajiahmadi, M., & Samavat, H. (2005). Living in high natural background radiation areas in Ramsar, Iran. Is it dangerous for health? *International Congress Series*, *1276*, 438–439. https://doi.org/10.1016/j.ics.2004.12.007

Mu, D.-S., Wang, S., Liang, Q.-Y., Du, Z.-Z., Tian, R., Ouyang, Y., Wang, X.-P., Zhou, A., Gong, Y., Chen, G.-J., Van Nostrand, J., Yang, Y., Zhou, J., & Du, Z.-J. (2020). Bradymonabacteria, a novel bacterial predator group with versatile survival strategies in saline environments. *Microbiome*, *8*(1), 126. https://doi.org/10.1186/s40168-020-00902-0

Murray, R. G. E. (1992). The Family Deinococcaceae. In *The Prokaryotes* (pp. 3732–3744). Springer New York. https://doi.org/10.1007/978-1-4757-2191-1_42

Narasimha, A., & Basu, B. (2021). New insights into the activation of Radiation Desiccation Response regulon in *Deinococcus radiodurans*. *Journal of Biosciences*, *46*(1), 10. https://doi.org/10.1007/s12038-020-00123-5

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood

Phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274. https://doi.org/10.1093/molbev/msu300

Omelchenko, M. V., Wolf, Y. I., Gaidamakova, E. K., Matrosova, V. Y., Vasilenko, A., Zhai, M., Daly, M. J., Koonin, E. V., & Makarova, K. S. (2005). Comparative genomics of Thermus thermophilus and *Deinococcus* radiodurans: Divergent routes of adaptation to thermophily and radiation resistance. *BMC Evolutionary Biology*, *5*, 57. https://doi.org/10.1186/1471-2148-5-57

Panitz, C., Frösler, J., Wingender, J., Flemming, H.-C., & Rettberg, P. (2019). Tolerances of *Deinococcus* geothermalis Biofilms and Planktonic Cells Exposed to Space and Simulated Martian Conditions in Low Earth Orbit for Almost Two Years. *Astrobiology*, *19*(8), 979–994. https://doi.org/10.1089/ast.2018.1913

Parks, D. H., Chuvochina, M., Rinke, C., Mussig, A. J., Chaumeil, P.-A., & Hugenholtz, P. (2022). GTDB: An ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Research*, *50*(D1), D785–D794. https://doi.org/10.1093/nar/gkab776

Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P. A., & Hugenholtz, P. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nature Biotechnology*. https://doi.org/10.1038/nbt.4229

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research*, *25*(7), 1043–1055. https://doi.org/10.1101/gr.186072.114

Pasternak, C., Ton-Hoang, B., Coste, G., Bailone, A., Chandler, M., & Sommer, S. (2010). Irradiation-Induced *Deinococcus radiodurans* Genome Fragmentation Triggers Transposition of a Single Resident Insertion Sequence. *PLOS Genetics*, *6*(1), e1000799. https://doi.org/10.1371/journal.pgen.1000799

Pavlov, A. K., Kalinin, V. L., Konstantinov, A. N., Shelegedin, V. N., & Pavlov, A. A. (2006). Hypothesis Paper Was Earth Ever Infected by Martian Biota? Clues from Radioresistant Bacteria. In *ASTROBIOLOGY* (6; Vol. 6). www.liebertpub.com

Raposo, P., Viver, T., Albuquerque, L., Froufe, H., Barroso, C., Egas, C., Rosselló-Móra, R., & da Costa, M. S. (2019). Transfer of Meiothermus chliarophilus (Tenreiro et al.1995) Nobre et al. 1996, Meiothermus roseus Ming et al. 2016, Meiothermus terrae Yu et al. 2014 and Meiothermus timidus Pires et al. 2005, to Calidithermus gen. Nov., as Calidithermus chliarophilus comb. Nov., Calidithermus roseus comb. Nov., Calidithermus terrae comb. Nov. And Calidithermus timidus comb. Nov., respectively, and emended description of the genus Meiothermus. *International Journal of Systematic and Evolutionary Microbiology*, *69*(4), 1060–1069. https://doi.org/10.1099/ijsem.0.003270

Reback, J., McKinney, W., jbrockmendel, Bossche, J. V. den, Augspurger, T., Cloud, P., gfyoung, Sinhrks, Klein, A., Roeschke, M., Tratner, J., She, C., Ayd, W., Hawkins, S., Petersen, T., Schendel, J., Hayden, A., Garcia, M., Jancauskas, V., … Kluyver, T. (2020). *pandas-dev/pandas: Pandas 1.0.0* (v1.0.0) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.3630805

Rosenberg, E. (2014). The family deinococcaceae. In *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea* (pp. 613–615). Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-642-38954-2_127

Sako, Y., Nakagawa, S., Takai, K., & Horikoshi, K. (2003). Marinithermus hydrothermalis gen. Nov., sp. Nov., a strictly aerobic, thermophilic bacterium from a deep-sea hydrothermal vent chimney. *International Journal of Systematic and Evolutionary Microbiology*, *53*(Pt 1), 59–65. https://doi.org/10.1099/ijs.0.02364-0

Scornavacca, C., Mayol, J. C. P., & Cardona, G. (2017). Fast algorithm for the reconciliation of gene trees and LGT networks. *Journal of Theoretical Biology*, *418*, 129–137. https://doi.org/10.1016/j.jtbi.2017.01.024

Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*, *30*(14), 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Selengut, J. D., Haft, D. H., Davidsen, T., Ganapathy, A., Gwinn-Giglio, M., Nelson, W. C., Richter, A. R., & White, O. (2007). TIGRFAMs and Genome Properties: Tools for the assignment of molecular function and biological process in prokaryotic genomes. *Nucleic Acids Research*, *35*(Database issue), D260-264. https://doi.org/10.1093/nar/gkl1043

Sharma, A., Gaidamakova, E. K., Grichenko, O., Matrosova, V. Y., Hoeke, V., Klimenkova, P., Conze, I. H., Volpe, R. P., Tkavc, R., Gostinčar, C., Gunde-Cimerman, N., Diruggiero, J., Shuryak, I., Ozarowski, A., Hoffman, B. M., Daly, M. J., Designed, M. J. D., & Performed, A. O. (2017). *Across the tree of life, radiation resistance is governed by antioxidant Mn 2+ , gauged by paramagnetic resonance*. https://doi.org/10.1073/pnas.1713608114

Shashidhar, R., & Bandekar, J. R. (2009). *Deinococcus* piscis sp. Nov., a radiation-resistant bacterium isolated from a marine fish. *International Journal of Systematic and Evolutionary Microbiology*, *59*(11), 2714–2717. https://doi.org/10.1099/ijs.0.003046-0

Shashidhar, R., Kumar, S. A., Misra, H. S., & Bandekar, J. R. (2010). Evaluation of the role of enzymatic and nonenzymatic antioxidant systems in the radiation resistance of *Deinococcus* . *Canadian Journal of Microbiology*, *56*(3), 195–201. https://doi.org/10.1139/w09-118

Sheppard, S. K., Guttman, D. S., & Fitzgerald, J. R. (2018). Population genomics of bacterial host adaptation. *Nature Reviews Genetics*, *19*(9), Article 9. https://doi.org/10.1038/s41576-018-0032-z

Sheridan, P. O., Raguideau, S., Quince, C., Holden, J., Zhang, L., Williams, T. A., & Gubry-Rangin, C. (2020). Gene duplication drives genome expansion in a major lineage of Thaumarchaeota. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-19132-x

Shimoyama, Y. (2021, October 21). *Moshi4/FastDTLmapper*. GitHub. https://github.com/moshi4/FastDTLmapper/blob/main/CITATION.cff

Slade, D., Lindner, A. B., Paul, G., & Radman, M. (2009). Recombination and Replication in DNA Repair of Heavily Irradiated *Deinococcus radiodurans. Cell*, *136*(6), 1044–1055. https://doi.org/10.1016/j.cell.2009.01.018

Slade, D., & Radman, M. (2011). Oxidative Stress Resistance in *Deinococcus radiodurans. Microbiology and Molecular Biology Reviews*, *75*(1), 133–191. https://doi.org/10.1128/mmbr.00015-10

Snipen, L., & Liland, K. H. (2015). micropan: An R-package for microbial pan-genomics. *BMC Bioinformatics*, *16*(1), 79. https://doi.org/10.1186/s12859-015-0517-0

Sturmbauer, C. (2013). Paraphyly. In S. Maloy & K. Hughes (Eds.), *Brenner's Encyclopedia of Genetics (Second Edition)* (pp. 225–226). Academic Press. https://doi.org/10.1016/B978-0-12-374984-0.01119-0

Suh, A. (2019). Genome Size Evolution: Small Transposons with Large Consequences. *Current Biology*, *29*(7), R241–R243. https://doi.org/10.1016/j.cub.2019.02.032

Suresh, K., Reddy, G. S. N., Sengupta, S., & Shivaji, S. (2004). *Deinococcus* indicus sp. Nov., an arsenic-resistant bacterium from an aquifer in West Bengal, India. *International Journal of Systematic and Evolutionary Microbiology*, *54*(2), 457–461. https://doi.org/10.1099/ijs.0.02758-0

Szöllősi, G. J., Rosikiewicz, W., Boussau, B., Tannier, E., & Daubin, V. (2013). Efficient Exploration of the Space of Reconciled Gene Trees. *Systematic Biology*, *62*(6), 901–912. https://doi.org/10.1093/sysbio/syt054

Tettelin, H., Riley, D., Cattuto, C., & Medini, D. (2008). Comparative genomics: The bacterial pan-genome. *Current Opinion in Microbiology*, *11*(5), 472–477. https://doi.org/10.1016/j.mib.2008.09.006

Tobin, G. J., Tobin, J. K., Gaidamakova, E. K., Wiggins, T. J., Bushnell, R. V., Lee, W.-M., Matrosova, V. Y., Dollery, S. J., Meeks, H. N., Kouiavskaia, D., Chumakov, K., & Daly, M. J. (2020). A novel gamma radiation-inactivated sabin-based polio vaccine. *PLOS ONE*, *15*(1), e0228006. https://doi.org/10.1371/journal.pone.0228006

Touchon, M., & Rocha, E. P. C. (2007). Causes of Insertion Sequences Abundance in Prokaryotic Genomes. *Molecular Biology and Evolution*, *24*(4), 969–981. https://doi.org/10.1093/molbev/msm014

Van Rossum, T., Ferretti, P., Maistrenko, O. M., & Bork, P. (2020). Diversity within species: Interpreting strains in microbiomes. *Nature Reviews Microbiology*, *18*(9), Article 9. https://doi.org/10.1038/s41579-020-0368-1

Vigil-Stenman, T., Ininbergs, K., Bergman, B., & Ekman, M. (2017). High abundance and expression of transposases in bacteria from the Baltic Sea. *The ISME Journal*, *11*(11), Article 11. https://doi.org/10.1038/ismej.2017.114

von Meijenfeldt, F. A. B., Hogeweg, P., & Dutilh, B. E. (2023). A social niche breadth score reveals niche range strategies of generalists and specialists. *Nature Ecology & Evolution*, *7*(5), Article 5. https://doi.org/10.1038/s41559-023-02027-7

Wang, L.-G., Lam, T. T.-Y., Xu, S., Dai, Z., Zhou, L., Feng, T., Guo, P., Dunn, C. W., Jones, B. R., Bradley, T., Zhu, H., Guan, Y., Jiang, Y., & Yu, G. (2020). Treeio: An R Package for Phylogenetic Tree Input and Output with Richly Annotated and Associated Data. *Molecular Biology and Evolution*, *37*(2), 599–603. https://doi.org/10.1093/molbev/msz240

Wang, X.-P., Li, C.-M., Yu, Y., Li, H.-R., Du, Z.-J., & Mu, D. 2019. (n.d.). *Deinococcus* arcticus sp. Nov., isolated from Silene acaulis rhizosphere soil of the Arctic

tundra. *International Journal of Systematic and Evolutionary Microbiology*, *69*(11), 3437–3442. https://doi.org/10.1099/ijsem.0.003636

White, O., Eisen, J. A., Heidelberg, J. F., Hickey, E. K., Peterson, J. D., Dodson, R. J., Haft, D. H., Gwinn, M. L., Nelson, W. C., Richardson, D. L., Moffat, K. S., Qin, H., Jiang, L., Pamphile, W., Crosby, M., Shen, M., Vamathevan, J. J., Lam, P., McDonald, L., … Fraser, C. M. (1999). Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. *Science (New York, N.Y.)*, *286*(5444), 1571–1577. https://doi.org/10.1126/science.286.5444.1571

Wick, R. R., & Holt, K. E. (2022). Polypolish: Short-read polishing of long-read bacterial genome assemblies. *PLOS Computational Biology*, *18*(1), e1009802. https://doi.org/10.1371/journal.pcbi.1009802

Wick, R. R., Judd, L. M., Cerdeira, L. T., Hawkey, J., Méric, G., Vezina, B., Wyres, K. L., & Holt, K. E. (2021). Trycycler: Consensus long-read assemblies for bacterial genomes. *Genome Biology*, *22*(1), 266. https://doi.org/10.1186/s13059-021-02483-z

Wick, R. R., Judd, L. M., & Holt, K. E. (2019). Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biology*, *20*(1), 129. https://doi.org/10.1186/s13059-019-1727-y

Williams, T. A., Davin, A. A., Morel, B., Szánthó, L. L., Spang, A., Stamatakis, A., Hugenholtz, P., & Szöllősi, G. J. (2023). *The power and limitations of species tree-aware phylogenetics* [Preprint]. Evolutionary Biology. https://doi.org/10.1101/2023.03.17.533068

Yamagishi, A., Kawaguchi, Y., Hashimoto, H., Yano, H., Imai, E., Kodaira, S., Uchihori, Y., & Nakagawa, K. (2018). Environmental Data and Survival Data of *Deinococcus* aetherius from the Exposure Facility of the Japan Experimental Module of the International Space Station Obtained by the Tanpopo Mission. *Astrobiology*, *18*(11), 1369–1374. https://doi.org/10.1089/ast.2017.1751

Yang, Y., Itoh, T., Yokobori, S., Itahashi, S., Shimada, H., Satoh, K., Ohba, H., Narumi, I., & Yamagishi, A. (2009). *Deinococcus* aerius sp. Nov., isolated from the high atmosphere. *International Journal of Systematic and Evolutionary Microbiology*, *59*(8), 1862–1866. https://doi.org/10.1099/ijs.0.007963-0

Yin, L.-Z., Li, J.-L., Liu, Z.-T., Fang, B.-Z., Wang, P., Luo, X.-Q., Dong, L., Duan, L., Li, S.-H., & Li, W.-J. (2022). *Deinococcus* aestuarii sp. Nov. And *Deinococcus* aquaedulcis sp. Nov., two novel resistant bacteria isolated from pearl river estuary. *Antonie van Leeuwenhoek*, *115*(1), 59–68. https://doi.org/10.1007/s10482-021-01680-x

Yu, G. (2020). Using ggtree to Visualize Data on Tree-Like Structures. *Current Protocols in Bioinformatics*, *69*(1), e96. https://doi.org/10.1002/cpbi.96

Zhang, K., Zhu, J., Li, S., Rao, M. P. N., Li, N.-M., Guo, A.-Y., Jiang, Z., Tang, Q.-Y., Wan, Y., Zhang, Z.-D., & Li, W.-J. (2020a). *Deinococcus* detaillensis sp. Nov., isolated from humus soil in Antarctica. *Archives of Microbiology*, *202*(9), 2493–2498. https://doi.org/10.1007/s00203-020-01920-0

Zhang, K., Zhu, J., Li, S., Rao, M. P. N., Li, N.-M., Guo, A.-Y., Jiang, Z., Tang, Q.-Y., Wan, Y., Zhang, Z.-D., & Li, W.-J. (2020b). *Deinococcus* detaillensis sp. Nov.,

isolated from humus soil in Antarctica. *Archives of Microbiology*, *202*(9), 2493–2498. https://doi.org/10.1007/s00203-020-01920-0

Zhou, Z., Tran, P. Q., Breister, A. M., Liu, Y., Kieft, K., Cowley, E. S., Karaoz, U., & Anantharaman, K. (2022). METABOLIC: High-throughput profiling of microbial genomes for functional traits, metabolism, biogeochemistry, and community-scale functional networks. *Microbiome*, *10*(1), 33. https://doi.org/10.1186/s40168-021-01213-8

5 Chapter V - Summary, Conclusions, and Future Work

5.1 Background and aims

The *Deinococcaceae* family is one of the most fascinating groups of microorganisms on Earth. However, there have been many misunderstandings in the past seven decades since they were discovered. The isolation of this group of bacteria went back to 1956 when scientists were studying gamma radiation to sterilise meat packages using doses that were then believed to be high enough to eliminate all life forms (Anderson et al., 1956). The most studied member of this group, *Deinococcus radiodurans*, was isolated from those experiments and named *Micrococcus radiodurans* because of its morphological similarities to the *Micrococcus* genus, especially tetrad formation (Dean et al., 1966; Krabbenhoft et al., 1965). It took 35 years for microbiologists to reclassify five radio-resistant "*Micrococcus*" species and propose a new family for this group of bacteria. In 1981, Brooks and Murray suggested renaming *Micrococcus radiodurans* to *Deinococcus radiodurans* and classified it under the *Deinococcaceae* family. The etymology used for naming this group indicated the peculiarity of this genus and family of bacteria since it was named after the Greek adjective *deinos*, which means strange or unusual (Brooks & Murray, 1981).

While scoping the research for this thesis, I noticed a potentially biased point of view about the *Deinococcaceae* family. The fascination of microbiologists about radiation resistance in this group of bacteria led to extensive research on mechanisms that confer this phenotype to bacteria. Therefore, for more than four decades, this family has been viewed only through the lens of radiation resistance, and their ecology, as well as their evolutionary history, has remained essentially unknown. Nonetheless, despite the tremendous effort put in by microbiologists, not all aspects of radiation resistance mechanisms have been fully discovered. In the latest comprehensive review article of *Deinococcaceae,* a pioneer of the field, Michael J Daly, wrote: "While the number of *Deinococcus* genomes is now sufficiently large to determine the core genome and pangenome of the family *Deinococcaceae,* extreme resistance persistently remains a phenotype without a genotype" (Daly, 2023).

My literature review of the *Deinococcus* species strongly indicated that researchers have always focused on the radiation resistance phenotype to the extent that radiation resistance is widely known as the defining characteristic of the *Deinococcaceae* family (Daly, 2023). This paradigm has shaped our understanding of this group of bacteria. As a consequence, except for a few studies, almost all isolation and cultivation reports of *Deinococcus* species used gamma radiation as a pre-culture treatment to eradicate all radiosensitive cells of the environmental samples to enrich radiation-resistant microorganisms (Asker et al., 2009; Ferreira et al., 1997; Lee et al., 2022; Mohseni et al., 2014; Rainey et al., 2005)

Nevertheless, some studies have accidentally isolated *Deinococcus* species without using gamma radiation as pre-culture treatment (Callegan et al., 2008; Dong et al., 2015; Wang et al., 2019). Some other studies used low levels of gamma radiation for cultivation studies (Jeon et al., 2016). As a result, several radiation-sensitive *Deinococcus* species have been isolated. For instance, *D. persicinus*, isolated from soil samples, had a D_{10} (A dose that 10 per cent of the population survives) value of 1.5 kGy, which is ten times more sensitive than well-studied species such as *D. radiodurans* and *D. geothermalis.* However, the species with low radiation resistance levels are seemingly perceived as not exciting enough, and no studies have explicitly compared them to highly resistant species.

5.2 Thesis summary.

In this thesis, I attempted to address this widespread misconception in the current view of microbiologists about the *Deinococcaceae* family and shift the focus to the ecological strategies and evolutionary history of *Deinococcus* species. I conducted comparative genomic analyses with three complementary approaches to address this general gap in our knowledge about the *Deinococcaceae* family. Each approach constituted a chapter of this thesis in the following order.

5.2.1 Chapter 2 – findings and conclusions

This chapter aims to explore the genomic diversity of the currently known *Deinococcus* species and to answer the following question: what factors contribute to the widespread distribution and diverse characteristics of the *Deinococcus* genus? I hypothesised that the ubiquity of *Deinococcus* is due to its genetic flexibility and large pangenome. To answer this question, I conducted a comparative genomic analysis with a pangenome approach and collated the genomic repertoire of *Deinococcus* members. My results revealed highly diverse genomic features that shape the open pangenome of *Deinococcaceae* and explain their widespread distribution. Notably, amino acid and carbohydrate transporters and various families of transposase enzymes constituted a significant portion of the accessory genome, indicative of phenotype variability among members. Comparing metabolic capabilities showed scarcity in cofactors, vitamins, and iron and manganese oxidation metabolisms. Moreover, specific ecological trends, such as nitrite reduction to ammonia in freshwater-isolated species and conserved sulfate reduction pathways in most members except a few host-associated species, underscored the diverse ecological adaptations within the *Deinococcus* genus.

This chapter concluded that available genomic data for *Deinococcus* species might be large enough to construct the core genome of this genus (Daly, 2023) but is likely not representative of the whole genomic repertoire of the *Deinococcus* genus. Therefore, it is necessary to develop novel methods for isolating *Deinococcus* species to avoid the disproportionate known *Deinococcus* species in terms of their ecological features. To my knowledge, this is the first pangenome study of all sequenced *Deinococcus* species.

5.2.2 Chapter 3 – findings and conclusions

As described in chapter two of this thesis, the accessory genome is associated with the phenotypic diversity of a species or genus. In this chapter, I build on this concept and try to answer the following question: What drives the diversity of radiation resistance in the *Deinococcus* genus and Is IR resistance predictable using comparative genomics methods? To answer these questions, I hypothesised that the analysis of diverse genomic samples in the *Deinococcus* genus can reveal a correlation between radiation

resistance phenotype and genomic content, suggesting a genomic basis for this phenotype. I demonstrated that *Deinococcus* has an open pangenome and high genomic diversity. One of the enigmatic facts about *Deinococcaceae* is that no unique genotype has been identified for the radiation resistance phenotype and that radiation resistance is highly inconsistent, even between the different strains of a species. The focus on the known radiation resistance mechanisms, which mainly originated from studies on *D. radiodurans*, has failed to address this phenotype.

This chapter aims to address this gap by combining phylogenomic analyses and statistical methods to identify the correlation between the presence and absence of genes and the radiation resistance of the corresponding genome. The phylogenomic relationship of *Deinococcus* species showed that radiation-sensitive *Deinococcus* species formed two distinct taxonomic clades and one single branch. Statistical analysis of the accessory genes showed that 188 gene families were evenly distributed between those two with clades and correlated with lower levels of radiation resistance.

Functional annotation of these genes highlighted their association with redox co-factors and oxidative agents, likely explaining the sensitivity of these species. This chapter also suggests that radiation resistance is likely an ancestral trait, as it is observed in the *Trueperaceae* family, a distant relative of the *Deinococcaceae* family. However, as discussed in Chapter 2, the *Deinococcus* genus possesses a flexible genome due to the high diversity of transposases, allowing it to acquire new functions in different environments. Consequently, in the absence of evolutionary pressure to maintain resistance to oxidative stress, this trait may be altered in favour of more beneficial functions, such as energy generation in oligotrophic environments. These functions can increase ROS in the presence of oxidative stress, leading to radiation sensitivity in some *Deinococcus* species. An in-depth analysis of evolutionary events and ancestral reconstruction could provide further insights into the evolutionary origins of radiation resistance in the common ancestor of *Deinococcaceae* and *Trueperaceae*.

In conclusion, the study suggests that specific gene acquisition induced radiation sensitivity in *Deinococcus* species. Due to the genomic flexibility of *Deinococcus*, and

since radiation resistance is not a natural evolutionary pressure on Earth, gaining certain genes that can help the organism evolve in new environments can become toxic under extreme levels of oxidative stress.

5.2.3 Chapter 4 – findings and conclusions

Based on the findings of chapters two and three, I hypothesised that the emergence of the *Deinococcaceae* family is associated with a notable genome expansion, which facilitated their genetic flexibility.

This chapter aims to gain insight into the evolutionary history of the *Deinococcota* phylum to infer evolutionary events that led to the emergence of *Deinococcaceae*. Previous comparative genomic studies had failed to address the radiation resistance phenotype of *Deinococcaceae* members, and the focus on radiation resistance arguably derailed further attempts. Since 2007, no comprehensive comparative genomics study has exclusively examined the evolution of *Deinococcaceae* as a family. To address this knowledge gap, I used genomes of the *Deinococcota* phylum to infer evolutionary events leading to the common ancestor of *Deinococcaceae*, which I called proto-*Deinococcus*.

The findings of this study revealed significant differences in genomic features of the two highly populated families, *Deinococcaceae* and *Thermaceae*. Members of *Deinococcaceae* exhibited more diverse and larger genome sizes, suggesting a broader adaptive potential and a generalist lifestyle. In comparison, *Thermaceae* members had more homogeneous and smaller genome sizes, indicating a specialist lifestyle with a preference for thermophilic environments.

The reconstruction of the common ancestor of *Deinococcaceae* highlighted a substantial gene gain event with 1,096 genes, particularly involving transcriptional regulators related to oxidative stress resistance response, amino acid, and carbohydrate transporters. Moreover, gene loss events were observed in the coenzyme and cofactors transport and metabolisms in the common ancestor of *Trueperaceae* and *Deinococcaceae* (N006), which can indicate the emergence of radiation resistance in

this lineage. These gained and lost genes likely played a crucial role in the survival of the N006 and the proto-*Deinococcus* under oxidative stress, providing proteome protection and ultimately enabling the expansion of its genome and adaptation to diverse habitats. The emergence of radiation resistance in the N006 node and conservation of this phenotype in the descendants confirms that it is an ancestral trait, but gaining certain genes can alter this phenotype in some species.

Overall, the study contributed valuable insights into the evolution of *Deinococcaceae*, shedding light on the genomic dynamics that have shaped the unique characteristics of this bacterial family.

5.3 Limitations

Since the beginning of this thesis, some limitations hindered us from conducting the research we planned in the first place. One big challenge in doing this PhD thesis was the COVID-19 pandemic. I had to restructure the whole thesis two years into my PhD programme because of some failed plans. The initial proposal of this thesis was based on three concepts: (i) comparative genomics and metagenomics, (ii) cultivation of *Deinococcus* without gamma radiation as pre-cultivation enrichment, and (iii) comparative transcriptomics between radiation-resistant and radiation-sensitive *Deinococcus* species to study the difference of response in those groups of *Deinococcus* species. However, it was impossible to achieve some of these goals due to multiple COVID-19 lockdowns, a lack of required instruments for gamma radiation exposure in New Zealand, and a long-time border closure. Instead, I focused on comparative genomics and restructured the thesis using available resources.

The main technical limitation of this study was the lack of diversity among known and sequenced *Deinococcus* species. First and foremost, the cultivation method bias described in this thesis has impacted the whole population of known *Deinococcus* species, and a large group of *Deinococcus* species have likely remained undiscovered. Second, we considered including MAGs in our analysis, but the limited number of

MAGs in older versions of GTDB and low genomic completeness could compromise the pangenome construction and further analyses.

The next issue is sequencing bias, where focusing on a few radiation-resistance species has resulted in an unbalanced availability of *Deinococcus* genomes in public databases. For instance, among the 180 available *Deinococcus* genomes in NCBI GenBank, 17 genomes belong to *Deinococcus radiodurans* R1, and 20 are related to *Deinococcus wulumuqiensis* R12. These sequences are usually modified genomes used for knockout experiments and form a large portion of genomic data in the public database. Meanwhile, many known *Deinococcus* species have not been sequenced. Out of 94 named and published *Deinococcus* members, only 65 have been sequenced. I sequenced two radiation-sensitive *Deinococcus* species for this study, which helped shed light on a phylogenomic pattern in radiation resistance. Another problem in studying the *Deinococcaceae* family is the scarcity of metagenomic data. Even though *Deinococcaceae* are ubiquitous and present in many metagenomic sequences, the low relative abundance of this family has made it difficult to draw a meaningful conclusion from those metagenomic data.

Another challenge of this PhD thesis, especially in chapter 4, was limited computational resources at the end of my PhD project. Gene tree-species tree reconciliation is a computationally expensive analysis. In the first run of my analysis, I used a popular tool called ALE (amalgamation likelihood estimation) to infer evolutionary events and reconstruct the common ancestor of *Deinococcaceae*. I used a broad range of phyla in that analysis, including *Deinococcota*, *Synergistota*, *Thermotogota*, and *Fusobacterota*. For this analysis using ALE, I used more than 24000 core hours in the HPC cluster, which is twice the amount of allocated node hours in my previous year, and the result was not properly interpretable (probably because of genome sampling bias). Given the superiority of probabilistic frameworks such as ALE, it is computationally more expensive and interpreting the data is more complicated.

Moreover, even though ALE provides a probability score for inferred genes, it is not necessarily the exact evolutionary history that happened in the real world, and all such

reconstructions are hypothetical and speculative. To address this limitation and deliver this thesis, I chose a more accessible tool (AnGST), which is based on a parsimony framework (David & Alm, 2011). I also limited the number of genomic samples to a subset of our main dataset. However, analysing the full dataset using both tools is in progress and is our goal for future publication. We also use ALE on the full dataset and include the result in an ensemble approach with AnGST.

5.4 Future work

This thesis proposed and validated a novel direction for studying *Deinococcus* species, and there are many more avenues to explore. The second chapter of this thesis provided a list of 188 gene families that significantly correlate with radiation sensitivity in two clades of *Deinococcus* species. I investigated their functional roles in silico, but one could formulate hypotheses based on my findings and use CRISPR-Cas or other cloning methods to investigate these genes and determine whether specific genes or combinations of them can cause radiation sensitivity. I submitted a proposal to the Australian Institute of Nuclear Science and Engineering (AINSE) in 2020 and suggested comparative transcriptomics analysis to compare the response of radiation-resistant and sensitive *Deinococcus* species to oxidative stress.

Chapter four of this thesis provides nuanced information on the evolutionary history of the *Deinococcota* phylum and the emergence of the common ancestor of *Deinococcaceae*. Ancestral reconstruction analyses should always be treated with caution. Due to computation resource limitations, I had to use a subset of our data for this analysis. Moreover, I used AnGST software for gene tree-species tree reconciliation analysis because of its accessibility through the FastDTLmapper pipeline. AnGST is a parsimony framework-based software (David & Alm, 2011). It relies on identifying the minimum number of evolutionary events required to reconcile a gene tree. Recently, probabilistic-based methods such as the ALE have gained popularity and can draw a more precise picture of evolutionary history by providing the probability of each gene present in the common ancestor (Boussau & Scornavacca, 2020).

Overall, this thesis represents a systematic and hypothesis-driven attempt to better understand an enigmatic bacterial phylum through an unconventional point of view. I argue that the same approach can be extended to other phyla with limited ecological information.

5.5 References

Anderson, A., Nordon, H., Cain, R. F., Parrish, G., Duggan, D., Nordan, H., Parish, G., & Cullum-Dugan, D. (1956). Studies on a radio-resistant micrococcus. I. Isolation, morphology, cultural characteristics, and resistance to gamma radiation. *Food Technology*, *10*, 575–578.

Asker, D., Awad, T. S., Beppu, T., & Ueda, K. (2009). *Deinococcus* aquiradiocola sp. Nov., isolated from a radioactive site in Japan. *International Journal of Systematic and Evolutionary Microbiology*, *59*(Pt 1), 144–149. https://doi.org/10.1099/ijs.0.65762-0

Boussau, B., & Scornavacca, C. (2020). Reconciling Gene trees with Species Trees. In C. Scornavacca, F. Delsuc, & N. Galtier (Eds.), *Phylogenetics in the Genomic Era* (p. 3.2:1-3.2:23). No commercial publisher | Authors open access book. https://hal.science/hal-02535529

Brooks, B. W., & Murray, R. G. E. (1981). Nomenclature for 'Micrococcus *radiodurans*' and other radiation-resistant cocci: *Deinococcaceae* fam. Nov. And *Deinococcus* gen. Nov., including five species. *International Journal of Systematic Bacteriology*, *31*(3), 353–360. https://doi.org/10.1099/00207713-31-3-353

Callegan, R. P., Noble, M. F., McTernan, P. M., Battista, J. R., Navarro-González, R., McKay, C. P., da Costa, M. S., & Rainey, F. A. (2008). Description of four novel psychrophilic, ionising radiation-sensitive *Deinococcus* species from alpine environments. *International Journal of Systematic and Evolutionary Microbiology*, *58*(5), 1252–1258. https://doi.org/10.1099/ijs.0.65405-0

Daly, M. J. (2023). The scientific revolution that unraveled the astonishing DNA repair capacity of the *Deinococcaceae*: 40 years on. *Canadian Journal of Microbiology*. https://doi.org/10.1139/cjm-2023-0059

David, L. A., & Alm, E. J. (2011). Rapid evolutionary innovation during an Archaean genetic expansion. *Nature*, *469*(7328), Article 7328. https://doi.org/10.1038/nature09649

Dean, C. J., Feldschreiber, P., & Lett, J. T. (1966). Repair of x-ray damage to the deoxyribonucleic acid in Micrococcus *radiodurans*. *Nature*, *209*(5018), 49–52. https://doi.org/10.1038/209049a0

Dong, N., Li, H.-R., Yuan, M., Zhang, X.-H., & Yu, Y. (2015). *Deinococcus* antarcticus sp. Nov., isolated from soil. *INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY*, *65*(Pt 2), 331–335. https://doi.org/10.1099/ijs.0.066324-0

Ferreira, A. C., Nobre, M. F., Rainey, F. A., Silva, M. T., Wait, R., Burghardt, J., Chung, A. P., & da Costa, M. S. (1997). *Deinococcus geothermalis* sp. Nov. And *Deinococcus* murrayi sp. Nov., two extremely radiation-resistant and slightly

thermophilic species from hot springs. *International Journal of Systematic Bacteriology*, *47*(4), 939–947. https://doi.org/10.1099/00207713-47-4-939

Jeon, S. H., Kang, M. S., Joo, E. S., Kim, E. B., Lim, S., Jeong, S. W., Jung, H. Y., Srinivasan, S., & Kim, M. K. (2016). *Deinococcus* persicinus sp. Nov., a radiationresistant bacterium from soil. *International Journal of Systematic and Evolutionary Microbiology*, *66*(12), 5077–5082. https://doi.org/10.1099/ijsem.0.001473

Krabbenhoft, K. L., Anderson, A. W., & Elliker, P. R. (1965). Ecology of Micrococcus *radiodurans*. In *Ecology of Micrococcus radiodurans. Appl. Microbiol* (6; Vol. 13, pp. 1030–1037). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1058391/pdf/applmicro00362-0200.pdf

Lee, J. H., Jung, J.-H., Kim, M.-K., & Lim, S. (2022). *Deinococcus taeanensis* sp. Nov., a Radiation-Resistant Bacterium Isolated from a Coastal Dune. *Current Microbiology*, *79*(11), 334. https://doi.org/10.1007/s00284-022-03044-8

Mohseni, M., Abbaszadeh, J., & Nasrollahi Omran, A. (2014). Radiation resistant of native *Deinococcus* spp. Isolated from the Lout desert of Iran "the hottest place on Earth". *International Journal of Environmental Science and Technology*, *11*(7), 1939–1946. https://doi.org/10.1007/s13762-014-0643-7

Rainey, F. A., Ray, K., Ferreira, M., Gatz, B. Z., Nobre, M. F., Bagaley, D., Rash, B. A., Park, M. J., Earl, A. M., Shank, N. C., Small, A. M., Henk, M. C., Battista, J. R., Kämpfer, P., & Da Costa, M. S. (2005). Extensive diversity of ionising-radiation-resistant bacteria recovered from Sonoran Desert soil and description of nine new species of the genus *Deinococcus* obtained from a single soil sample. *Applied and Environmental Microbiology*, *71*(9), 5225–5235. https://doi.org/10.1128/AEM.71.9.5225-5235.2005

Tettelin, H., Riley, D., Cattuto, C., & Medini, D. (2008). Comparative genomics: The bacterial pangenome. *Current Opinion in Microbiology*, *11*(5), 472–477. https://doi.org/10.1016/j.mib.2008.09.006

Wang, X. P., Li, C.-M., Yu, Y., Li, H.-R., Du, Z.-J., & Mu, D. (2019). *Deinococcus* arcticus sp. Nov., isolated from Silene acaulis rhizosphere soil of the Arctic tundra. *International Journal of Systematic and Evolutionary Microbiology*, *69*(11), 3437–3442. https://doi.org/10.1099/ijsem.0.003636